A Comparison of Mobile Search Interfaces for isiXhosa Speakers

By

MOREBODI MODISE

Department of Computer Science
UNIVERSITY OF CAPE TOWN

A dissertation submitted to the University of Cape Town in accordance with the requirements of the degree of MASTER OF SCIENCE IN COMPUTER SCIENCE BY COURSEWORK AND DISSERTATION in the Department of Computer Science, Faculty of Science.

JANUARY 2018

Supervised By

ASSOCIATE PROFESSOR HUSSEIN SULEMAN

Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

PUBLICATIONS

Part of this research appears in the following IEEE conference publication:

Modise, M. and Suleman, H. (2017) "Mobile Search Interfaces for isiXhosa Speakers: A Comparison Between Voice and Text", in IEEE Information Communications Technology and Society. Umhlanga Durban, South Africa: IEEE, pp. 47-52.

Dedication

To my family:

I dedicate this research to my mother for her love, encouragements and wisdom that builds me on a daily basis. To my brothers who have been with me through tough times; Thusoetsile, Rebaone, and Kagiso. My nephews Omphile and Lebogang for always bringing a smile to my face whenever I took time off writing this thesis. To my son Botlhale, the best is yet to come.

To my friends:

For the encouragement and support guys; Nwabisa, Sharon, Thabang, Ame, Bruce, Bazamile and Vincent.

To my late father:

Thank you for the hard work you have put forth to ensure you raise me to become the man you have always desired me to be. You are greatly missed and forever appreciated.

ACKNOWLEDGEMENTS

First and foremost I would like to thank the Almighty God for giving me the strength and ability to undertake this research. Furthermore, I am grateful to my supervisor Associate Professor Hussein Suleman for his invaluable support, advice and guidance from the beginning to the end of this research. I am grateful for the time you were always willing to spare to discuss my research with me. Thank you for the patience and kindness you have shown me through all this time. I sometimes look back on my research journey and realize I would have not achieved this much if it was not of your very valuable input. Much appreciation to Dr Melissa Densmore, Dr Brian DeRenzi, Dr Maria Keet and Associate Professor Michelle Kuttel for sparing a few minutes to answer questions about my research. Special thanks to Vaal Reefs Disaster Trust Fund for the funding for the project.

I extend my appreciation to the University of Cape Town for allowing me to access its students. Additionally, I wish to appreciate all the students who took part in this study; thank you for your support. I would also like to thank the Language Resource Management Agency (LRMA) for allowing me access to their research and findings on speech systems in South African languages. To my colleagues in the ICT4D laboratory, thank you for your valuable contributions, in particular, Josiah Chavula, Blessing Ojeme, Lighton Phiri, Thomas Reitmaier, Dr Charity Chao Mbongo, Sarah Brittan, Jorgina Paihama, Fiona Ssozi, Haji Ali, Job King'ori Maina, Richard Maliwatu, Yamiko Msosa, Maletsabisa Molapo, Bhavana Harrilal and Olojede Ayodeji. I am also thankful for the DL reading group for the amazing academic discussions.

Special thanks to Nwabisa Mfaxa who worked with me tirelessly to develop the isiXhosa voice recognition system. I appreciate your time and suggestions and, more importantly, the translations and tutoring on the isiXhosa language. I also thank Dr Tessa Dowling for providing material on isiXhosa language vocabulary. I am also thankful to Barbara Fourie for assisting in designing the graphics used in building the mobile prototype. I am also grateful to Zola Mahlaza for his valuable contributions to the design of the mobile prototype. My heartfelt appreciation also goes to my brothers Thusoetsile Modise and Kago Kebotsamang for reviewing and correcting sections of this thesis.

Abstract

Search interfaces have for a long time been targeted at the resource-rich languages, such as English. There has been little effort to support African (Bantu) languages in search interfaces when compared to languages such as English, particularly the isiXhosa language. However, due to the increase in use of mobile phones in developing countries, these interfaces can now be adapted to languages in these settings to support information access on the Web. This study proposes mobile search interfaces to support isiXhosa speakers to search for information on the Web using isiXhosa as a discovery language. The isiXhosa language is considered a low-resourced African (Bantu) language spoken in resource-constrained environments in South Africa. The language is spoken by over eight million people. Yet, there has been no search interface specifically targeted at supporting isiXhosa speakers.

Two mobile search interfaces were developed on an Android application. The interfaces were text based and voice based. The design of the interfaces was based on feedback from 4 native isiXhosa speakers in a design focus group, and guidelines from the literature. Using the developed interfaces, an experiment was conducted with 34 native isiXhosa speaking students at the University of Cape Town, South Africa. This was done to investigate, which interface could better support isiXhosa speakers to search for information on the Web using mobile phones.

Quantitative data was collected using application log files. Additionally, user feedback was then obtained using the standard Software Usability Measurement Inventory (SUMI) instrument, and both interfaces were confirmed as usable. In contrast to what was expected, users preferred the text interface in general, and according to most SUMI subscales. This could be because of greater familiarity with text search interfaces or because of the relative scarcity of voice interfaces in African (Bantu) languages. Where users are not literate, the voice interface may be the only option, so the fact that it was deemed usable is an important independent finding. Search in African (Bantu) language collections is still a largely unexplored field, and more work needs to be done on the interfaces as the algorithms and collections are developed in parallel.

Table of Contents

List of Tables

Table **Page**

LIST OF FIGURES

List of Abbreviations

AM	Acoustic Model
ASR	Automatic Speech Recognition
AST	African Speech Technology
DL	Digital Library
HCI	Human Computer Interaction
HCIR	Human Computer Information Retrieval
ICT4D	Information and Communication Technologies for Development
IR	Information Retrieval
LIS	Library and Information Science
LM	Language Model
LRMA	Language Resource Management Agency
PC	Personal Computer
PD	Phonetic Dictionary
PM	Pronunciation Model
SUI	Search User Interfaces
SUMI	Software Usability Measurement Inventory
UxD	User Experience Design

CHAPTER 1

INTRODUCTION

The use of mobile phones in developing countries has increased significantly in recent years. This has made it possible for people in these settings to have an alternative means to the desktop computer that enables them to search for information on the Web. Mobile phones can be extended to better address information needs for users in developing countries who search for information on the Web using these devices [10]. However, searching for information on the Web using mobile phones presents several challenges and limitations [18]. These challenges and limitations present an opportunity for further research, especially given that these devices also offer additional benefits compared to their desktop counterparts, such as searching the Web while walking or even driving.

While interacting with information on the Web using mobile phones, information seekers make use of search user interfaces commonly referred to as search interfaces, a notion that will be adopted throughout this study. Most search interfaces today focus on supporting English literate users [8]. As such, they have become unusable to some communities in developing countries where English literacy is considerably low, in spite of accessibility to mobile phones [45]. For example, in most African contexts, users may own a mobile phone and may not be fluent in English but rather in their mother tongue only. It is against this background that this study compares mobile search interfaces for isiXhosa speakers to help them search for information on the Web using isiXhosa as a discovery language. Because isiXhosa speakers reside in a developing country, where it might not always be possible to have access to a Personal Computer (PC), use of mobile phones is proposed as an alternative.

The study focuses primarily on comparing two mobile search interfaces: text and voice. The mobile text interface allows for isiXhosa speakers to search for information on the Web by typing isiXhosa text queries into a mobile phone. The mobile voice interface allows for isiXhosa speakers to search for information on the Web by speaking isiXhosa voice queries into a mobile phone. The study proposes that the mobile text and voice interfaces can be adapted to the requirements of isiXhosa speakers to assist them in searching for isiXhosa content on the Web using their mobile phones.

1.1 Research Scope

The scope of this study is on comparing mobile search interfaces for isiXhosa speakers that have been developed following a 4-step iterative process, and not the development of speech technology. The isiXhosa speakers make up 16% of South Africa's total population, with more than eight million first language speakers [61]. Despite this high figure, few Information Retrieval (IR) systems have been built to support them. This study is a part of a bigger project that aims to develop IR systems, algorithms and tools to help improve access to isiXhosa content on the Web.

There were two mobile search interfaces developed based on text and voice - on the Android platform. The Android platform was chosen as this is the most used phone operating system amongst the isiXhosa speaking communities. Additionally, the mobile text interface was chosen to support text-based information retrieval, the most popular technology in IR systems today [38].

The increase in presence of multimedia content on the Web, and the increase in number of smart-phones with enhanced processing power has resulted in the necessity for voice-based information retrieval, which is increasingly becoming popular [38]. For this reason, the mobile voice interface was chosen.

1.2 Problem Statement

A lot of research on search interfaces has been predominantly focused on designing these interfaces to support English literate users [70]. There has been little effort to support African (Bantu) languages when compared to the English language, especially the isiXhosa language, which is a low resourced language spoken in resource-scarce environments [16][49]. This raises a concern as the novice and non-English literate isiXhosa speakers are unable to make use of current interfaces for the purpose of acquiring critical information that involves their health, environment and day to day decision making processes. The perspective presented in this study is based mainly on the following assumptions:

a) African languages: today's search interfaces have little to no support for African (Bantu) languages, particularly isiXhosa. Even in cases where the language support is offered, it is offered in part as a user submits a query in one language and receives the results in another language [75].

b) Human intermediary: it is almost impossible for a new user to use the current search interfaces without the need of a human intermediary to aid the information search process [7]. This study argues that, due to this claim, novice and non-English literate isiXhosa speakers are unable to use these interfaces when searching for information on their own.

1.3 Research Questions

The study aimed to answer the following research questions:

> *a) Is a mobile text interface more effective than a non-mobile text interface for isiXhosa speakers when searching for information on the Web?*

This research question compares the mobile text interface against the non-mobile baseline desktop interface to determine the most effective interface. The effectiveness metric is discussed in Chapter 4 and is made up of three components, namely: satisfaction, speed and relevance.

> *b) Is a mobile voice interface more effective than a non-mobile text interface for isiXhosa speakers when searching for information on the Web?*

This research question compares the mobile voice interface against the non-mobile baseline desktop interface to determine the most effective interface between the two. Similarly, the effectiveness metric is made up of three components, namely: satisfaction, speed and relevance.

> *c) Is a mobile voice interface preferred over a mobile text interface for isiXhosa speakers when searching for information on the Web?*

This research question compares the mobile voice interface against the mobile text interface to determine the most preferred interface by isiXhosa speakers.

1.4 Research Methodology

The comparison of mobile search interfaces for isiXhosa speakers could only be possible if such interfaces existed, as such, the first step in the study was to design and develop a proof-of-concept mobile application with the two interfaces. The study followed an iterative 4-step design approach to build the mobile application prototype with the two mobile interfaces. The design step involved designing and developing each interface according to the design guidelines suggested during a design focus group by 4 native isiXhosa speakers, and design guidelines recommended in the reviewed literature. The next step - testing - focused on the developer testing the usability of each interface, thereafter passing it on to fellow researchers in the ICT4D laboratory for peer review. As the testing phase was underway, design and feature suggestions were made and, where possible, they were incorporated into the design.

The refinement step involved revisiting the design of the interfaces based on design modifications suggested during the testing phase. This resulted in efforts to modify key features to make them as close as possible to the general requirements of search interfaces [24]. In the evaluation stage, target users played a vital role in ensuring that the interfaces were well suited to isiXhosa speakers. At the end of the evaluation, the design process was repeated, including new features. For each mobile interface, the 4 steps discussed above were repeated 3 times.

The desktop interface, that was also the baseline, was initially developed separately from this study. The interface was built as part of an unpublished isiXhosa search engine project that focused on developing a Web crawler that could collect and index isiXhosa documents on the Internet. The Web crawler distinguishes isiXhosa documents from other languages through a Language Model [35]. The prototype with the two mobile interfaces and the baseline interface was used in a pilot study with 3 participants who were native isiXhosa speakers to verify the practicality of the experimental protocol for this study. Thereafter, the interfaces were evaluated in the main experiments by 34 native isiXhosa speakers.

In order to address the first research question for the study, qualitative feedback was collected using Software Usability Measurement Inventory (SUMI) questionnaires. SUMI questionnaires are made up of 50 usability questions that have 3 options the respondent had to choose from. These questionnaires are developed by the SUMI company and are provided free of charge for academic use. The qualitative feedback data was collected after the participants evaluated the mobile text interface and the baseline desktop interface. The data collected was analyzed by SUMI to determine the most effective interface between the two interfaces. In addition to this, quantitative data for speed and relevance for the two interfaces was collected. Speed data was collected to determine the total amount of time the participants took evaluating each of the interfaces. Clicks-through data was collected to measure relevance feedback of each interface. Speed and relevance data was monitored through application log files while participants interacted with the two interfaces. The analyzed SUMI results from the collected data were then statistically tested to determine the most effective interface between the mobile text interface and the baseline desktop interface.

In order to address the second research question, qualitative feedback was also collected using SUMI questionnaires. These questionnaires were similar to those used to answer the first research question. However, the two questionnaires were separated by a uniquely identifiable password that was provided by SUMI. The qualitative data used to address the second research question was collected together with quantitative data for speed and relevance of the mobile voice interface. Speed captured the amount of time the participants spent while evaluating the mobile voice interface while relevance was determined by the number of clicks-throughs on the voice results. Speed and relevance were monitored through application log files while participants interacted with the interfaces. The analyzed SUMI results from the collected data were then statistically tested to determine the most effective interface between the mobile voice interface and the baseline desktop interface.

In order to address the third research question of the study, qualitative data from the SUMI questionnaires discussed above was used to compare the mobile text interface to the mobile voice interface. The analyzed results from the collected data were statistically compared to determine the most preferred mobile interface. Therefore, to address the first two research questions of the study, both qualitative and quantitative data were used. To address the third research for the study, only the qualitative data was used.

1.5 Research Contributions

It was anticipated that, by answering the research questions for this study, the following research contributions would be made:

1. A proof-of-concept mobile application prototype, which isiXhosa speakers can use to search for information on the Web using isiXhosa as a discovery language.

2. Usable mobile voice and text interfaces that could support isiXhosa speakers whose English literacy is considerably low.

3. Mobile search interfaces that contribute to search in African (Bantu) languages.

1.6 Thesis Structure

The thesis is arranged into six chapters as outlined below:

Chapter 2: Background and Related Work

This chapter discusses previous research that focuses on search interfaces, in particular, mobile search interfaces. The chapter is divided into sections, namely: definition of search interfaces and previous research on search interfaces; benefits and challenges of the transition from desktop to mobile devices as primary information access tools; benefits and challenges of mobile text and voice interfaces in supporting users in the search process; and guidelines for designing mobile search interfaces.

Chapter 3: Design and Implementation

The design of a mobile application prototype that has the mobile text and voice interfaces is discussed in this chapter. The design process of the two mobile interfaces followed an iterative design approach. How the iterative design was applied to the design and development of the mobile text interface is outlined. Thereafter, the iterative design process that was used to design and develop the mobile voice interface is discussed. The baseline desktop interface is briefly discussed as well.

Chapter 4: Experiment Design and Evaluation Methods

The chapter focuses on how the evaluation was conducted to address the three research questions for the study. First, the participants who took part in the study are described, as well as the sampling strategy used to select them. The data collection tools used to collect data while participants evaluated the interfaces is discussed. Further, how the experiments were conducted is discussed, with emphasis on the environment the experiments were conducted in, the search tasks used and the pilot study conducted prior to the final experiments. Additionally, the evaluation criteria used to address the research questions for the study is discussed. The chapter ends with a summary of the evaluation process.

Chapter 5: Results and Discussion

In this chapter, the results are presented and discussed for each interface. The chapter begins by describing the participants' involvement in the experiments. Thereafter, the results of the pilot study are discussed as per the three interfaces, namely: mobile text interface; mobile voice interface; and the baseline desktop interface. Subsequently, the final experiments' results are discussed as per the three interfaces. A statistical comparison of the three interfaces is then presented and discussed. The chapter ends with a summary of the research findings.

Chapter 6: Conclusion

This chapter starts by revisiting the research problem and the research questions for the study. Thereafter, how the results from the experiments addressed the research questions for the study is discussed. This is followed by a discussion of implications of the study and future research recommendations.

CHAPTER 2

BACKGROUND AND RELATED WORK

Mobile search interfaces have been extended to developing countries due to the increase in use of mobile phones in these areas. However, use of these interfaces in these settings is faced by many challenges. It is against this background that this chapter discusses search interface definitions in Section 2.1.1 as perceived by experts in Information Retrieval (IR). Section 2.1.2 discusses trends in search interfaces. In Section 2.2, the isiXhosa search engine, which forms the foundation for this study is discussed. Section 2.3 reflects on the transition of designing search interfaces from desktop computers to mobile phones by discussing the transition benefits in Section 2.3.1, and the challenges in Section 2.3.2. Section 2.4 focuses on mobile text interfaces and, particularly, research that has aimed to better support users to search for information on the Web using mobile phones. In like manner, Section 2.5 looks into mobile voice interfaces. In Section 2.6 the different design approaches aimed at designing mobile search interfaces to better support users are explored. Section 2.7 concludes the chapter by discussing limitations and opportunities that have been identified in the background and related work.

2.1 Search Interfaces

2.1.1 Defining Search Interfaces

The field of search interfaces is multidisciplinary and has over the years received a lot of attention from various researchers of different communities [70]. These communities are mainly interested in how users make use of information as retrieved from IR systems. The different communities include but are not limited to Human Computer Interaction (HCI), User Experience Design (UxD), IR, Human Computer Information Retrieval (HCIR), and Library and Information Sciences (LIS) communities, as depicted in Figure 2.1. Although there are different communities studying search interfaces, there remains a common understanding of what search interfaces are and what they ought to be, regardless of the mode of access, such as mobile phones and desktop computers. The following definitions derived from different experts in the domain of search interfaces have been adopted by many researchers around the world to share a common understanding of what search interfaces are.

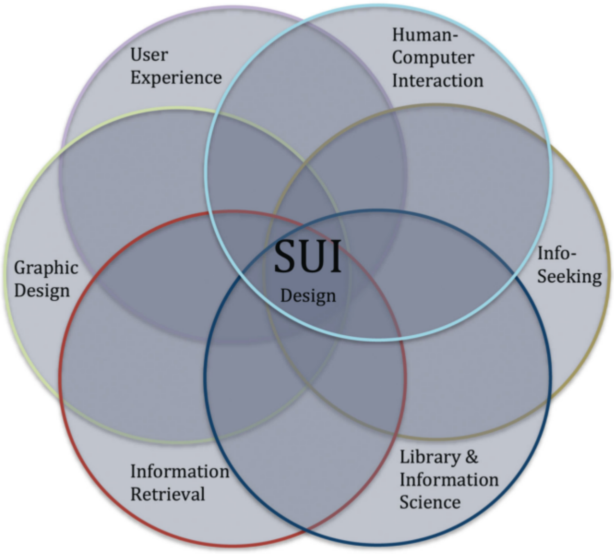


FIGURE 2.1. A screen-shot showing the different disciplines involved in Search User Interface (SUI) design, from Wilson et al. [70]

Search interfaces have been viewed as tools serving the purpose of aiding information seekers to express their information needs, formulate search terms, analyze and make sense of the retrieved results as well as assist them to track their search progress [24]. In her further work, Hearst [25] defines search interfaces as guiding tools that help the searcher to understand and know what information is needed, and how to retrieve and interpret the information once it is available. In like manner, Wilson et al. [70] provides a high level definition by denoting search interfaces as an in-between tool for people who have an information need and the different data sources of the needed information on the Web across the world.

2.1.2 Trends in Search Interfaces

Early studies in search interfaces put a lot of emphasis on text-search interfaces [42][41][27]. Some studies focused on identifying key interface features that are better at supporting end users in the information seeking process [34][42] while others were aimed at reducing the complexity encountered when designing search interfaces, by coming up with guidelines for design [24][41][70]. Later studies experimented on different techniques that could be incorporated in search interfaces to enhance their capabilities. Shiri et al. [59] investigated different IR interfaces that have been enhanced by including a standard thesaurus as part of their searching and browsing facilities.

As information increased on the Internet, IR researchers began to design search interfaces that would actively involve the information seeker in the retrieval process. This field of research is regarded as HCIR [43]. The HCIR paradigm is aimed at empowering information seekers to explore large-scale information bases [44]. Despite the success with text-search interfaces, it is argued that almost all the text facilities can be accomplished using voice [38]. This type of information retrieval is referred to as voice-based information retrieval, where a search query and/or content is retrieved in the form of voice. Voice-based information retrieval has resulted in the design and development of voice-based search interfaces. These interfaces have already gained attention from different communities. Gilbert and Hall [20] investigated spoken query systems by measuring user satisfaction and language model coverage for such systems, while Banu and Ummuhanysifa [67] experimented on voice based search interfaces by developing an interface that allowed for the user to query information by speaking into the computer and thereafter the computer responding to the user in voice.

In an attempt to investigate multiple language support in search engines, Zhang and Lin [75] investigated the different search engines that exist on the Internet to identify those that have multiple language support capabilities. The results of the study show that users who were familiar with one language struggled to interpret information presented to them in a different language than the one they understand. The study marked this as one of the main challenges of multiple language support in search interfaces. Lee and Pan [38] also experimented on two IR search interfaces: text and voice. Their study was able to identify that voice-based information retrieval is lacking in two aspects: retrieval accuracy and user system interaction.

2.2 isiXhosa Search Engine

As part of the isiXhosa Information Retrieval project, Kyeyune [35] developed an isiXhosa search engine that could crawl and index isiXhosa documents on the Internet. The search engine made use of a Language Model to differentiate isiXhosa documents from other language documents on the Web. The developed search engine was deployed with a Web interface for public access at `http://isixhosa.alir.cs.uct.ac.za`. The search engine formed a basis for this study, as the mobile interfaces connected to it to fetch results for search queries.

2.3 From Desktop to Mobile

Desktop computers, for example laptops and PCs, have been for many years the primary means for people to access information on the Web. This influenced how IR systems' interfaces in these devices were designed [70]. However, the recent increase in usage of mobile devices, in particular smart-phones with network capabilities, has presented new ways for people to access information on the Web and in the same fashion influenced the design of IR search interfaces. Mobile devices present a quick means for people to access information from the comfort of their homes, offices and while mobile [55]. Schalkwyk et al. [55] argue that the increase in mobile devices as information access tools creates new information access expectations and needs. Some of these expectations are for constant access to up-to-date information at any time of the day and in any place in the world.

Foreseeing the move from desktop to mobile devices as primary information access tools, Schilit [56] claims that the increase in number of mobile devices will lead to more people accessing information on the Web using mobile devices as primary tools rather than the traditional desktop computers. He further points out that due to the pervasiveness of these devices, they have come to be recognized as utilities, like water and electricity. In complementing Schilit [56], a more recent report has revealed that mobile search interfaces have become popular, with Google announcing that more than half of its searches now happen on mobile devices[1].

2.3.1 Benefits

Searching for information on the Web using mobile devices has presented several benefits, which might have been almost impossible with the traditional desktop computers. The features of these devices, for example, low weight, wireless networking capabilities and smaller sizes allowing them to fit in a pocket hence being accessible at any time and in any place, has given these devices distinguishing characteristics from their desktop counterparts. Amongst the many benefits, Long and Chang [40] discuss the benefit of location based searches. Through mobile devices, users are now able to search for location based information such as the closest restaurant or gas station, which has been for many years impossible to do while mobile. Additionally, the recent developments in mobile technology, such as the ability to support voice input allows for users to use their voice to search for information on the Web while walking or even driving.

Anderson and Rainie [2] argue that the use of voice search on touch based smart-phones will become widespread by the year 2020. According to Schalkwyk et al. [55], this is influenced by the type of services and expectations users of these devices have, which includes the need to access real time information while mobile. The use of voice based search on mobile devices presents several other benefits, notably the possibility of reducing literacy barriers. Addressing literacy barriers in mobile interfaces has been well researched in the HCI community [45]. Although most of the research is not directly specific to search interfaces but rather mobile user interfaces and how they could be designed to best suit novice and low literacy users, to enable them to access the vast amount of services and utilities that are increasingly available to them.

Mobile devices further present an opportunity to extend search systems into the developing world to help people in these settings to search for information on the Web [54]. This alone addresses issues such as the digital divide, which might not have been possible without mobile devices. Rajput and Nanavati [50] investigated different mobile voice solutions that have been developed to target the developing world with the aim to improve both the social and economic landscapes. The results of the study indicate that speech on mobile devices is yet to benefit the developing world economically and socially.

[1] `http://adwords.blogspot.co.za/2015/05/building-for-next-moment.html`

2.3.2 Challenges

Although mobile devices have brought many benefits with them, such as the ones outlined in the above section, they have also brought some challenges. These challenges have affected the design of mobile search interfaces (both text and voice) to ensure they appropriately suit the mobile phone specifications. Amongst the many challenges affecting search interface design on mobile devices is the small display, which limits the amount of information to be displayed at a time, and the small keyboard needed for typing search terms [46]. While interacting with mobile search interfaces, in particular text-search interfaces, users are unlikely to have the efficiency and effectiveness they would have while typing queries in desktop interfaces as the typing keyboard is smaller in size. Similarly, Dunlop and Brewster [18] have argued that although mobile device displays may increase in color and pixel density to better support end users, the smaller displays on mobile devices will remain a challenge as these devices are designed for portability. They argue that due to the smaller display, input keyboards on mobile devices are limited in the number of keys presented to the user while typing on a mobile device.

In regard to mobile voice interfaces, Dunlop and Brewster [18] further argue that the quality of input mechanisms such as microphone on mobile devices is very poor and thus limits the ability of these devices to recognize spoken queries. Additionally, searching like most mobile operations, relies heavily on the battery life, which has proven to be an issue in modern smart phone technology. Want [69] elaborates on this point by reflecting on the ongoing developments that have been targeted at improving the battery life of most smart phones, which to date has proven not sufficient. In order to successfully deliver the promise of searching entirely on the mobile phone, smart phones should be able to sustain enough power to support both text and voice searches.

2.4 Mobile Text Interfaces

Mobile text interfaces are by far the most common interfaces for people accessing information on the Web through IR systems such as search engines [38]. They usually involve the user typing in the search term and getting the results in text. This section of the chapter puts emphasis on research that has been done in relation to text interfaces on mobile devices.

2.4.1 Language Support

In an effort to support access to non-English languages' informational sources that are rapidly increasing on the Web, mobile text interfaces have been adapted to other languages, including the low resourced languages [8]. Hattab et al. [23] present an Arabic search engine called Addaall, which uses different levels of Arabic morphological knowledge to enhance the quality of the search engine. The results of the study have shown promise that the study's approach is practical and it offers a significant improvement to the search engine, therefore its interface.

Ricardo et al. [51] describe the need to comprehend what people around the world search for in their mobile devices so as to design the best interface to support them. Presenting a study on how mobile devices are used to search for information on the Web by the Japanese, Ricardo et al. [51] investigated the usage of mobile devices in searching for information in Japanese by collecting sample log data from Yahoo!. The collected data was from both the mobile and desktop query logs as provided by Yahoo! in 2006. The results of the study show that mobile queries are shorter than desktop queries, which is mainly due to the text-input restrictions caused by the small display on mobile devices. Additionally, they discovered that most unique search terms were made up of two terms in both the desktop and mobile interfaces.

2.4.2 Addressing Small Display

In addressing the challenges presented by the small displays of mobile devices, Ziefle [76] studied the increase in use of mobile devices by aged users, through studying information presentation on mobile devices. He identifies two important factors to consider when designing text-based interfaces for displaying information on mobile devices, namely font size and the size of the preview. He argues that when font size and size of the preview are larger, navigational performance on mobile devices increases. In a similar study, Sweeney and Crestani [64] investigated the effects the retrieved results' summary length has on different screen sizes of mobile devices. Their study aimed to find out whether screen size has any particular effect in presenting search results, and whether there is an optimal summary size to be taken into consideration for a given mobile device screen. The results of their study show that the search results summary size should not be chosen based on screen size, however, screen size plays a major role in designing presentation of personalized results as determined by user needs.

Jones et al. [30] investigated the impact of usability on small display devices during the search process. The results of the study show that while interacting with small displays, users are 50% less effective compared to when using the larger screen display. Additionally, the study has shown that, in small displays, users tend to scroll often in order to complete a search task. Based on these findings, the study suggests a set of design guidelines for improved user interaction on small displays such as a) place important information at the top of the results interface b) reduce the amount of information that can be viewed by users at a given time and c) fix the navigation menu at the top of the interface.

2.4.3 Location Based Search

Mobile text interfaces are by far the most commonly used tools for searching for information that is based on the user's location while mobile. This is influenced by the fact that most users are always carrying their mobile devices with them [28].

Liu et al. [39] carried out two experiments to investigate use of mobile devices to search for information based on a particular location; an airport. The first experiment focused on the user's search performance while searching for information on different topics under different user contexts. The results of the experiment show that the relevance of the first result was highly considered, the number of results clicked decreased and the ability to recall information was improved in comparison to searching for non-location based information. The second experiment measured the user's search performance in differing levels of information requirement pressure. The results of the experiment show that users who searched under minimal information pressure clicked and viewed more search results than those with high information requirement pressure.

Google and Nielsen [22] companies reported on an analysis of search queries submitted to Google from mobile devices and found that 77% of these queries happen while users are at home or working while 17% happen while users are mobile. Interesting enough, the 77% of the queries that happen at home or at work are done by users who have a PC available to them. Additionally, they point out that searching for information on mobile devices is strongly associated with the user's context, for example, shopping queries are more likely to be taking place in stores. Furthermore, users prefer fast and convenient ways of searching for information, hence turn to mobile interfaces for searching the Web. This is confirmed by the results of the study, which reflects that 81% of searches on mobile devices are fuelled by speed and convenience.

2.5 Mobile Voice Interfaces

Voice search systems and their interfaces have been in existence for a long time, with their early versions focusing on business services such as Directory Assistance [55]. Voice interfaces on mobile devices, as we know them today, became more popular in 2008 when iPhone and Android phones introduced a new way for people to search for information on the Web by speaking queries into their phones [55]. According to Mishra et al. [47], mobile voice interfaces present an alternative means to unnatural and tedious typing as found on mobile text interfaces due to smaller typing keyboards.

Voice interfaces follow the architecture outlined in the Figure 2.2. According to Wang et al. [68], a user speaks a search query into the voice system that is then recognized by the Automatic Speech Recognizer (ASR) using the Language Model (LM), Pronunciation Model (PM) and Acoustic Model (AM). Thereafter, the results are passed on from the ASR to the search component to find a list of search entries in the database. These entries are then passed on to a Dialog Manager, which decides how best to present the results. The Disambiguation Module is used in the case where there has not been a match, where the system prompts the user to speak again.

2.5.1 Language Support

Mobile voice interfaces have also been extended beyond the resource-rich languages like English to other languages in both the developed and developing countries. Schuster and Nakajima [57]

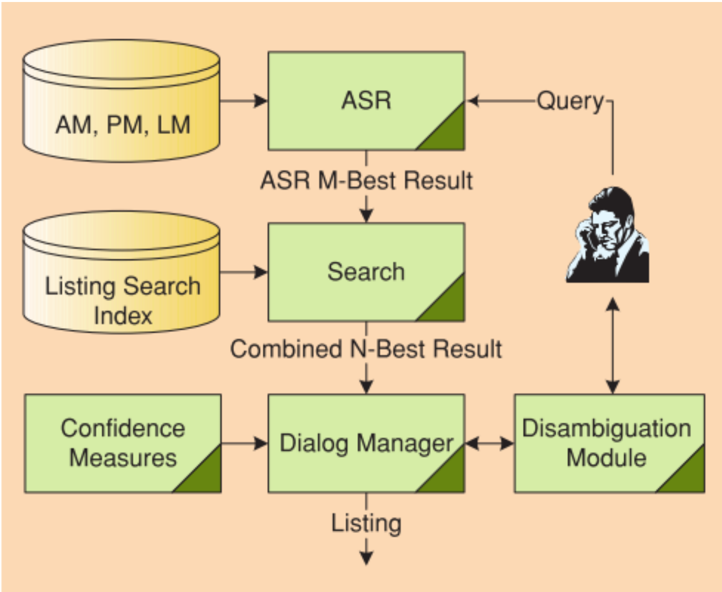


FIGURE 2.2. Voice interfaces architecture, from Wang et al. [68]

highlight problems and solutions they encountered while building the mobile voice search system for the Japanese and Korean languages. One of the challenges they faced was adapting words, numbers and symbols from the English language to the Japanese and Korean languages, as these languages have complicated character sets. The results of the study have shown that they were successful in building a mobile voice search system for the Japanese and Korean languages at Google. Based on their results, they proposed a simplified standard for building mobile voice search systems in other international languages. Similarly, Shan et al. [58] carried out a detailed investigation into searching on a mobile phone using voice in Mandarin Chinese. They detail the process of collecting audio data for building language, pronunciation and acoustic models. The results of the study reflect a successful voice search mobile application, which has been well-received by end users and performs well with different user accents.

Sung et al. [63] detail an attempt to build a mobile voice search system for the Cantonese language, a language spoken in the southern parts of China. They found out that most users find it ideal to speak rather than type out search queries when searching for information on the Web. This is because the Cantonese characters are difficult to type out on a mobile phones' smaller keyboard. In associating voice search with the potential of improving lives in the developing world, Barnard et al. [6] investigated the impact of speech technology, in particular voice-search, in improving the presence of speech technology in the developing world. They argue that speech technology is necessary to help address some of the challenges faced by people in the developing world, such as illiteracy and the ability to access information. The results of the study highlight a successful journey in building a mobile voice search system in two of South Africa's eleven official languages.

Heerden et al. [26] investigate the performance of a multilingual directory enquiry (DE) application operating in four of South Africa's official languages. The DE application was then subjected to usability testing, both informally and formally by first language speakers, to test whether the system was usable to a wide variety of users. The results of the study show that although the application has some limitations most users found the application to be easy and simple to use.

2.5.2 Miscellaneous Applications

Trippas et al. [66] developed an Interactive Information Retrieval (IIR) system where all input and output is strictly voice based. The system presents the user with the ability to express an information need as well as acquire results based on the expressed information need and, to finally, refine voice search results. The results of their study show that varying design methodologies from different fields are still needed to have a more interactive search in voice based systems in mobile devices.

2.5.3 Query Refinement

According to Allauzen et al. [1], query refinement has been mostly for text-based systems and not for voice-based systems. In discussing voice query refinement, Alluazen et al. [1] investigated a system with a refinement model that could refine voice queries by updating a given voice query with a shorter query rather than giving a completely new query. In this system, a user initially speaks a query such as "North Italian restaurants in New York" and later on refines it by speaking a shorter query such as "Korean instead". The results of the study show that the system outperforms the language model-based system and the three human judges who were used as a baseline as there were no systems for spoken query refinement.

2.5.4 isiXhosa Speech Systems

A voice search interface is dependent on the Automatic Speech Recognition (ASR) system to recognize the user's spoken query, as outlined in Section 2.5. For this reason, this section of the chapter focuses on discussing developments in ASR in isiXhosa, an important component considered when building the mobile voice interface for this study. However, this is done with care noting that the focus of the study is not on ASR systems nor is it on Speech browsers but rather on mobile search interfaces and how these can be adapted to help isiXhosa speaking communities to access information on the Web using their mobile devices. With this in mind, research efforts have been made to build ASR systems in isiXhosa and these are discussed below.

Roux et al. [52] discuss the African Speech Technology (AST) project, which was aimed at developing an annotated speech database for isiXhosa and 4 other South African languages. The database was then used to train and develop a speech recognition engine used in a multilingual telephone-based hotel booking system. The study's findings show that isiXhosa users responded with long sentences and in noisy environments, which thus affected the ability of the recognition engine to recognize words and, as such, resulted in recognition errors and poor system performance compared to other language groups. In like manner, Tait et al. [65] investigated the architecture and performance of 3 spoken dialogue systems operating in isiXhosa and 2 other South African languages. The systems were then evaluated by a group of 3 users from the 3 different language groups. The results of the study show that spoken dialogue systems in isiXhosa are feasible. With regards to performance, it was observed that isiXhosa users' interaction responses with the system were longer and more varied when compared to the other language groups. Additionally, the isiXhosa users switched to English for numerical information like times, numbers, amounts and dates.

Collecting acoustic data for languages spoken in the developing context for ASR purposes has been viewed as a challenging task [17]. In an attempt to resolve this, De Vries et al. [17] developed a mobile application, Woefzela, which they used to collect acoustic data for the 11 South African languages including isiXhosa. Woefzela works without the need for Internet connectivity and enables users to record audio on multiple sessions at the same time thus suitable for the developing contexts [17]. Furthermore, it performs basic quality checks on the data while recording is ongoing hence ensuring collection of good quality data. Using Woefzela, De Vries et al. collected 798 hours of acoustic data for all the 11 South African languages, of which, 76.48 hours was for isiXhosa. In a similar study, Badenhorst et al. [4] discuss the Lwazi corpus for ASR for isiXhosa and 8 other Southern Bantu languages spoken within South Africa. The results of the study show that due to practical constraints, little speech data was collected (1 to 10 hours of data) when compared to major corpora in world languages. Furthermore, the study reports on results obtained from phone recognisers that were built from the collected speech data. For isiXhosa, phone recognisers had an accuracy of 57.24% for 210 isiXhosa speakers.

Davel and Martirosian [16] highlight a design process they followed to develop the pronunciation dictionaries for isiXhosa and 10 other South African languages. They then used the developed pronunciation dictionaries to build ASR systems for the 11 South African languages. The ASR systems had 66% to 76% of phone recognition accuracies. In the same way, Barnard et al. [5] report on the National Centre for Human Language Technology (NCHLT) project, which was aimed at collecting speech corpora for each of the 11 official languages in South Africa including isiXhosa [5]. The NCHLT project expands on the Lwazi and AST projects discussed above. The NCHLT aimed to collect more than 50 hours of speech data for each of the 11 official languages of South Africa. The results of the study show that more than 50 hours of speech data was collected and has been released for public access. According to Barnard et al. [5], the data is available as a collection of audio wave files together with their corresponding transcriptions in XML format.

2.5.5 isiXhosa Translation Systems

Some efforts have been made to translate English content into isiXhosa to support native speakers on online systems. The following section discusses some projects focused at translating content into isiXhosa to help increase the language presence on the Internet even though this is not the focus of the study.

Dalvit et al. [15] present Pootle, a Web application used for collaborative translation online. The application was deployed and used by native isiXhosa students at two South African universities. The results of the study indicate that freely-available open-source software can be localised in any language in a cheap manner. Also, the results of the study show that the enthusiasm and wiliness of native isiXhosa students to use the application demonstrated a need by them to use their mother-tongue in ICT systems. Similarly, Google South Africa has recently announced support for isiXhosa translation on their platform [2]. The translation is in its early stages as some quick and informal tests of everyday language use yielded incorrect results at the time of testing, for example, *"molweni"* translated into "in the law" instead of formal plural greeting such as "Good day" or "Good morning" or "Good evening".

2.6 Mobile Interface Design

Designing mobile search interfaces requires a careful thought process, which is different from when designing for desktop interfaces, as replicating the design style can lead to cluttered and frustrating interfaces that users may find not useful [53]. In addition to this, Russell-Rose and Tate [53] discuss the three main design principles that should be adhered to when designing mobile search interfaces to ensure the best search experience, namely: content should be prioritized over controls; the interface should provide answers rather than results; and cross-channel continuity.

With regard to voice interface design on mobile devices, Yankelovich and Lai [73] recommend involving users in the design process to ensure a well planned conversational style, like interaction with the interface as it might happen naturally between two people. Additionally, they encourage that a voice interface provides a means to resolve an error should it be encountered by the user or caused by the system itself. In a more practical approach, Yankelovich et al. [74] further highlight the need to develop speech interfaces from scratch rather than copying them from their textual counterparts. They outline the importance of informative feedback to keep the user updated with their interactions with the interface. Suhm [62] adopts a methodology to bring together the different design guidelines for different speech interfaces. In his study, he identifies ten design guidelines, which enables practitioners to utilize usability engineering methodologies related to speech interfaces, such as heuristic evaluation and early consultation in the design and planning

[2] `http://www.itweb.co.za/index.php?option=com_content&view=article&id=162936`

phase, prior to engaging in the actual design process. He further outlines the need to ensure that the short term memory is not overloaded in cases of having large query results.

Presenting minimal options per interaction as well as providing confirmations at each stage of interaction are good means to reduce overloading the working memory for users of speech interfaces [71]. Detailing the general search interface design process, Shneiderman et al. [60] propose a four phase framework to think about when designing search interfaces. The framework consists of four components, namely: formulation, action, review of results and refinement. The formulation phase involves considering all the decisions that need to be taken by the user before initiating the search process. Such decisions may include sources of information to search for and variants of text accepted. The action phase involves the user starting the search process; this can be through pressing a search button to search for information and thereafter pressing it again to get the relevant retrieved results. The reviewing results phase deals with the actual interaction with the results by the user as well as changing the query to get better results.

Hearst [24] draws on user interface design guidelines proposed by Shneiderman et al. [60] to tailor search interface specific guidelines, which include: a) giving users informative and efficient feedback, b) balancing what the user should do with the automated tasks, c) minimizing errors, d) providing users with keyboard shortcuts, e) paying attention to little details, and f) aesthetics design.

2.6.1 Summary of Mobile Interface Design Principles

Designing mobile interfaces for search has been viewed as a process that needs careful planning. There are some general principles that ought to be followed when designing interfaces for search on mobile devices, as discussed in Section 2.6. For example, search interfaces must prioritise content over controls and minimise room for errors. However, there are also unique and specific principles that are dictated by the mode of search, that is, voice or text. In voice search, the interface needs to be designed to enable conversational style like interaction, as it might happen naturally between two people. Further, it is recommended to develop such interfaces from scratch rather than copying them from their textual counterparts.

The study has adopted design principles from the reviewed literature to aid the design process of the proposed interfaces. For the voice interface, the following specific principles are adopted:

- Involve users in the design process to ensure a well planned conversational style, like interaction with the interface as it might happen naturally between two people [73].
- Develop the interface from scratch rather than copying it from its textual counterpart [74].

For both interfaces, the following principles are adopted:

- Provide a means to resolve an error should it be encountered by the user or caused by the interface itself [73].

- Do not overload the user's short term memory [62].
- Place important information at the top of the results interface [30].
- Pay attention to little details [24].
- Give users informative and efficient feedback [24].
- Reduce the amount of information that can be viewed by users at a given time [30].

2.7 Chapter Summary

Search interfaces have been defined as an in-between tool for people who have an information need and the different data sources of the needed information on the Web across the world. To date, search interfaces have been researched by different communities such as HCI, HCIR, UxD, IR and LIS. Early studies in search interfaces focused a lot on text interfaces. This was influenced by the type of devices used to search for information, which were mainly desktop computers. However the increase in use of mobile phones, especially smart-phones with Internet capabilities saw the shift from desktop to mobile phones as primary information access tools.

By their nature, mobile phones offer several benefits compared to their desktop counterparts. These benefits include the ability for people to search and access information while driving or walking, and for people to search for information based on their current location. The most important of these benefits, however, is the opportunity of extending mobile phone functionality, in particular the ability to search for information on the Web, to the developing countries. This presents opportunities such as developing mobile search interfaces that support searching for information in one's native language thus breaking down the digital divide caused by language barriers.

Mobile text interfaces are the most common way for people to access information on the Web. Similarly, mobile voice interfaces have been in existence for a long time and have overtime improved. Both text and voice interfaces have been extended to support languages other than the resource-rich English language. Despite this not being the focus of the study, projects such as AST, Lwazi and NCHLT in South Africa were focused on building ASR systems for isiXhosa and 10 other official languages in South Africa. Furthermore, there has been some translation work done to translate English content into isiXhosa so as to improve the presence of isiXhosa on online systems. Nonetheless, the successful development of the above projects does not mean the presence of isiXhosa on the Internet has become the norm: research work that utilizes the produced ASR speech corpora to build ASR systems, particularly IR systems in isiXhosa is still needed.

Designing mobile search interfaces for use on a mobile phone differs from the desktop counterparts. There are several distinguishing characters of mobile phones that make them different from the desktop computer, therefore influencing the design of mobile search interfaces.

The background and related work has discussed some gaps and opportunities that have motivated the following opinions:

1. Most of today's search interfaces are mainly targeted at English literate users and thus become unusable to most people in developing countries, where English literacy is considerably low.

2. The use of complete voice search interfaces in isiXhosa is lacking, that is, querying information using voice and retrieving the results in voice as well and not in text.

3. The widespread of mobile phones in developing countries presents an opportunity to extend search interfaces to communities in these settings.

The design and implementation of the mobile search interfaces that this study proposes is discussed in the next chapter, taking into consideration the above highlighted points.

Chapter 3

Design and Implementation

The focus of this study was to compare mobile search interfaces that have been designed and developed specifically for isiXhosa speakers to help them search for information on the Web. As such, the first step in the design process involved designing the mobile interfaces based on isiXhosa speakers' needs. To successfully achieve this, an invitation to participate in a design focus group was sent out through Facebook, word of mouth, and emails to native isiXhosa speaking students across the University of Cape Town, South Africa. The design focus group resulted in several design contributions that aided the design and implementation of the two mobile interfaces used in this study. The first interface was a mobile text interface, which allows for isiXhosa speakers to search for information on the Web by typing text queries on a mobile phone. The second interface was a mobile voice interface, which allows for isiXhosa speakers to search for information on the Web by speaking queries into a mobile phone.

The most distinguishing feature about the two interfaces is that the main communication language in all interfaces is isiXhosa. The isiXhosa language is a click language, and is categorized as a low resourced language spoken in resource-scarce environments [16]. Therefore, developing an IR system for such a language as isiXhosa may present an IR researcher with challenges unique to working with low-resourced languages in general, as discussed in Chapter 2. Additionally, there are few Web resources available to provide data needed to build such IR systems. In the case that there is enough data available, researchers are usually presented with unique challenges, which include complicated phonetic structure of these languages, and little research published on these languages [16].

It against this background that this chapter reflects on the process of designing, developing and implementing the two mobile interfaces on the Android platform, presenting examples in the form of figures, challenges encountered, solutions and hopefully a road-map for similar future projects. The chapter begins with Section 3.1, which discusses the design approach followed when designing the two mobile interfaces. Section 3.2 covers the design process followed while developing the mobile text interface while Section 3.3 details the design and implementation of the isiXhosa speech recognition system, which complements the voice interface. Section 3.4

gives details on the desktop interface that was used as a baseline for this study, followed by a concluding Section 3.5 that summarizes the chapter.

3.1 Design Process

The design process of the two mobile interfaces followed an iterative 4-step design approach, as shown in Figure 3.1 [48]. The first step involved designing and developing each interface as close as possible to the design guidelines raised during the design focus group. The next step involved the developer informally testing the usability of each interface, thereafter passing it to fellow researchers in the ICT4D laboratory to do the same. As the testing phase was underway, design and feature suggestions were made, and where possible they were incorporated into the design.

3.1.1 Design Principles

The interfaces were developed with the aim of incorporating these design requirements, as further discussed in the next section.

The refinement stage involved revisiting the design of the interfaces based on design modifications suggested during the testing phase. This resulted in modifying key features to make them adhere to the general requirements of search interfaces [24]. In the evaluation stage, target users played a vital role in ensuring that the interfaces were well suited for isiXhosa speakers. At the end of the evaluations, the design process was repeated, including new features and in some cases removing some. The next sections cover the iterative design process as used in developing

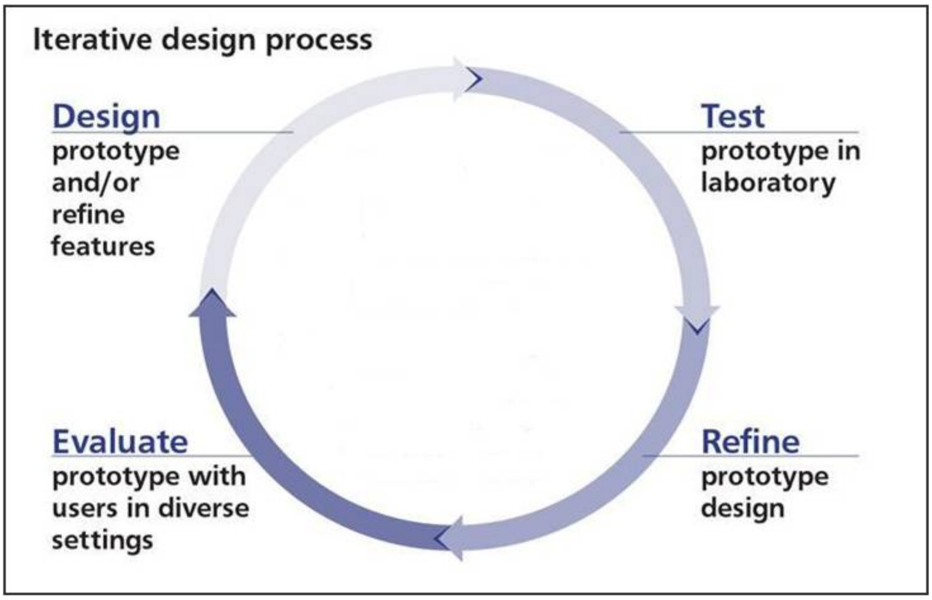


FIGURE 3.1. Iterative design stages as followed to design and develop the mobile text and voice interfaces

both the mobile text and voice interfaces, highlighting key changes, challenges and solutions at each iteration of the design.

3.2 Mobile Text Interface

The mobile text interface was built by the researcher on the Android mobile platform using the Android Studio Integrated Development Environment (IDE). The interface was connecting to a Solr search engine built specifically for searching isiXhosa content on the Web as part of the Computer Science honors project [35]. The interface submitted search queries to the search engine and received text results via Hyper-Text Transfer Protocol (HTTP). The design of the interface was initiated by a design focus group that aimed to involve target users in the design of the interface. There were four participants in the design focus group. Three were from the Computer Science honours class, and they had been trained on interface design and search engines through HCI and IR courses offered within their curriculum. The last participant was a second year student from the Mechatronics Department, and had been trained on search engines and speech systems. Additionally, the four participants had prior experience with search engines such as Google and Yahoo!. Two of the participants had used Google to search for lyrics of isiXhosa songs, one had used Google to search for translations of isiXhosa words and another to search for isiXhosa assignments, mostly in high school.

The participants complained that while using search engines like Google and Yahoo! to search for isiXhosa content on the Web, they do not get the results they are looking for and this needs to be improved. Moreover, they have observed that the isiXhosa language support still remains a challenge in most of today's search engines. This is because they barely get access to isiXhosa content they want even when they are positive that there are websites that have the content. The result of the focus group was a list of design guidelines, which were summarized to the following: a) have a search button with a magnifying glass next to the search box b) auto complete or suggest words as users type c) blink cursor within the search box prior to a search, and give the search box a familiar and clear colour d) include clear simple instructions to begin the search process, and avoid using complex and lengthy isiXhosa words for search instructions e) neatly space out results f) for each result, have blue link and black description of each returned result and g) no adverts.

The participants' design guidelines listed above could be considered as obvious and straightforward. This could have been influenced by participants' past experiences with the already existing search interface designs such as that of Google and Yahoo!. However, given that the objective of the focus group was to gather the essential design guidelines from native isiXhosa speakers to begin the design phase, the above design guidelines together with those identified in the literature (see Section 2.6.1) met such an objective and were used for the initial design iteration.

3.2.1 First Prototype

The first prototype of the mobile text interface focused on delivering a search interface as soon as possible while utilizing the design guidelines outlined above. This resulted in the design shown in Figures 3.2 and 3.3. Figure 3.2 shows the query submission interface, and Figure 3.3 shows the search results interface. As seen from these figures, the interface consisted of a mix of English and isiXhosa instructions. This was done to serve as a guide for native isiXhosa speakers to give correct prompting instructions in isiXhosa during the evaluation phases. During the design focus group session, the mobile application was given an isiXhosa name - *Khangela ngesiXhosa* (translated: Search in isiXhosa). The name was aimed at highlighting the purpose of the application while maintaining the objective of having a completely isiXhosa interface.

The interface had a help menu even though at this stage the menu did not serve any purpose. The help menu was made up of a simple question mark icon as illustrated in Figures 3.2 and 3.3. The position and simplicity of the help menu adhered to the design principles reviewed in the literature, that is, keeping important information at the top as well as reducing amount of information that can be viewed at a given time. The query submission interface had the search box as recommended during the design focus group, which was also positioned at the top of the interface for similar reasons as discussed above. The search box was used by participants to type in search terms. They then clicked on the button *'khangela'* to submit the query to the search engine. The results interface would then be presented to the user with the title and URL of each result. Upon clicking on one of the listed results, the application would open the default browser of the mobile phone and display the contents of the visited result, as illustrated in Figure 3.4. This simple design was based on the design guidelines of the focus group and the reviewed literature as explained above, although at this stage the focus was not to achieve every design guideline due to time constraints.

At the end of the first design and development phase, the interface was informally tested and evaluated by fellow researchers in the ICT4D laboratory in the Department of Computer Science at the University of Cape Town, South Africa. The evaluation focused on usability and resulted in recommendations to include new design features and in some cases remove some of the existing features. There were 3 Computer Science PhD students who evaluated the interface, and the feedback received was that the interface had a good starting point but more design needed to be done. The ICT4D laboratory researchers were an ideal choice for two reasons: a) they are experienced researchers who focus on building prototypes that serve a similar purpose as the prototype of this study and b) they were easily accessible. The informal testing and evaluation phase raised several design modifications that led to the refinement of the design, discussed in the next section.

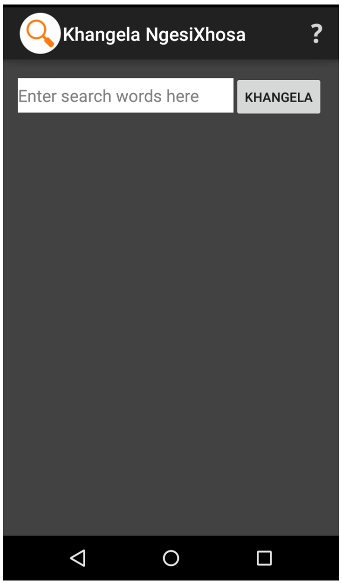


FIGURE 3.2. Mobile text interface's query submission interface

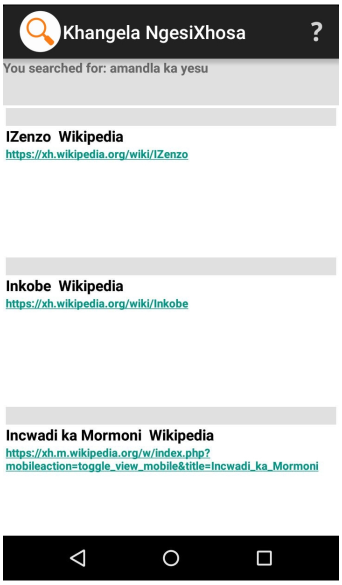


FIGURE 3.3. Mobile text interface's most relevant results interface

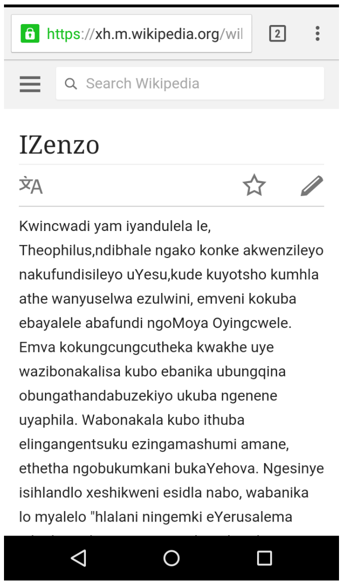


FIGURE 3.4. Single result opened in mobile phone's default browser

3.2.2 Second Prototype

The second prototype was designed and built based on the design modifications that were brought forth during the testing and evaluation of the first prototype. These modifications resulted in the new design shown in Figures 3.5 and 3.6. There were two design suggestions made on the first prototype: a) include an information menu to inform users about the purpose of the interface and b) change the help menu icon to a more familiar icon. The two suggested changes were therefore included in the second prototype, as reflected in Figures 3.5 and 3.6.

At the end of the refinement of the interface design, the interface was informally evaluated by 2 native isiXhosa speakers. This time around the evaluation focused not only on usability but also on whether the interface delivered correct content as most of the content was translated from English to isiXhosa. The evaluation phase led to the refinement of the design, which is discussed in the next section.

3.2.3 Final Prototype

The third and final prototype focused on delivering the interface to be used in the final experiments for this study. The prototype had a lot of changes, which came about during the testing and evaluation of the first and second prototypes. The most important or noticeable change is the prompting hint message, shown in Figure 3.7, which reads *"Khangela apha..."* (translated: search here). The second change is that of highlighting the search box to make it clear to the user

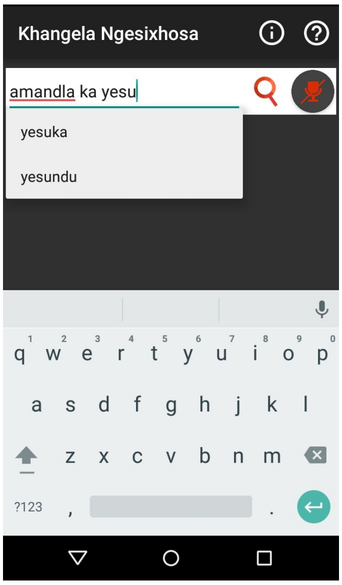


FIGURE 3.5. Mobile text query submission interface

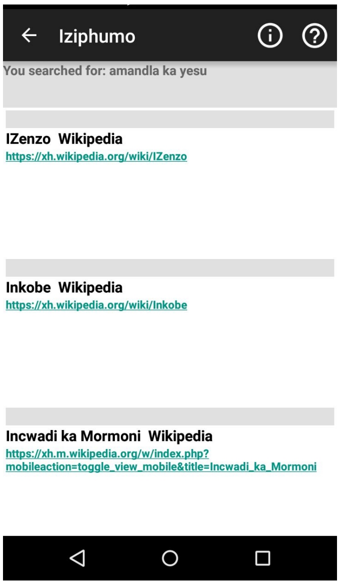


FIGURE 3.6. Mobile text interface results' list view

where exactly the search should happen. Following the paying attention to little details guideline, the application name was kept and correctly spelt from *Ngesixhosa* to *ngesiXhosa*. In regard to the results interface shown in Figure 3.8, the back navigation link was removed to force the user to use the device's back button so as to capture the amount of time spent on each result. Furthermore, to avoid bias towards the text interface the auto-complete feature was removed. This feature was impossible to implement in the voice interface in the given time-frame of this project thus there was going to be bias towards the mobile text interface as participants could use this feature and be more productive with the text interface and not the voice interface that did not have the feature. Therefore, the user's efficiency on the two interfaces could be judged on a feature that exists on one interface and not the other hence creating biasness. The summary snippets were also not implemented for the same reason.

The results interface further included the amount of time the search engine took to return the search results; this is denoted by the statement *"into oyikgangeleyo"* (translated: what you searched for). The number of results returned per query was denoted by the statement *"30 iziphumo (0.025 kwimizuzwana)"* (translated: 30 results in 0.025 seconds). However, the number of results returned and the time taken varied per search query. This information was a fulfilment of the give users informative and efficient feedback guideline, which was adapted from the reviewed literature as users were kept informed of the search process. The interface further offered users the ability to page through the results using the two navigation buttons *"Elandelayo"* (translated: next) and *"Emva"* (translated: previous), as shown in Figure 3.8.

FIGURE 3.7. Mobile text query submission interface

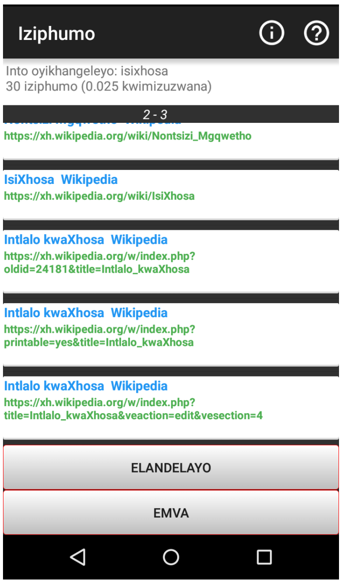


FIGURE 3.8. Mobile text interface results list view

FIGURE 3.9. Mobile text interface single result view

The results interface presented ten results per page, and the user could tap on any of the results to view the result in detail. This would open the result in another window as shown in Figure 3.9. Unlike in the results view menu in the previous two prototypes, the new menu now opens within the application, and does not make use of the mobile phone's default browser. This was done so as to capture the amount of time the user spent on viewing a single result. Noticeably, the menu did not present any back navigation button; this was meant so that the user was forced to use the device's back button to allow the application to capture the result exit time.

The contents of the help and information menu were also added to the interface at this stage. The information menu shown in Figure 3.10 consisted of a pop-up menu with short text. This was to keep the information menu as simple as possible, therefore aligning its design with the recommended design guidelines as per the design focus group as well as those reviewed in the literature as discussed above. The information menu has the title *"Iingcombolo"* (translated: information), and its body reads *"Le-system ikuvumela ukuba ukhangele nantoni na ofuna ukuyikhangela ngesiXhosa"* (translated: this system allows for you to search for whatever information you need in isiXhosa). Figure 3.11 shows the help menu with title *"Uncebo"* (translated: help), and its body reads *"Faka umbhalo kwibhokisi emhlophe, wandule ukucofa iqosha elisekhohlo kwebhokisi leyo"* (translated: enter words to search for in the search box, then press the button on the right to search).

The help menu was also kept as simple as possible so as to align its design with the requirements of isiXhosa speakers, as pointed out in the design focus group. Noticeably, the help and information menus were also positioned at the top of the interfaces to align their design with the design guidelines reviewed in the literature (see Section 2.6.1). Similarly, the number of results returned and the time it took were also position at the top.

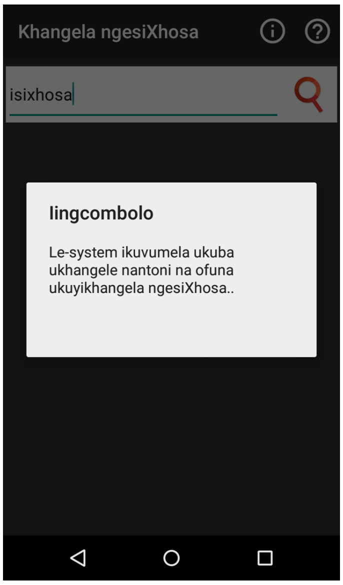


FIGURE 3.10. Mobile text interface Information view

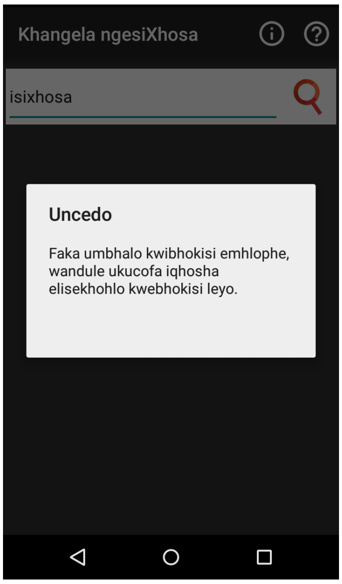


FIGURE 3.11. Mobile text interface's Help view

3.3 Mobile Voice Interface

Similar to the mobile text interface, the mobile voice interface was built by the researcher on the Android mobile platform using the Android Studio Integrated Development Environment (IDE). The interface was complemented by an Automatic Speech Recognition (ASR) system which had 3 components, namely: a) a Language Model b) an Acoustic Model and c) a Pronunciation Model, which were incorporated into an Android application. The design of the interface began with a design focus group session that involved 4 native isiXhosa speakers in the design process. As discussed in Section 3.2, the 4 native isiXhosa speakers were well experienced in search systems - both voice and text systems. The result of the focus group was a list of design guidelines, which were summarized into the following: 1) the user must press a button or use voice activation to begin searching for information on the Web; 2) as users speak, display search terms in text format in a search box as a form of feedback; 3) use clear icons to represent functions, for example, a human head with sound waves emitted from the mouth to inform the user to begin speaking; 4) have the interface respond back to the user in voice; 5) do not put any adverts in any of the interfaces; and 6) use more visuals than words.

The ASR system for this study was built using the Carnegie Mellon University (CMU) Sphinx speech recognition library. This is a large-vocabulary speaker-independent continuous ASR system [37]. As depicted in Figure 3.12, the user speaks a voice query that is recognized by the speech recognition engine. Firstly, the engine extracts speech features from the speech signal in the voice query. These are then passed on to the ASR Decoder, which decodes them using a combination of the Language Model (LM), Pronunciation Model (PM) and an Acoustic Model (AM). Finally, the speech recognition engine sends back the decoded voice query to the application if there was a match in the Language Model.

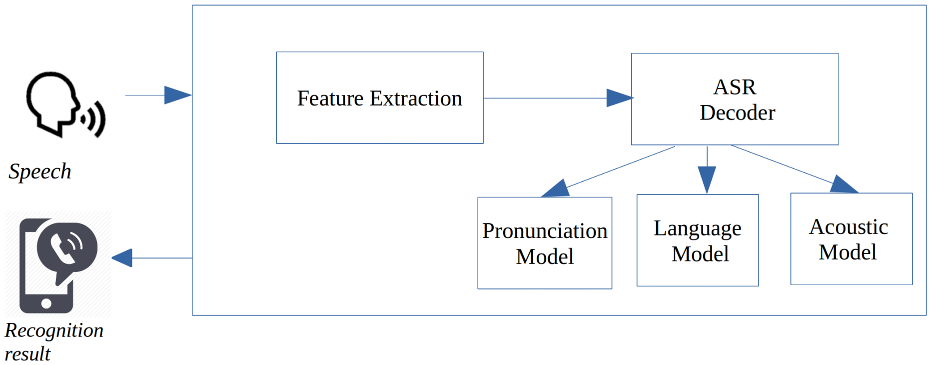


FIGURE 3.12. The overview of the back-end of the speech recognition engine

Following the outline in Figure 3.12, the design of the mobile voice interface was initiated by building the LM discussed in the next section.

3.3.1 Statistical Language Model

There are different types of Language Models that are used in speech recognition systems [21]. This study adopted a statistical Language Model. The purpose of the Language Model (LM) in speech recognition systems is to compute the probability of words and their combinations based on training data [21][1]. This type of LM is best suited for free-form speech input, where it is not possible to predict well in advance all the possible words (and their sequences) the user may speak at a given time. The LM therefore became an ideal choice to include in the ASR system training as the aim was to support as many unknown voice queries as possible.

The LM contained the probabilities of different words and their combinations, as shown in Figure 3.13. The numbers on the left and right of the isiXhosa words represent Good-Turing frequency estimations, that is, estimated probabilities of unknown words occurring before or after the given words [2]. Furthermore, the LM was a 3-gram language model, based on a vocabulary of 20 000 words generated from 108 594 isiXhosa sentences.

[1] http://cmusphinx.sourceforge.net/wiki/tutoriallm
[2] http://www.speech.cs.cmu.edu/SLM/toolkit_documentation.html

The process to build the LM was initiated by collecting isiXhosa text corpora. The text corpora were retrieved from the National Centre for Human Language Technology (NCHLT) project that collects text documents, raw corpora and genre classified documents for building speech systems for ten of South Africa's official languages [19]. The text corpora were then processed to suit the requirements of the CMU language modeling toolkit [12], which was used to train the LM. Further processing involved expanding abbreviations and removing mispelt words, removing duplicate words and numbers as well as names of places and people. The next step involved building a closed vocabulary Language Model by feeding the corpora into the CMU language modeling toolkit that produced the needed LM.

```
-2.5246 amathathu amaqhosha 0.6992
-3.0772 amathathu amashumi 0.0836
-1.5391 amathathu amawaka 0.1800
-2.1026 kwaye xa ingxaki
-2.1026 kwaye xa isebenza
-0.8238 kwaye yandisa iinkqubo
-0.8238 kwaye yanelisekile ukuba
-1.8652 ungubani na kwaye
-1.8652 ungubani na lo
-1.3009 wombuso okanye elo
-1.3009 wombuso okanye ke
-0.8238 zovavanyo nokhuseleko lovavanyo
-0.8238 zovavanyo okanye ulwazi
```

FIGURE 3.13. An extract of statistical probability of words and their combinations in the Language Model used in the study

The process to build the Language Model was faced with several challenges. Firstly, the researcher who was also the developer of the LM is not a native isiXhosa speaker. This led to prolonged hours of writing scripts in order to clean up the text corpora, as in most cases there had to be a third party person (a native isiXhosa speaker) to verify spelling mistakes and remove words that were not part of the isiXhosa vocabulary. Furthermore, the isiXhosa language has a complicated grammar and syntax since it is a click language with few documents available online to help non-native speakers to quickly learn the language [16].

3.3.2 Pronunciation Model

The Pronunciation Model (PM) was made up of a Phonetic Dictionary (PD), which serves the purpose of providing the ASR Decoder with the data needed to match vocabulary words to their respective phonemes, as shown in Figure 3.14 [3]. The PD was built out of all the twenty thousands words in the vocabulary file generated while building the statistical Language Model in Section 3.3.1. As shown in Figure 3.14, the PD had each word in the LM with its respective pronunciation

[3] http://cmusphinx.sourceforge.net/wiki/tutorialdict

transcription. The initial plan regarding the PD was to use the pre-built dictionary from the Lwazi project discussed in the reviewed literature in Section 2.5.4 of Chapter 2, and shown in Figure 3.14. However, this attempt failed as most of the isiXhosa click phoneme characters were not accepted by the CMU Phonetic Dictionary training tool[4].

The main challenge with building the PD was that the CMU dictionary trainer did not accept case-sensitive characters as well as special characters, which were used to represent clicks by the pre-built dictionary retrieved from the Lwazi project (*abasisigxina, wagqiba, and wagqibela*), as shown in Figure 3.14. An attempt was made to convert all special characters and case-sensitive variants to case-insensitive and alphanumeric characters, as recommended by CMU Sphinx while the researcher was engaging them in possible alternative solutions [5]. However, this attempt failed as the trainer did not correctly map the phonemes to the words in the dictionary during training. This approach resulted in low recognition accuracies by the speech recognizer of below 40% and high sentence error rate of 91% and word error rate of 75%.

The 40% accuracy discussed above was rather too low to assume any voice recognition by the speech recognizer. A solution to this problem was to adopt the English PD from the CMU speech system to isiXhosa, as recommended by CMU Sphinx and shown in Figure 3.15. This yielded a good recognition accuracy rate of 64%, a Sentence Error Rate (SER) of 43% and a Word Error Rate (WER) of 46%. These figures are rather to demonstrate the impact the PD had on the overall speech recognition accuracy. However, the study does not allude that this was the only factor that resulted in lower accuracies and higher percentage rates for SER and WER. The details into why the recognition accuracy was low, SER and WER were high are the primary focus of Speech focused projects, which is beyond the scope of this study.

```
abasebenza   a b_< a s E b_< E n z a
abasebenzi   a b_< a s E b_< E n z i
abasisigxina a b_< a s i s i |\|\g_0 i n a
abasweleyo   a b_< a s w E l E j O
abathanda    a b_< a t_h a n d a
wagqiba      w a !\g_0 i b_< a
wagqibela    w a !\g_0 i b_< E l a
wagwetywa    w a g w E c_> w a
```

FIGURE 3.14. An extract of the Phonetic Dictionary from the Lwazi project showing vocabulary words with their respective phonemes

[4] https://github.com/cmusphinx/g2p-seq2seq

[5] https://sourceforge.net/p/cmusphinx/discussion/help/thread/44d21cc8/?limit=25

abasebenza	ae b ah s ah b eh n z ah
abasebenzi	ae b ah s ah b eh n z iy
abasisigxina	ae b ah s ah s ih g k s ah n ah
abasweleyo	ae b ah s w iy l ey ow
abathanda	ae b ah th ae n d ah
wagqiba	w ae g k ah b ah
wagqibela	w ae g k ih b ah l ah
wagwetywa	w ae g w eh t ah w ah

FIGURE 3.15. An extract of adapted Phonetic Dictionary from CMU sphinx used in the study

3.3.3 Acoustic Model

The Acoustic Model (AM) serves the purpose of computing the statistical representations of all sound units that make up a given word from the Language Model[6]. The Acoustic Model relies on a speech corpus (and its corresponding transcripts) using a training algorithm to generate statistical representations for every given phoneme in a given language. These statistical representations use Hidden Markov Models (HMM), therefore, each phoneme has a corresponding HMM. The CMU Sphinx has three types of Acoustic Models for usage in its speech system. These are the Continuous, Semi-Continuous and Phonetically Tied Acoustic Models (PTM). The difference in the three models lies in speed and accuracy. The Continuous model is the most inefficient while the Semi-Continuous model is fast but less accurate. The PTM has similar accuracy as the Continuous model, and is the fastest [7]. Initially, the study adopted the PTM model, however, after a month of training and getting low accuracies the study used semi-continuous model, which in a short period of time gave good accuracy results discussed below.

The process to build the AM started off by collecting a speech corpus from the Lwazi project [4]. The Lwazi project aimed to collect speech corpora for the purpose of building speech recognition systems for nine of South Africa's eleven official languages [4]. The next step involved processing the corpus to meet the requirements of the CMU Sphinx acoustic model trainer. The CMU Sphinx trainer relies on the speech corpus to extract acoustic model statistics from the speech[8]. The trainer makes use of several files to successfully train the AM. These files are: 1) transcripts 2) phoneset 3) fillers 4) language model and 5) phonetic dictionary. The language model file was made up of the Language Model data discussed in Section 3.3.1 while the filler file contained phonemes that were not included in the Language Model, such as laughs, breaths and silences, as shown in Figure 3.16.

The transcript files consisted of two text files listing the transcriptions for each audio file in the speech corpus. The first transcription file was the training file. In order to prepare this file, a shell script was written that looped through all the audio files, mapping each audio file

[6] http://cmusphinx.sourceforge.net/wiki/tutorialam
[7] http://cmusphinx.sourceforge.net/wiki/acousticmodeltypes
[8] http://cmusphinx.sourceforge.net/wiki/tutorialam

```
<s>       SIL
</s>      SIL
+um+  ++um++
+noise+  ++noise++
```

FIGURE 3.16. Extract of the Fillers file used in the study showing silence and noise notes

to its respective transcription. The end result was a text file listing all the sentences that had occurred in each audio file alongside their respective audio file name, as shown in Figure 3.17. The second transcript file contained transcriptions of audio files that were used for testing the Acoustic Model at the end of the training. The phoneset file contained all the phonemes for the words used in the dictionary, with one phoneme per line, as shown in Figure 3.18.

```
<s> ulumko ukuba uza </s> (nchlt_xho_001f_0001)
<s> omnye ngaphadle kwesizathu </s> (nchlt_xho_001f_0002)
<s> yakokwethu iinkomo zakokwethu </s> (nchlt_xho_001f_0003)
<s> lo mfana athanda </s> (nchlt_xho_001f_0004)
<s> eminye imizekelo yombabazo </s> (nchlt_xho_001f_0006)
<s> utyiswa namasi ukaze </s> (nchlt_xho_001f_0008)
<s> ukuba sisiphi na </s> (nchlt_xho_001f_0009)
<s> yhinito entle mpelele </s> (nchlt_xho_001f_0010)
<s> asekwe phezu kwalo </s> (nchlt_xho_001f_0014)
<s> nalo ke eli </s> (nchlt_xho_001f_0016)
<s> wokugqibela wengcambu yesenziwa </s> (nchlt_xho_001f_0017)
<s> kuhlamba ngantelezi nasiidlo </s> (nchlt_xho_001f_0019)
<s> uneendawo ezifuna ukushiiywa </s> (nchlt_xho_001f_0020)
<s> zona zafika iimini </s> (nchlt_xho_001f_0021)
<s> nakho icuba kwakumnyama </s> (nchlt_xho_001f_0022)
<s> kaloku kulapho be </s> (nchlt_xho_001f_0023)
<s> kuba umfi uphumile </s> (nchlt_xho_001f_0024)
<s> ongezelelekileyo kuba thina </s> (nchlt_xho_001f_0025)
<s> ngayo esithi akamthembi </s> (nchlt_xho_001f_0026)
<s> kakhulu lanela nje </s> (nchlt_xho_001f_0027)
<s> okokuqala emhlabeni uba </s> (nchlt_xho_001f_0030)
```

FIGURE 3.17. Extract of the Transcript file used in the study showing sentences that had occurred in separate audio files alongside their respective audio file names

```
ae
ah
b
eh
g
ih
iy
k
l
n
s
w
z
```

FIGURE 3.18. Extract of the Phoneset file used in the study showing the phonemes for words in the dictionary

Once the above files were ready for use, some configurations were made, namely: a) configuring the type of Acoustic Model to be generated at the end of the training and b) the path to the speech corpus on the computer, as recommended in the CMU tutorial[9]. Thereafter, training of the Acoustic Model was started, which resulted in the generation of Acoustic Model files. The CMU Acoustic Model is made up of seven files, namely: mdef, feat.params, mixture weights, means, noisedict, transition matrices and variances file. These are the files that were added to the Android mobile application to support off-line isiXhosa speech recognition. The accuracy of the trained Acoustic Model is discussed in the next section.

3.3.4 ASR Training and Accuracy Results

While ASR systems are not the focus of this study (see Section 2.5.4), an ASR system was necessary as part of the voice interface. The first step in addressing this involved investigating open-source libraries that could enable the researcher to quickly develop the ASR system that could support offline speech recognition, and without the need to understand the internal workings of ASR systems. This led to the successful adaptation of the CMU Sphinx speech library. Following the CMU Sphinx tutorials on building ASR systems [10], the developer downloaded the NCHLT isiXhosa speech corpus from the RMA website [11]. The speech corpora was made up of around 56 hours of isiXhosa speech data from 209 speakers [5]. Unfortunately, after several training sessions the trainer produced poor accuracy rates of below 40%, which was less than the recommended (around 70%) accuracy rate.

After failing to get the recommended accuracy rate with the NCHLT isiXhosa speech corpora, the researcher downloaded the Lwazi speech corpus, which was rather smaller in size (around 20 hours) of isiXhosa speech corpus. However, on first few iterations of training, this corpus produced good recognition accuracy rate of 64% to 72%. Given the objective of the study and recommendation by CMU Sphinx, this accuracy was good enough to enable the development of the mobile voice interface, which was the main focus of the study. Therefore, the study adopted the Lwazi speech corpus over the mobile phone collected speech corpus from the NCHLT project because it worked at an acceptable level of accuracy.

The details of the mobile voice interface prototypes are discussed accordingly in the next sections.

3.3.5 First Prototype

Similar to the mobile text interface, the voice interface's first design consisted of a mix of English and isiXhosa words. All of the instructions were written in English to guide isiXhosa speakers

[9]`http://cmusphinx.sourceforge.net/wiki/tutorialam`
[10]`https://cmusphinx.github.io/wiki/tutorialam/`
[11]`https://rma.nwu.ac.za/index.php/nchlt-speech-corpus-xh.html`

who would evaluate the interface to translate the instructions or rewrite them, in order to decide what the expected actions should be.

Figure 3.19 shows the interface with a rotating wheel icon that informs the user that the application is still loading the speech recognizer settings. Additionally, there are help and information menus, which at this stage did not serve any purpose as they were not yet implemented.

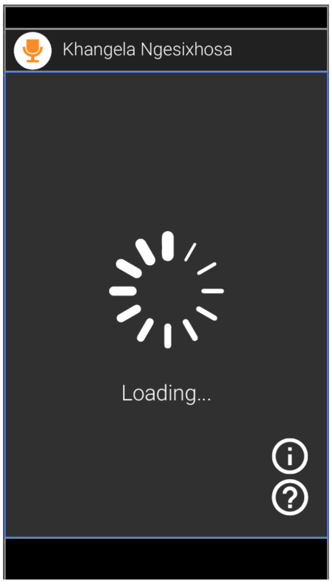


FIGURE 3.19. Mobile voice interface initializing speech recognizer settings

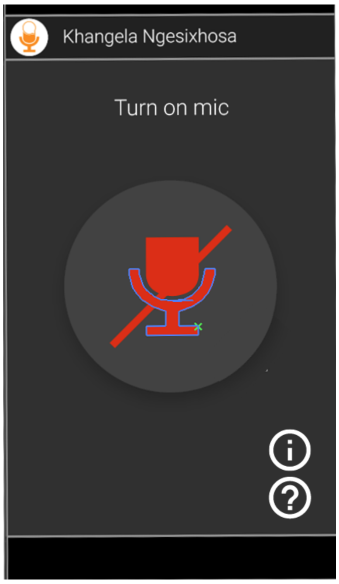


FIGURE 3.20. The prompting interface for the mobile voice interface

In Figure 3.20, the user would tap on the mute-microphone icon to enable the application to access the device's built-in microphone. In Figure 3.21, the user is prompted to tap on the microphone icon, where they can now start speaking into the phone. The next interface, Figure 3.22, would then inform the user to tap on the microphone button again to get the search results. At this stage, the interface did not provide any search results, but rather displayed the spoken query, provided the query was recognized by the ASR Decoder.

After the design and development phase was completed, the interface was informally tested by fellow researchers in the ICT4D laboratory in the Department of Computer Science at the University of Cape Town. These were the same 3 researchers who had evaluated the mobile text interface. The evaluation focused on usability and resulted in recommendations to include new design features and in some cases modify some of the existing features. The interface was then refined based on the new suggested features, which are discussed in the next section of the chapter.

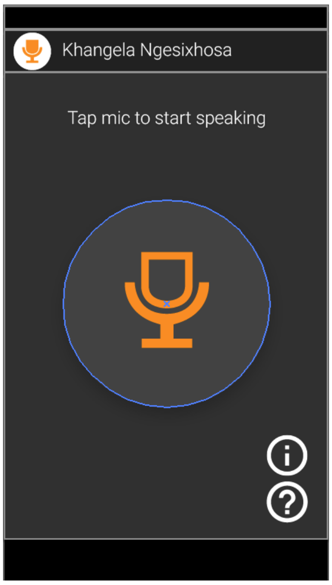


FIGURE 3.21. The mobile voice interface prompting the user to tap microphone to start searching

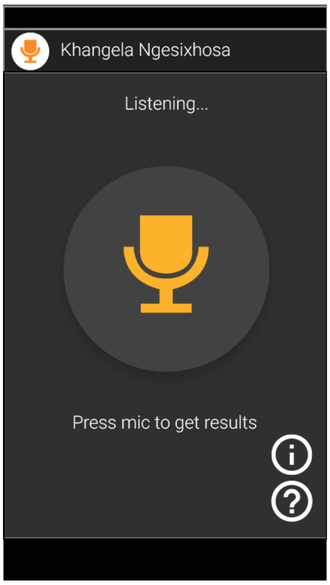


FIGURE 3.22. Mobile voice interface alerting the user to start speaking search queries into a mobile phone

3.3.6 Second Prototype

During the testing phase of the first prototype, it was recommended that the help and information menus should remain consistent in both the text and voice interfaces. This led to moving both menus to the action bar of the application as is the case in the mobile text interface. The help and information menus, however, did not serve any purpose at this stage. Furthermore, it was recommended that the initial loading message should be changed to a more descriptive message, preparing the recognizer, as shown in Figure 3.23. The step of activating the microphone prior to starting the search was also removed, and the suggestion was to let the application do the activation on behalf of the user. This led to the change in the prompting message to now read as "Tap the button and speak to search", as shown in Figure 3.24.

To keep the user informed while speaking voice queries into the mobile phone, the interface messages were changed by giving more descriptive messages, as shown in Figure 3.25. Additionally, the instructions were now in a single line to avoid overloading the user with lots of information, and in turn their short-term memory. This resulted in the instructions, which read "Listening....tap to get results". Similar to the previous prototype, the results interface had not been implemented yet so the user would only see the search term in the case where the term was detected by the speech recognizer. The search term would then be displayed in the same location as the prompting message, as shown in Figure 3.25.

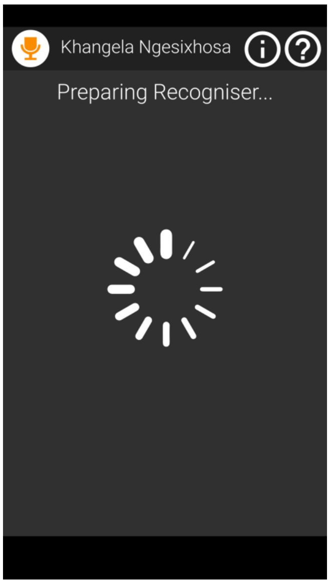


FIGURE 3.23. The mobile voice interface with changes from the testing phase

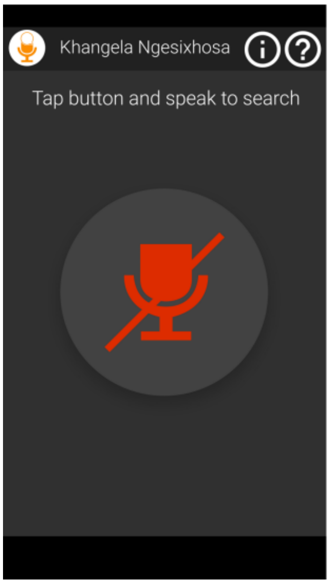


FIGURE 3.24. The mobile voice interface with the new prompting instructions

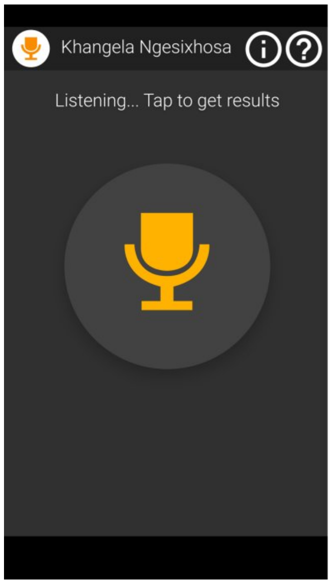


FIGURE 3.25. The mobile voice interface informing the user to speak and retrieve results

The interface in the second prototype was informally tested and evaluated by the two native isiXhosa speakers who had earlier evaluated the mobile text interface. The evaluation focused not only on usability but also on whether the interface delivered correct content as most of it was translated from English to isiXhosa. The testing and evaluation phase led to further changes in the design, as discussed in the next section.

3.3.7 Final Prototype

The final prototype was the end result of the design guidelines from the focus group, as discussed in Section 3.3, as well as modifications that were brought forth during the testing and evaluation phases in the first and second prototypes. The final prototype consisted of more visuals than words, as recommended by native isiXhosa speaking participants during the design focus group session.

In the first interface (Figure 3.26), the user is presented with a rotating-wheel icon to signal the loading process, and the speech clip *"Ukulungiselela i-system, nceda ulende"* (translated: preparing the system, please wait) is played. Once the application finalizes preparing the speech recognizer, the user is presented with the query submission interface in Figure 3.27. While this interface is being loaded, a speech clip is played that states *"cofa iqosha uthethe, cofa iqosha*

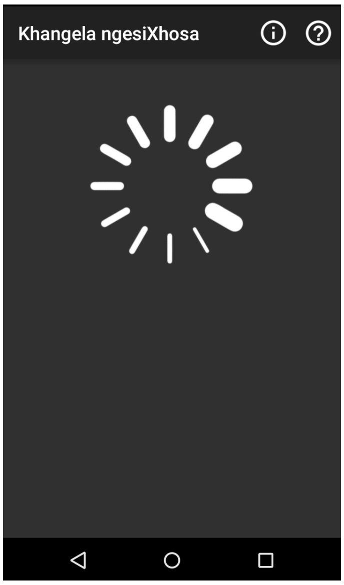


FIGURE 3.26. A rotating wheel icon to inform the user the application is loading the speech recognizer settings

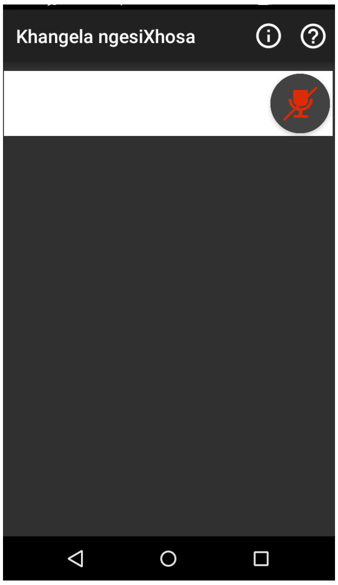


FIGURE 3.27. The home view for the voice interface where the user starts the search process

kwakhona ukufumana iziphumo" (translated: tap the button and speak, tap the button again to get results). The user can then tap on the search button with the red microphone icon and start speaking the search terms, as shown in Figure 3.28.

Upon tapping on the search button, the interface immediately attempts to interpret every sound signal. Additionally, the search button starts to blink and changes from a muted-microphone icon to a human-head icon with emitted sound waves. The interface keeps the user informed by displaying recognized search terms in the white box, as shown in Figure 3.28. The user can stop the search by tapping on the search button to retrieve results relative to the recognized search term. In the case where the search term does not return any results, the interface plays the speech clip *"Uxolo, akukho ziphumo, phinda kwakhona"* (translated: sorry, there are no results, please search again). The user can then tap the search button and speak a new search query. In the case where the search query returns results, the interface will inform the user that there are relevant results for the given query by playing the speech clip *"iziphumo zakho zikhona, cofa kulamaqosha alandelayo uzifumane"* (translated: your results are available, use the buttons below to access them).

The mobile voice interface served similar roles as the text interface. This was to avoid bias towards the voice interface over the text interface, and vice versa. Therefore, the user could use the left and right arrows to navigate and play each of the returned results. This behaviour is

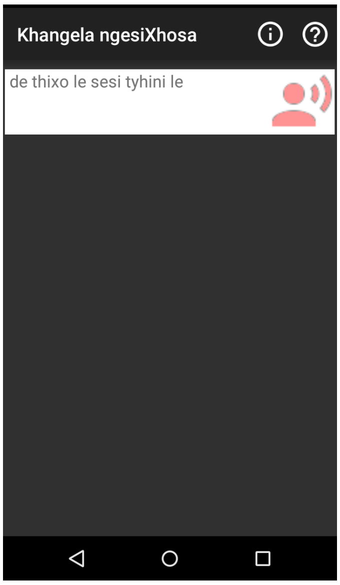


FIGURE 3.28. The mobile voice interface showing detected voice queries as user speaks

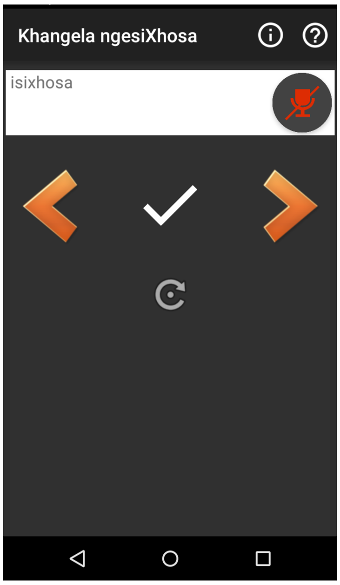


FIGURE 3.29. The mobile voice interface showing recognized voice query after the user spoke it

similar to the user scrolling or paging through the textual results in the text interface. If the user felt the currently played result is relevant to the search they made, the user would simply tap on the button with the tick icon to mark the result as relevant. This feature was added for research purposes in order to keep track of the results the user felt were relevant, similar to the visited results in the text interface. The user had the ability to replay the current result, similar to the text interface where the user could visit the same result over and over again. Lastly, the help and information menus from the mobile text interface were converted to speech. This meant that when the user clicked the help or information menu the same text in the mobile text interface would be played in audio format.

3.4 Desktop Interface

The desktop interface, that was also used as the baseline, was initially developed separately from this study (see Section 2.2). However, the interface was connected to the same search engine as the mobile text and voice interfaces, to avoid bias towards one interface over the other because of content.

3.4.1 First Prototype

The query submission interface for the desktop interface consisted of English words, as shown in Figure 3.30. Therefore, the interface did not meet the requirements of this study, which was to have a complete isiXhosa interface.

To meet the requirements of the study, the interface was redesigned with minor improvements such as changing font style, styling components like buttons, input fields and adding background card style behind text. The interface was then given to two native isiXhosa speakers to evaluate. This led to the new and final interface, which is discussed in more detail in the next section.

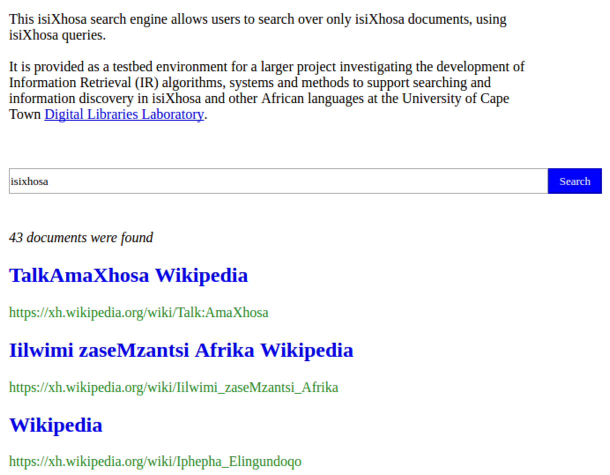


FIGURE 3.30. The initial desktop interface with complete English instructions

3.4.2 Final Prototype

The final prototype was the end product of modifications made to the first prototype with the aim of meeting the requirements of this study, that is, to have a complete isiXhosa interface, as shown in Figure 3.31. The interface name was changed to *Khangela ngesiXhosa*, the same name that is in the mobile voice and text interfaces. Additionally, the description of the interface was changed to read *"Khangela ulwazi ngesiXhosa malunga nantoni na ofuna inkcukacha ngayo"* (translated: search for information in isiXhosa that you need details about). The instructions in the search box were also changed to the same instructions as those on the mobile text interface: *"Khangela apha.."* (translated: search here). The button's name was also changed to *"Khangela"* (translated: search).

The results interface was also changed to neatly space out results similar to the mobile text interface. In addition, the results found message was also changed to isiXhosa, as shown in Figure 3.32.

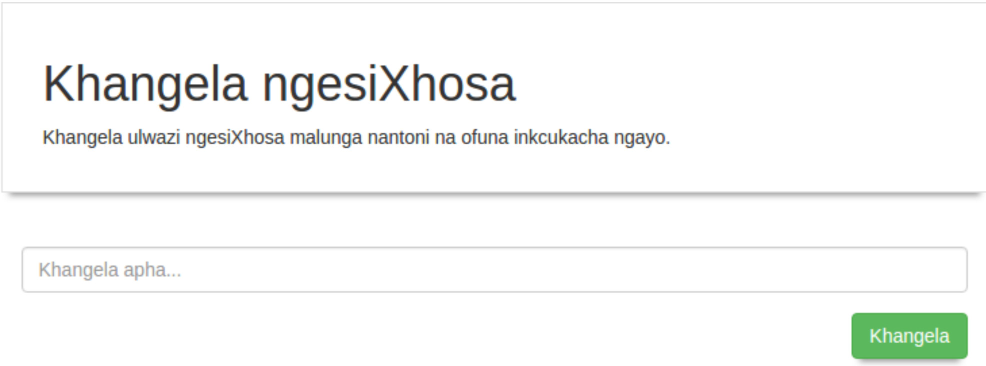


FIGURE 3.31. The query submission interface that has been translated to isiXhosa with refinements on the design

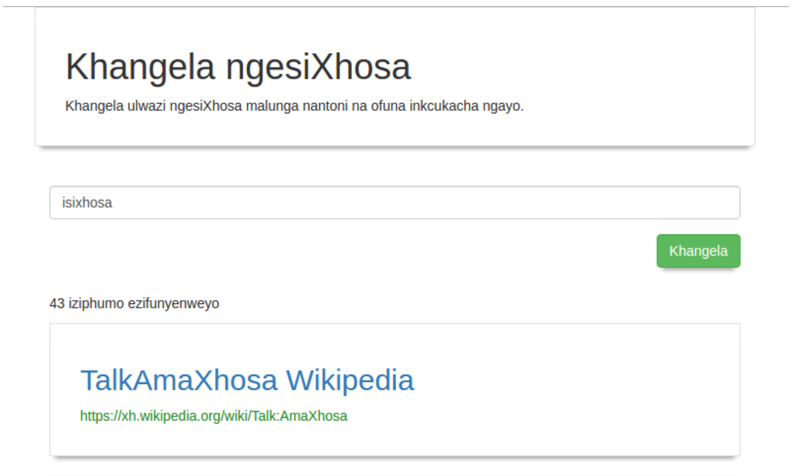


FIGURE 3.32. The results interface with improved modifications from the previous interface

3.5 Chapter Summary

Two mobile interfaces, text and voice, were developed using an iterative design approach. The two mobile interfaces each had unique design requirements. However, both interfaces were developed based on design guidelines raised during a design focus group session with native isiXhosa speakers and those adopted from the reviewed literature.

Both the mobile text and voice interfaces had two prototypes that led to the final prototype used in the final experiments for this study. The first prototypes were informally tested for usability by ICT4D researchers in the Department of Computer Science at the University of Cape Town, South Africa. This resulted in the second set of prototypes that had refinements to the design that were then evaluated by native isiXhosa speakers. The first and second prototype modifications led to the development of the final prototypes that were used in the final experiments of this study. The desktop interface, which was also used as the baseline, was improved in design to suit the requirements of this study.

The successful completion of designing and developing the two mobile interfaces plus the baseline desktop interface meant that the study could move onto the evaluation of the proposed interfaces. This is discussed in the next chapter.

Chapter 4

Experiment Design and Evaluation Methods

This chapter discusses the evaluation process used in this study by first describing the design of the experiment, participants in the study, and the sampling strategy used to select them, in Sections 4.1.1 and 4.1.2. Thereafter, Section 4.2 describes the different data collection tools used to gather data needed to answer the research questions for the study. In Section 4.3, how the experiments were conducted is discussed, with Section 4.3.1 discussing the environment experiments were conducted in. The search queries and search tasks used to evaluate the interfaces are discussed in Sections 4.3.2 and 4.3.3, respectively. The pilot study conducted before the final experiments is discussed in Section 4.3.4, and the procedure followed during the final experiments is presented in Section 4.3.5. Section 4.4 discusses the prototype that had the two mobile interfaces, which were evaluated during the final experiments. In Section 4.5 the criteria used to address research questions for the study is presented. The chapter concludes with Section 4.6, that gives a summary of the chapter.

4.1 Design of the Experiment

4.1.1 Sampling

This study adopted a convenience sampling strategy because most of the participants who met the requirements of the study were already accessible within the university. Participation was on a voluntary basis, which meant participants participated in the study by choice and were free to withdraw participation at any given time. The recruitment process was initiated by sending out an email to the entire university, targeting native isiXhosa speakers to invite them to participate in the experiment. A total of 10 native isiXhosa speaking students were targeted for the pilot study and a minimum of 30 native isiXhosa speaking students were targeted for the final experiments. The number of participants needed for the final experiments was computed using SurveyMonkey's[1] sample calculator. Additionally, to ensure a large turnout of participants:

[1] `https://www.surveymonkey.com/mp/sample-size-calculator/` (population-size: 8,000,000, confidence-level (%): 95, margin of error (%) : 17)

a) participants were invited to the study as close to the experiment date as possible, and b) each participant was given a R40 cash incentive for participating in the study.

4.1.2 Participants

There were three groups of participants who took part in the study, namely: a) those who took part in the design focus group, b) those who participated in the pilot study and c) those who participated in the final experiment to evaluate the mobile interfaces. All these groups of participants were students so an ethics clearance application was made to the Science Faculty and Department of Student Affairs to request access to the students, as shown in Appendix A. The focus group participants were selected to come up with the design of the two mobile interfaces: text and voice. The objective was to have the interface co-designed with native isiXhosa speakers. This approach is supported by the literature that emphasizes the need to engage target users in the design process of interfaces, starting from the initial design until the end product is achieved [73][74].

A total of 10 native isiXhosa speaking students were invited to the design focus group and only 4 turned up for the session held on 13 October 2015. Three of the participants were from the Computer Science honours class, and had been trained in interface design and search engines through HCI and IR courses offered within their curriculum. The other participant was a second year student from the Mechatronics Department, and had been trained in search engines and speech systems and, in particular, speech to control robots. Therefore, the participants were a right fit for the design focus group.

The next set of participants were undergraduate and postgraduate native isiXhosa speaking students, of which 3 participated in the pilot study and 34 participated in the final experiments. However, the 3 students who participated in the pilot study did not form part of the 34 who participated in the final experiments due to other commitments. Considering the aim of the study, which was to compare mobile search interfaces for isiXhosa speakers, the chosen participants were deemed the right fit for the study.

In every group of participants, questions regarding the study were dealt with and the participants were given informed consent forms, which they signed, acknowledging their participation as voluntary, as shown in Appendix B.

4.2 Data Collection Tools

The data collection techniques used in the study were carefully selected to ensure they enabled the collection of as much data as possible to address the key research questions for the study. The following methods were used and are discussed in detail below: SUMI questionnaires and application log files.

4.2.1 SUMI Questionnaires

Electronic questionnaires were presented to the participants to get their feedback regarding the three interfaces developed in this study, as discussed in Chapter 3. These were electronic questionnaires completed using a laptop, and were constructed and provided for use in the study by SUMI. SUMI[2] is a tried-and-tested standard method of extracting information about software quality as viewed by the user [33]. It is used across commercial and academic organizations all over the world. SUMI is scored and interpreted with reference to a standardized database of about 2000 profiles, ranging from office systems and communications software to computer-aided design packages [31]. The SUMI questionnaires were used to separately collect information on the participants' experiences with each of the three interfaces.

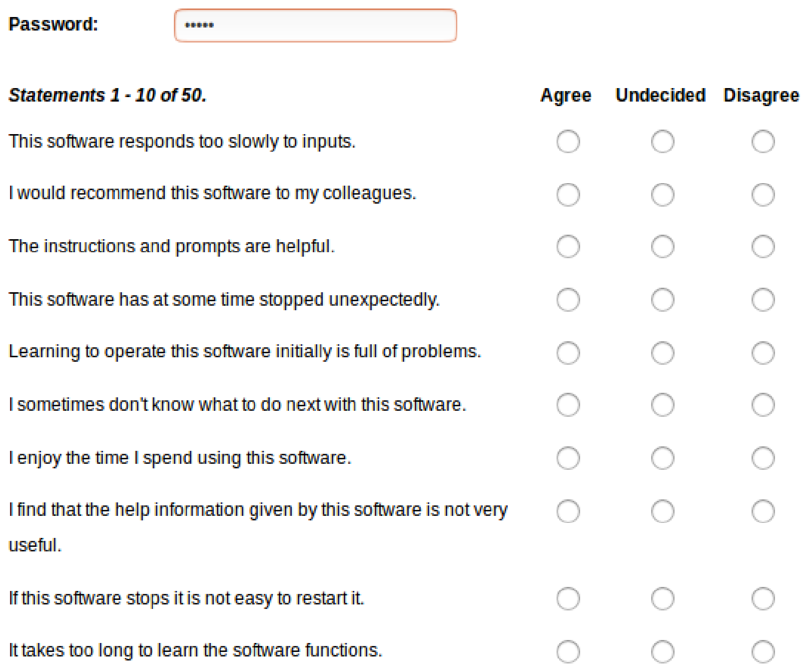
The SUMI questionnaire contains 50 questions, from which participants had three options to select, namely: agree, disagree and undecided. The agree option implies that the participants feel the questionnaire statement relates to their experiences with the interface while the disagree option means the user feels the questionnaire statement does not relate to their experiences with the interface. The undecided option implies that the participants could not decide on the answer to the given questionnaire statement. SUMI questionnaires provide analyzed results in two forms: firstly, the Global scale score, which reflects general usability of the given interface; and, secondly, the 5 usability subscales: a) Helpfulness b) Affect c) Control d) Learnability and e) Efficiency [31].

Helpfulness measures whether the given software is self-explanatory and has enough help facilities such as documentation and help menus. Affect measures the user's emotional behaviour towards a given software, that is, the likability aspect. Control measures the degree to which the user feels in command of the software as opposed to the software controlling the user while performing a given task. Learnability measures the adaptiveness and speed with which the user feels they were able to master the given software. Efficiency measures the degree to which the user feels the given software helped them to do their tasks. A sample questionnaire from the experiment is shown in Figure 4.1, with a completed questionnaire presented in Appendix C. The questions have been arranged cyclically to address Efficiency, Affect, Helpfulness, Control and Learnability subscales. Therefore, question 1 measures Efficiency, 2 measures Affect, 3 measures Helpfulness, 4 measures Control, 5 measures Learnability, 6 measures Efficiency, and so on.

SUMI measures the Global scale with 25 questionnaire statements from the 5 subscales [33]. The 25 questionnaire statements are 7, 8, 15, 18, 19, 21, 23, 24, 26, 27, 28, 31, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 48, and 49, as shown in the completed questionnaire in Appendix C. The Global scale represents the single construct of perceived quality of use rather than a mere average of all the 50 questions on the questionnaire [33]. These were later on confirmed in replication studies [31]. Therefore, they formed a general satisfaction scale with a high reliability coefficient. An average software system scores between 40 and 60 on the Global scale and all

[2] `http://sumi.cc.ie/whatis.html`

the 5 sub-scales discussed above [33]. Any software that rates less than 40 is considered below average and any software that rates above 60 is considered above average. Furthermore, each SUMI Global scale and 5 sub-scales scores can be computed statistically and independently from one another to determine which users to call for an in-depth interview [31].


Password: •••••

Statements 1 - 10 of 50.	Agree	Undecided	Disagree
This software responds too slowly to inputs.	○	○	○
I would recommend this software to my colleagues.	○	○	○
The instructions and prompts are helpful.	○	○	○
This software has at some time stopped unexpectedly.	○	○	○
Learning to operate this software initially is full of problems.	○	○	○
I sometimes don't know what to do next with this software.	○	○	○
I enjoy the time I spend using this software.	○	○	○
I find that the help information given by this software is not very useful.	○	○	○
If this software stops it is not easy to restart it.	○	○	○
It takes too long to learn the software functions.	○	○	○

FIGURE 4.1. SUMI questionnaire with 10 of 50 statements, and the corresponding options for the user to choose from

The participants were told before answering the questionnaire that the word software in the questionnaire refers to the interface they had just evaluated (see Figure 4.1). In addition, a single SUMI questionnaire was used on each interface. In order to separate questionnaires by their respective interfaces, each questionnaire had a unique password corresponding to a particular interface. The passwords were entered by the researcher prior to each participant answering the questionnaire. Participants took an average of 10 minutes to answer the 50 questions on a single questionnaire. Therefore, it took each participant about 30 minutes to complete answering questionnaires for the three interfaces.

4.2.2 Application Logs

Further user interactions with the mobile interfaces and the baseline desktop interface were monitored using application log files. This assisted in collecting and analyzing log data about the series of steps the participant went through while evaluating each interface. The user's privacy and confidentiality was protected as the information collected did not include their personal information but rather their interactions with each interface. Additionally, the application log files were captured within the mobile phone for the two mobile interfaces, and a laptop for the baseline desktop interface. This made it easy to extract the information immediately after each participant finished working with any of the interfaces. This was done to ensure availability of the results from each part of the experiment. Figure 4.2 shows the log data from the baseline desktop interface. The log file captured the search term as shown in line 1 and the time the page took to be loaded in line 2.

The time the user started typing a search term to the time they submitted it was also captured, as shown in lines 3 to 4 respectively. The time the user started viewing a single result was also captured, as well as the results they had viewed, as shown in lines 5 and 6.

```
1.  "Search Term: " umhlaba
2.  "Page Loaded At": 14:11:23
3.  "Typing Finished At: "14:16:16
4.  "Submitted At: "14:16:19
5.  "Result Viewing Started At: "14:16:30
6.  "Result Visited: " [Umhlaba  Wikipedia https://xh.wikipedia.org/wiki/Umhlaba]
7.  "Result Viewing Finished At: "14:17:17
8.
9.  "Search Term: " isixhosa
10. "Page Loaded At": 14:16:19
11. "Typing Finished At: "14:22:41
12. "Submitted At: "14:22:42
13. "Result Viewing Started At: "14:22:53
14. "Result Visited: " [TalkAmaXhosa  Wikipedia
    https://xh.wikipedia.org/wiki/Talk:AmaXhosa]
15. "Result Viewing Finished At: "14:23:05
```

FIGURE 4.2. Log file showing logged data from the baseline desktop interface

The log files for the mobile text interface captured similar times and events to that of the baseline desktop interface discussed above. Firstly, the time the application was fully loaded was captured, as shown in line 2 in Figure 4.3. The time the user took typing or correcting the search query was also captured, as shown in lines 3 to 4. The time the search term was submitted was also captured in line 7 with the corresponding search query captured in line 6. The time the user started viewing results was also captured, as well as the results viewed, as shown in lines 8 and 9 respectively.

The mobile voice interface also had log files, which captured the user interactions with the interface similarly to the two interfaces discussed above. Figure 4.4 shows the log file for the mobile voice interface, with line 70 showing the spoken and recognized search term. The most relevant results were denoted by their position in the list of 10 results, as shown in line 72. Furthermore, the log file captured the time the user attempted speaking a given query, as illustrated in lines 73 and 74.

In essence, the log files captured the amount of time the participant took to submit a single search query and act on the results for the given query. Additionally, the log files captured the

```
2. "App Started/Loaded At" : 2:17:54 PM
3. "Began Typing/Corrections At" : 2:17:58 PM
4. "Began Typing/Corrections At" : 2:18:03 PM
5.
6. "Search Term" : umhlaba
7. "Search Initiated At" : 2:18:06 PM
8. "Result Viewing Started At" : 2:18:10 PM
9. ["Umhlaba  Wikipedia" : https://xh.wikipedia.org/wiki/Umhlaba]
10."Viewing Finished At" : 2:18:26 PM
11."Time on result (s)" : 16
12.
13."App Exited At" : 2:18:34 PM
14.
15."App Started/Loaded At" : 2:36:48 PM
16."Began Typing/Corrections At" : 2:36:52 PM
```

FIGURE 4.3. Log file showing logged data from the mobile text interface

```
70."Search Term" : amanzi
71."Result 1 Access Started At" : 15:37:32
72."Relevant result " : [1]
73."Attempt Speak Search Term At" : 15:37:39
74."Attempt Speak Search Term At" : 15:37:47
75.
76."Search Term" : emzantsi afrika
77."Result 1 Access Started At" : 15:37:53
78."Relevant result " : [1]
79."Result 2 Access Started At" : 15:38:01
80."Result 3 Access Started At" : 15:38:03
81."Result 4 Access Started At" : 15:38:17
82."Relevant result " : [4]
83."Attempt Speak Search Term At" : 15:38:21
84."Attempt Speak Search Term At" : 15:38:27
85."Attempt Speak Search Term At" : 15:38:32
86."Attempt Speak Search Term At" : 15:38:36
```

FIGURE 4.4. Log file showing logged data from the mobile voice interface

total number of clicks-throughs on the results for a given query. This information will help in answering key research questions of the study, as discussed in Section 4.5. How and where the experiments were conducted is discussed in the next section.

4.3 Conduct of the Experiment

4.3.1 Environment

All parts of the experiment were conducted one after the other in a quiet and comfortable guest lecture room in the Computer Science Building at the University of Cape Town. The room met the requirement of a quiet location. A quiet location was required to cater for all parts of the experiment and, in particular, the mobile voice interface, which could not work as expected in a noisy environment due to the noise interference that hinders the ability of the speech recognizer to recognize spoken queries. Although the mobile text and baseline interfaces did not require a quiet environment, they were also evaluated in the same room as the mobile voice interface to ensure consistency in the experiment. The room was equipped with a laptop with the baseline desktop interface and two mobile phones, one having the text interface and another the voice interface. The mobile phones and the laptop were connected to the university's network.

4.3.2 Search Queries

A total of 10 short queries were formulated and given to the participants to use when evaluating each of the three interfaces. This effort was intended to make sure that the participants focus on the core task of evaluating the interfaces rather than worrying about expressing their information need through formulating queries [3]. Furthermore, it is sometimes possible for the user to produce poor search queries mainly because they find it difficult to convey their information need through search queries [14]. The queries were formulated from the isiXhosa words in the Language Model developed as part of the speech recognizer as well as terms that produced more document hits on both the desktop and mobile text interfaces. Therefore, the search queries were structured as shown in Table 4.1 to ensure that they were recognizable by the voice interface and that they produce more relevant documents per interface. Unfortunately, the search engine could not collect as many documents as is the case in other search engines such as those of the English language, hence, this limited the quality of the search queries. The researcher was forced to construct queries that could yield relevant results and not necessarily of high quality to avoid bad user experience in the case where the interface did not provide relevant results.

Table 4.1: isiXhosa search queries with their respective English translations

isiXhosa Query	**English Translation**
isiXhosa	The isiXhosa language
emzantsi afrika	South Africa
inkxaso	Support of the females / Support from the females
ndihlala emthatha nabantwana	I live in Mthatha with my children
Isebe lezemihlaba eMpumalanga	The Department of Land Affairs in Mpumalanga
amanzi	Water
isixeko sasekapa	City of Cape Town
umhlaba	Soil
hamba uye emaXhoseni	Go to the land of the Xhosa people
ndikhathazekile ndoda	I am a troubled man

The queries in Table 4.1 were used during the evaluation of each of the three interfaces. The queries were given to the participants without their corresponding English translations (see Appendix D) and, where questions arose about the queries, the researcher answered them to the participants' satisfaction.

4.3.3 Search Tasks

A search task per participant was made up of 10 search queries, discussed in Section 4.3.2. The queries were printed on an A4 page, and a brief overview of the project was highlighted in the Informed Consent form. The participants were notified that there are many relevant documents per query, and as such must try to find as many of these documents as they could.

4.3.4 Pilot Study

A pilot study was conducted prior to the final experiment. The pilot study was conducted to verify the experimental protocol for this study.

The within-groups experimental approach was adopted, where each participant evaluated each of the three interfaces using the 10 search queries in Section 4.3.2. Each participant who came through first evaluated the baseline desktop interface, and at the end of the evaluation they completed the SUMI questionnaire shown in Appendix C. They then evaluated the mobile text interface, and at the end of the evaluation they completed a corresponding SUMI questionnaire. Finally, they evaluated the mobile voice interface, and at the end of the evaluation they completed the corresponding questionnaire. The second participant then evaluated the baseline interface, thereafter the mobile voice interface, then finally, the mobile text interface. These alternations were applied until the last participant, that is, alternating between the mobile text and mobile voice interfaces. This counterbalanced measure was practised to avoid chances of the order of experiments adversely influencing the results. In each part of the pilot study, each participant was given the same search tasks and completed the SUMI questionnaire at the end of the evaluation.

The successful execution of the pilot study meant that the experimental design was feasible and the study could move on to the final experiments.

4.3.5 Procedure

Following the successful completion of the pilot study, a single experiment session was conducted to address the research questions for the study. The experiment was executed under the same conditions as the pilot study, including the environment, prototype, experimental approach and search tasks.

During the experiment, the following guidelines were applied:

- Each participant was given an overview of the study as well as the purpose of the experiment.
- Any questions regarding the research or the experiment were answered to the participant's satisfaction.
- An informed consent form was given to the participant, and the participant guided on how to complete it.

- The participant was given a laptop with the baseline desktop interface together with the printed search tasks to evaluate the interface. Upon completion, the participant answered a SUMI questionnaire for the interface.
- The participant was then handed a mobile phone with the text interface to evaluate using the same search tasks as used to evaluate the baseline desktop interface. Upon completion, the participant answered a SUMI questionnaire for the mobile text interface.
- The participant was given another mobile phone with the mobile voice interface to evaluate using the same search tasks used to evaluate the mobile text interface. Upon completion, the participant answered a SUMI questionnaire for the mobile voice interface.
- The second participant evaluated the baseline interface, thereafter the mobile voice interface then finally the mobile text interface. The order of evaluating the mobile interfaces changed until the last participant in this manner.
- The participants' logged data was extracted from each device - laptop and mobile phone - each time they moved to another device to avoid losing data or data inconsistencies that may arise as more participants evaluated the interfaces.
- Finally, the facilitator thanked the participant as a sign of appreciation, and handed them the R40 cash incentive.

4.4 Prototype

The mobile prototype, which had the two mobile interfaces was developed using the Android platform, and targets Android 3.0 (Honeycomb) and above, as discussed in Chapter 3. The prototype is made up of two search interfaces: text and voice. The main communication language in each interface is isiXhosa, as this is the target group's language. For experimental purposes, the two interfaces were separated into two mobile applications to make a clear distinction between the two interfaces to participants in the study. There were two mobile phones pre-installed with the prototype, one having the text interface and another having the voice interface. The two mobile interfaces needed an Internet connection in order to fetch the search results from the search engine server hosted in the Department of Computer Science at the University of Cape Town, South Africa. In the case where Wi-Fi was not available, mobile data was used.

As shown in Figure 4.5, the mobile text interface allows the user to search for information by typing in text queries in the white search box, then pressing the search button to retrieve a list of search results in text, as shown in Figure 4.6. The participant can then scroll through the results and access those most relevant to their search.

In the mobile voice interface shown in Figure 4.7, the participant taps the microphone icon and speaks a search term, thereafter obtaining a list of the most relevant search results in voice.

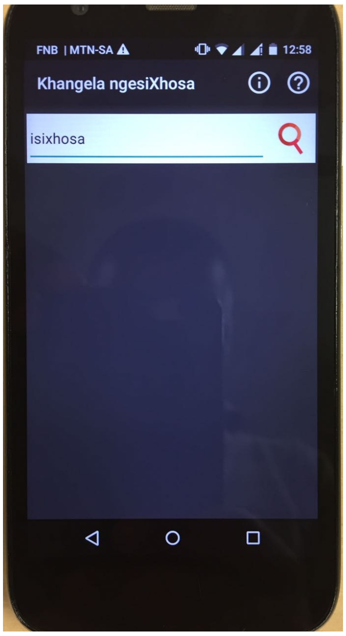


FIGURE 4.5. The query submission interface for the mobile text interface, where users submit text queries

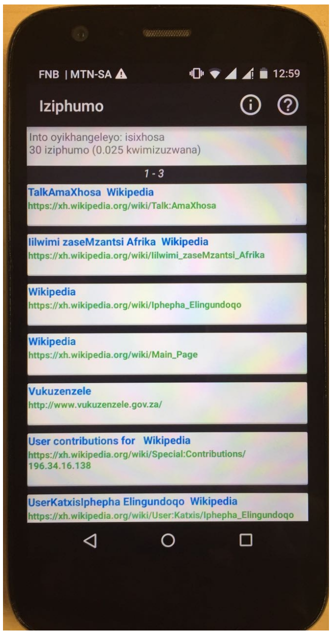


FIGURE 4.6. The search results interface for the text interface with a list of most relevant results

The participant can opt to listen, replay, mark results as relevant or skip through results as they are being played, using the buttons shown in Figure 4.8. The two orange arrow buttons provide back and forth navigation through the list of the voice results. The tick button marks the currently listened to result as the most relevant.

The desktop interface acting as a baseline for the experiment was a Web application, which was accessed from a laptop that was provided to participants during the evaluation, as shown in Figure 4.9. The interface was connected to the same search engine server for fetching the search results as the mobile interfaces using Hypertext Transfer Protocol (HTTP), thus ensuring consistency in the experiment.

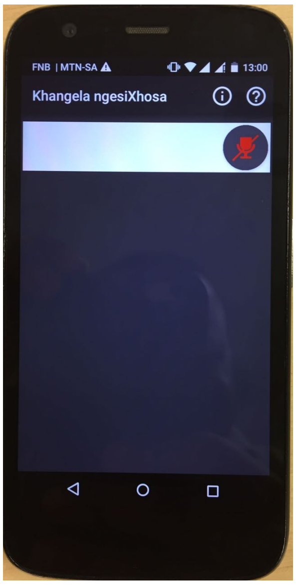


FIGURE 4.7. The query submission interface for the mobile voice interface prompting the user to start speaking

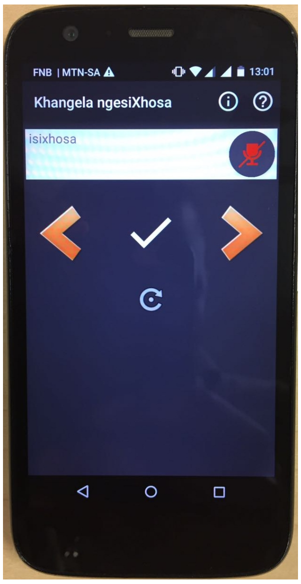


FIGURE 4.8. The search results interface for the mobile voice interface with a list of most relevant voice results

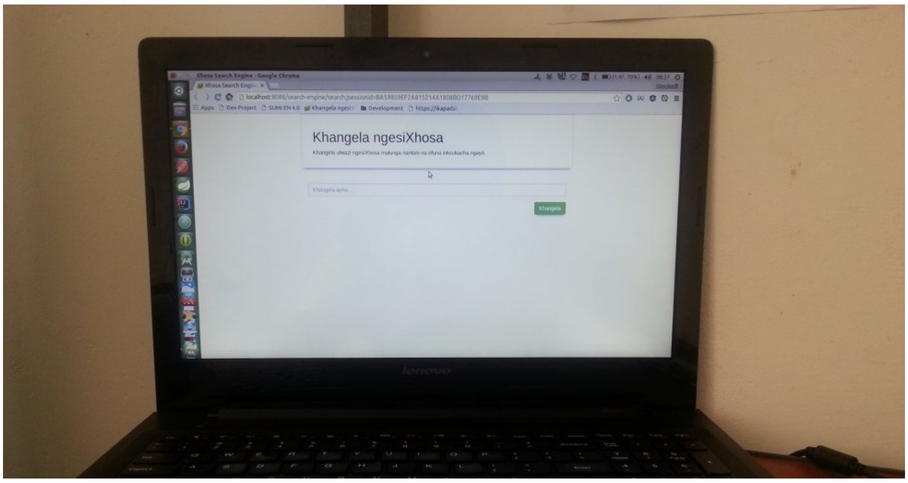


FIGURE 4.9. The query submission interface for the baseline desktop interface opened inside the Chrome browser on a laptop

4.5 Research Questions Evaluation Criteria

The experiment design discussed above was used during the evaluation, where each participant interacted with the three interfaces. Each participant's interaction with an interface was measured using the following effectiveness metrics: speed, relevance and satisfaction. The following section discusses how these metrics were used to address the research questions for the study.

4.5.1 First and Second Questions

a) Is a mobile text interface more effective than a non-mobile text interface for isiXhosa speakers when searching for information on the Web?

b) Is a mobile voice interface more effective than a non-mobile text interface for isiXhosa speakers when searching for information on the Web?

The two questions above compare the mobile text and voice interfaces against the baseline desktop interface. The three interfaces have been developed from scratch, utilizing best practices in IR, as discussed in Chapter 3.

In order to address the effectiveness of the mobile text and voice interfaces against that of the baseline desktop interface, the following metrics were assessed: speed, relevance and satisfaction. Speed was measured by considering the amount of time each participant took on the interface for a given query. The time was calculated from the moment the participant submitted a search query by clicking on the search button to the moment the participant acted on the search results. The distribution of speed for a single search query was calculated based on the individual data for the query across all the participants. The overall distribution of speed for the interface was calculated using all the individual data for all the search queries across all the participants. The experiment was controlled; the participants were observed to ensure that they were not distracted by anything upon receiving the search results, thus ensuring consistency in the amount of time the participants took in acting on the search results.

In this study relevance is measured based on the assumption that a click on listed results is assumed to indicate that the interface provided relevant documents for the submitted query. A further assumption made is that the title and URL displayed in the results list (as discussed in Chapter 3) are indicative of relevance of the returned documents. Therefore, the study assumes click-through data as a form of relevance feedback to determine the effectiveness of the interfaces as part of an IR system. The study considers this measure of relevance as weak-relevance as the focus is not on the user's explicit feedback on the returned results nor is it on the submitted query but rather on the number of clicks per search results. Similar assumptions have been made in several studies, where click-through data is used to interpret relevance feedback [9][29].

In all the interfaces - baseline, mobile text and mobile voice - the results viewed were logged in a text file when the user tapped (or clicked in the case of the baseline desktop interface) the results to view them, as discussed in Section 4.4. More clicks demonstrate that the interface provided relevant results and, as such, effectively assisted participants in their search process. Participants may not click on any results for a given query and this will be interpreted as a count of zero (0) across all queries. However, participants may click to open a single result for a given query and this will be interpreted as a count of one (1) across all queries.

To address user satisfaction, SUMI questionnaires were used to collect qualitative feedback from the participants in the study, as discussed in Section 4.2.1. This was used to measure satisfaction based on the SUMI Global scale as well as the 5 subscales, namely: Learnability, Controllability, Efficiency, Helpfulness and Affect. A Shapiro-Wilk test for normality will be used to test the data for normality. In the case where the data is not normally distributed, the Wilcoxon signed-rank test will be used to compare the mobile interfaces with the baseline desktop interface to determine statistical significant difference. In the case where the data is normally distributed, a paired t-test will be used. A significant difference in the compared interfaces would imply the interface with a higher SUMI Global scale rating is the most effective interface.

4.5.2 Third Question

> *c) Is a mobile voice interface preferred over a mobile text interface for isiXhosa speakers when searching for information on the Web?*

This research question looks into determining user preferences for the two mobile interfaces: text and voice. To address this question, SUMI questionnaires were used to collect qualitative feedback from the participants after they evaluated the two mobile interfaces in addition to the speed and relevance metrics, as discussed in Section 4.5.1. Similarly, a paired t-test/Wilcoxon signed-rank test will be used to compare the mobile text interface against the mobile voice interface, and a significant difference in the SUMI Global score for the interfaces would imply the interface with a higher SUMI Global scale rating is the preferred one.

4.6 Chapter Summary

The study was designed by first considering the groups of participants who would participate in the design focus group, pilot study and the final experiments separately. Convenience sampling was used to select the needed number of participants because most of the participants who met the requirements of the study were already accessible within the university. For data collection tools, SUMI questionnaires and application log files were used to collect as much data as possible, so as to address the research questions for the study.

An experimental protocol was designed and then tested during a pilot study to ensure its feasibility. Further, the pilot study and the final experiments were conducted in a quiet and comfortable guest lecture room in the Computer Science Building at the University of Cape Town, South Africa. The room catered for all parts of the experiments. A total of 10 isiXhosa search queries were formulated and given to participants to use when evaluating each of the three interfaces. The within-groups experiment approach was used during evaluations, where each participant evaluated all the three interfaces, starting with the baseline interface, then alternating between the mobile interfaces.

The evaluation criteria discussed in this chapter was aimed at answering the three research questions for the study using the following effectiveness metrics: speed, relevance and user satisfaction. Additionally, a paired t-test or Wilcoxon signed-rank test will be used to compare the interfaces in terms of preference.

The results will be discussed in the next chapter, based on the data collected during the experiments, following the above outlined experiment design and evaluation methods.

Chapter 5

Results and Discussion

This chapter focuses on the results from the experiment, where participants evaluated the two mobile interfaces and the baseline interface. The data collected during experiments is analyzed as per the effectiveness metrics discussed in Chapter 4, namely: speed, relevance and satisfaction. The results analyzed from the application log files are presented and discussed to address both the speed and relevance metrics for each interface. The results from SUMI are presented and discussed to address the satisfaction metric for each interface.

Section 5.1 starts the chapter by giving an overview of the participants who took part in the experiments, while Section 5.2 focuses on the pilot study results conducted prior to the final experiments. In Section 5.3, the effectiveness metrics are discussed for the mobile voice interface. In Section 5.4, the effectiveness metrics are discussed for the baseline interface, while Section 5.5 discusses the effectiveness metrics for the mobile text interface. In Section 5.6, tests for normality are discussed. In Section 5.7, the baseline interface is compared to the mobile text interface while in Section 5.8 the baseline interface is compared to the mobile voice interface. In Section 5.9, the mobile interfaces are compared. The chapter concludes with Section 5.10, which provides a summary of the chapter.

5.1 Participation

A total of 10 native isiXhosa students showed interest in the pilot study after an invitation email was sent to all first year Computer Science students within the University of Cape Town. However, only 3 students participated in the pilot study, as reflected in Table 5.1. The study targeted a minimum of 30 native isiXhosa speakers for the experiments. However, recruiting only first year Computer Science native isiXhosa students did not yield the expected number of participants. Therefore, a decision was made to recruit native isiXhosa students across the university. An invitation email was sent to all University of Cape Town students, which resulted in a total of 71 respondents who showed interest in the study. Out of the 71 respondents, a total of 34 respondents participated in evaluating the three interfaces.

Table 5.1 shows the distribution of participants per interface in both the pilot study and final experiments.

Table 5.1: Distribution of the participants per interface in the experiments

Interface	No. of participants in the experiment	No. of participants in the pilot study
Mobile textual interface	34	3
Mobile voice interface	34	3
Baseline desktop interface	34	3

5.2 Pilot Study Results

A pilot study was conducted with 3 participants from the first year Computer Science group to verify the feasibility of the experimental protocol. The data collected while users interacted with the interfaces during the pilot study showed that the experimental protocol was feasible. The participants were able to use all the three interfaces without any prior training within the 1 hour alloted time. The mobile voice interface was able to recognize spoken voice queries, confirming the venue was the right fit for the experiment. Additionally, users did not have any challenges completing the SUMI questionnaires, as in most cases they completed the questionnaires without any assistance.

The following sections discuss the results obtained from the pilot study for each interface, which further confirm the feasibility of the experimental protocol discussed in Chapter 4.

5.2.1 Baseline Desktop Interface

The feedback received through the SUMI questionnaire for the baseline interface shows that users had an understanding of what the baseline interface's purpose was. This is shown by the participants' comments on their reflections on using the interface, as shown in Table 5.2. It is clear from the participants' comments that they fully understood that the interface was used to search for isiXhosa information on the Internet. Further, the response of participant 3 shows that the participant had a deeper understanding of the purpose of the baseline interface.

Table 5.2: Answers to usage of the baseline desktop interface

What, in general, do you use this software for?	
Participant	**Comment**
1	To search for information from the internet
2	For searching for various information using isiXhosa language
3	For searching for information from a desktop computer and having the results returned in Xhosa

The results further show that the participants were able to establish a personal connection to the interface. This is supported by the fact that all 3 participants found the baseline interface to be very important when asked in the SUMI questionnaire, 'how important for you is the kind of software you have just been rating?', as shown in Table 5.3.

Table 5.3: Answers to the importance of the baseline desktop interface question

How important for you is the kind of software you have just been rating?	
Extremely important	3
Important	0
Not very important	0
Not important at all	0

5.2.2 Mobile Text Interface

The results obtained from the pilot study indicate that the participants were able to successfully evaluate the mobile text interface. The feedback received through the SUMI questionnaire for the mobile text interface shows that participants understood the purpose of the interface. This is shown by the participants' comments on their reflections on using the interface, as shown in Table 5.4.

Table 5.4: Answers to usage of the mobile text interface

What, in general, do you use this software for?	
Participant	**Comment**
1	To search for information from the internet
2	Searching for information
3	To search for information using isiXhosa

Although the participants were able to relate to the importance of the mobile text interface, it was interesting to notice that one participant did not find the mobile text interface as extremely important as the baseline interface, although the interfaces were similar, with the only difference being the mode of device used to access the interface. Table 5.5 shows that two participants found the interface to be extremely important while one participant found it to be important.

Table 5.5: Answers to the importance of the mobile text interface question

How important for you is the kind of software you have just been rating?	
Extremely important	2
Important	1
Not very important	0
Not important at all	0

5.2.3 Mobile Voice Interface

Similar to the mobile text interface and the baseline interface, the mobile voice interface was successfully evaluated by the participants in the study. The feedback obtained from the SUMI questionnaire for the mobile voice interface indicated that participants had an understanding of the mobile voice interface. As depicted in Table 5.6, all the 3 participants understood the usage of the mobile voice interface by answering that the interface is used to search for information on the Internet using voice.

Table 5.6: Answers to usage of the mobile voice interface

What, in general, do you use this software for?	
Participant	**Comment**
1	To voice-search for information from the internet
2	To search for information on the internet
3	To search for information using voice

Participants were able to relate to the importance of the mobile voice interface as all participants showed that the interface was important and, in some cases, extremely important to them. It is, however, interesting to notice that the 3 participants had different opinions as to the degree of importance of the voice interface with answers evenly distributed between extremely important, important and not very important. Similar to the mobile text and baseline interfaces, no participant found the voice interface to not be important at all, thus all users understood the importance and, therefore, the purpose of the interface, as reflected in Table 5.7.

Table 5.7: Answers to the importance of the mobile voice interface question

How important for you is the kind of software you have just been rating?	
Extremely important	1
Important	1
Not very important	1
Not important at all	0

The results from the final experiments are discussed in the next sections of this chapter.

5.3 Results Overview

The number of participants who participated in the final experiments and those who managed to submit their questionnaires is outlined in Table 5.8. Some participants failed to submit their questions due to failure of Internet connection during the submission time, as the questionnaires were administered online. The SUMI questionnaire asked participants 50 usability questions, which led to the analysis of the results by SUMI in the Global scale and the 5 subscales, namely: a) Efficiency, b) Affect, c) Helpfulness, d) Control and e) Learnability.

Table 5.8: Distribution of participants who submitted questionnaires per interface in the experiments

Interface	No. of participants in the experiment	No. of participants who submitted questionnaires
Mobile textual interface	34	32
Mobile voice interface	34	33
Baseline desktop interface	34	34

For each interface, the SUMI Global scale results are discussed first, which determine the overall satisfaction factor of the interface, followed by a break-down of the 5 subscales. The distribution of responses per the Global scale and the 5 subscales are in the groupings of below average, average and above average. Each interface is considered to be average if it rates between 40 and 60, below average if it rates less than 40 and above average if it rates greater than 60. Further, the statistical summary of the SUMI Global scale and the 5 subscales for each interface is discussed. The detailed participants' responses, which the statistical summary analysis is based on for each individual scale, is found in Appendices E.1, E.2 and E.3 for the 3 interfaces.

The speed results from the experiment are also discussed. As noted in Chapter 4, speed was monitored using application log files and speed was calculated from the time the participant submitted the voice or text query to the time they acted on the returned results. The distribution of speed for a given query was measured based on the individual data for the aforementioned query across all the 34 participants. The overall distribution of speed for each interface was then measured using all the individual speeds for all the queries across all the 34 participants. The raw speed data for the interfaces is shown in Appendices G.1, G.2 and G.3, with the related R script used to create the box-plots from the data.

The relevance results for the interfaces are further discussed. From Chapter 4, relevance was measured based on the number of voice/text results marked as relevant, that is, a tap on the tick button for the voice interface or clicks-throughs for the text interface. The overall relevance results for each interface were made up of combined results over all trials from all 34 participants. The results were not analysed on a per-query basis. Further, it was assumed that the more the results were marked as relevant or clicked the more it demonstrated that the given interface provided relevant results and, as such, effectively assisted participants in the search process. The results marked as relevant were captured in the application log files. The raw data used to determine whether the interfaces provided relevant results is shown in Appendices H.1, H.2 and H.3.

5.4 Evaluation of Mobile Voice Interface

5.4.1 Global

The Global scale for the voice interface had a rating of 49.70, as shown in Table 5.9. Further, the results from Figure 5.1 show that 60.6% of the participants rated the voice interface as average, 24.2% rated it as above average and 15.2% rated it as below average. These statistics indicate that the participants in the study found the interface to be satisfactory, as the Global scale rating (49.70) falls within average value range. Although the interface scored within the average rating, it can be noted from the results that 15.2% of the participants were somewhat unsatisfied with the usability of the interface. Although this may be true, it is clear that the views of the 84.8% participants who rated the voice interface as average and above average greatly influenced the final average rating for the mobile voice interface.

The strength and weakness analysis shows that the positive usability questionnaire statements received a 99% agree verdict while negative usability questionnaire statements received a 98% disagree verdict, as shown in Appendix F.2. Comparatively, the 5 subscales that make up most of the SUMI Global scale show that as a whole participants were satisfied with the interface, as discussed below.

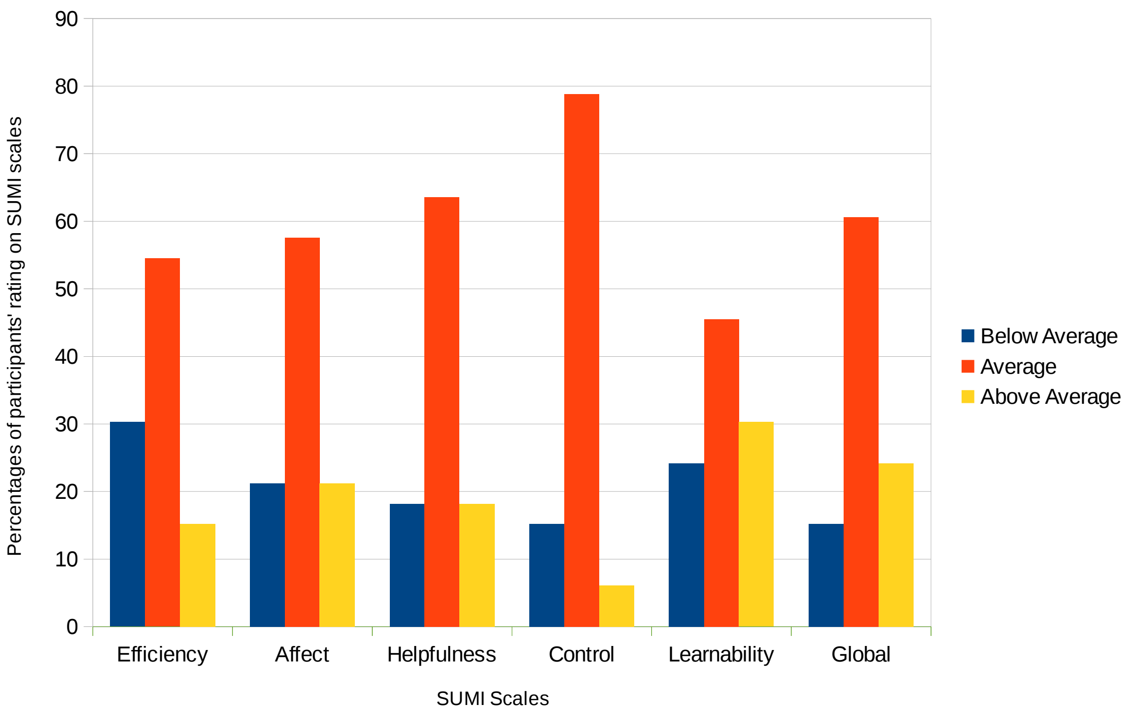


FIGURE 5.1. Distribution of responses on the Global scale and 5 SUMI subscales for the mobile voice interface

Table 5.9: A statistical summary of the SUMI results for the mobile voice interface

	Mean	St Dev	Median	IQR	Minimum	Maximum
Global	49.7	12.07	50	19	18	70
Efficiency	45.39	14.38	48	24.5	12	66
Affect	50.06	12.52	52	19.5	22	72
Helpfulness	52.12	9.88	53	11.5	30	70
Controllability	49.58	8.53	50	12	30	65
Learnability	51.27	11.87	54	22	23	66

5.4.2 Efficiency

From Figure 5.1, 54.5% of the participants rated the interface's Efficiency as average while 15.2% of the participants rated it as above average. Similarly, the statistical analysis in Table 5.9 show that the interface scored a mean value of 45.39 for the Efficiency subscale, with most of the participants' ratings ranging between 12 and 66. These statistics show that most of the participants were happy with the extent to which the interface assisted them in their search.

However, as seen from Figure 5.1, 30.3% of the participants felt the interface's Efficiency rated below average. Even so, the overall Efficiency score of 45.39 was influenced by the views of most participants who rated the interface's Efficiency as average and as above average. In fact, most of the participants, when asked, 'what do you think is the best aspect of this software, and why?', responded thus:

'It is very fast'

'It works very fast and good.'

'The speed is great!'

'Not having to type anything but just using my voice to search makes searching for information more convenient and easy.'

As much as the participants were happy with the rate at which the interface assisted them in their search, some participants (30.3%, see Figure 5.1), notably pointed out that the weakness of the interface is that it responds slowly to inputs, does not favour straight-forward execution of tasks, does not support searching as expected and that it is sometimes inconsistent, as shown by the high percentage of responses to the following questionnaire statements:

a) This software responds too slowly to inputs; verdict: agree, strength: 98%.

b) Tasks can be performed in a straight forward manner using this software; verdict: disagree, strength: 97%.

c) The software hasn't always done what I was expecting; verdict: agree, strength: 61.7%.

d) I think this software is inconsistent; verdict: agree, strength: 99.9%.

However, this was perhaps to be expected as in most cases users, while interacting with voice interfaces, often attempt to speak queries in a conversational style as they would to a fellow human being with the expectation to be easily understood as fellow humans would. This statement is supported by a study that points out that users interacting with voice systems are often not aware of the limitations of such systems, and in turn use a conversational style, which voice systems then fail to recognize [13]. This point of view is further reinforced by the participants' responses to the SUMI questionnaire statement 'what do you think needs most improvement, and why?', which shows that most participants were indeed not pleased by the rate at which the voice queries were detected by the interface:

'The problem with it is not being able to identify what I have said because of the speed at which I am speaking'

'It needs to be able to pick up different accents and intonation of voices'

'Improved voice detector and also correct words because sometimes it gives words that are out of context'

'The recognizer of the voice can be improved because it is slow'

5.4.3 Affect

According to the results in Figure 5.1, most of the participants were pleased with the interface, with 57.6% of the participants rating the interface's Affect as average and 21.2% rating it as above average. Most of the participants recorded ratings that fell within the average range of 40 and 60, which resulted in an overall Affect rating of 50.06, as shown in Table 5.9. Further, this is also confirmed by the high percentage number of responses to the following questionnaire statements:

a) Working with this software is mentally stimulating; verdict: agree, strength: 86.5%.

b) I enjoy the time I spend using this software; verdict: agree, strength: 86.6%.

Regardless, a reasonable number of participants (21.2%) felt the interface was somewhat awkward and frustrating leading to a tense feeling while interacting with it, and as such, rated it as below average. This is also shown by the high percentage of responses to the following questionnaire statements:

a) Using this software is frustrating; verdict: agree, strength: 69.1%.

b) Working with this software is satisfying; verdict: disagree, strength: 94.9%.

c) There have been times in using this software when I have felt quite tense; verdict: agree, strength: 78.4%.

d) This software is really very awkward; verdict: agree, strength: 97.9%.

In addition to the above, some participants when asked, 'what do you think needs most improvement, and why?', responded thus:

'The lady does not have a warm voice, she makes the software less friendly'

'The voice comes too strong and can be disruptive at times'

'Navigation from results screen back to the home page and the error feedback is annoying'

The comments above suggest that the participants feel there is a need to change the tone of the feedback speech on the interface by ensuring consistent volume levels on speech sounds. Additionally, the above results show that the navigation from the results menu back to the query submission interface should be redesigned, based on these comments. The error feedback, that is, the message that informs the user that the interface did not find any relevant results for a given voice query also needs improvement to avoid annoying users, especially, in the case where the interface fails to detect the spoken query in more than 3 attempts.

5.4.4 Helpfulness

From Figure 5.1, 63.6% of the participants rated the interface's Helpfulness as average while 18.2% rated it as above average. Most of the participants in the study found the voice interface to be self-explanatory, with enough help facilities. This is evidenced by the Helpfulness subscale mean value of 52.12, as shown in Table 5.9, which further proves the helpfulness of the interface. The strength and weakness analysis results show that most participants indeed agree that the mobile voice interface was helpful:

a) The instructions and prompts are helpful; verdict: agree, strength: 97.8%.

b) The organisation of the menus seems quite logical; verdict: tend to agree, strength: 90.6%.

c) The software has helped me overcome any problems I have had in using it; verdict: slightly agree, strength: 93.9%.

Despite most participants finding the interface helpful, as pointed out above, 18.2% of the participants felt the information given on the interface was not very helpful, the interface itself was not clear, there was never enough information on the interface when needed, and they could not understand and act on the information provided by the interface. This is indicated by the negative number of responses to the following questionnaire statements:

a) I find that the help information given by this software is not very useful; verdict: agree, strength: 97.2%.

b) There is never enough information on the screen when it's needed; verdict: agree, strength: 97.5%.

c) I can understand and act on the information provided by this software; verdict: slightly disagree, strength: 72.8%.

The participants' responses to the above questionnaire statements suggest that the way the help menu was designed was somewhat not helpful and, as such, needs to be redesigned to better support users. Moreover, the prompting speech instructions that inform the user to perform a given task on the interface needs to be reviewed and redesigned to better accommodate users.

5.4.5 Controllability

Figure 5.1 shows that 78.8% of the participants rated the interface's ability to be controlled as average while 6.1% of the participants rated it as above average. Moreover, the voice interface scored a mean value of 49.58 for the Control subscale. These statistics indicate that most of the participants felt in command of the interface. This is further supported by the strength and weakness analysis, which shows that the participants felt it was easier to make the voice interface do exactly what they wanted and that they felt safer using fewer functions of the interface that they were familiar with, as highlighted below:

a) I feel safer if I use only a few familiar functions; verdict: agree, strength: 99.9%.

As seen from Figure 5.7, 15.2% of the participants rated the interface's ability to be controlled as below average. This could potentially explain why some of the participants felt safer using fewer functions of the interface, as illustrated in point a) above. Additionally, most of the participants stated that the interface was awkward in the case where they tried different search queries from the ones they were given; as such more words should be supported by the interface. This is verified by the high percentage for the agree verdict to the questionnaire statement:

a) This software is awkward when I want to do something which is not standard; verdict: agree, strength: 91.2%.

5.4.6 Learnability

As seen from Figure 5.1, 45.5% of the participants rated the interface's Learnability as average while 30.3% rated it as above average. Similarly, the results from Table 5.9 show that most of the participants were able to learn how to use the voice interface features with an average value of 51.27 on the Learnability subscale. These statistics demonstrate that the voice interface accommodated 75.8% of the participants as they were able to learn how to use most of its features. Some participants, when asked, 'what do you think is the best aspect of this software, and why?', responded thus:

> 'The instructions, it directs and tells you exactly what you must do and how'

However, 24.2% of the participants rated the Learnability subscale of the interface below average. This is because these participants felt learning how to use the mobile voice interface at first is full of problems, it takes too long to learn some of the functions, and they had to look for assistance most of the times when using the interface. This is supported by the high negative number of responses to the following questionnaire statements:

a) Learning to operate this software initially is full of problems; verdict: agree, strength: 88.6%.

b) It takes too long to learn the software functions; verdict: agree, strength: 77.6%.

c) Learning how to use new functions is difficult; verdict: agree, strength: 93%.

d) I have to look for assistance most times when I use this software; verdict: agree, strength: 99.3%.

In addition to the above responses, some participants when asked by the SUMI questionnaire, 'what do you think needs most improvement, and why?', responded:

> 'There should be more clearer helping buttons on the screen of this software'

The above comment shows that the design of the help menu needs to be reviewed and redesigned to better help the users learn how to use the interface without much hassle.

5.4.7 Speed Results

Figure 5.2 shows a box-plot of the speed results for the mobile voice interface. The box-plot for the 10 voice queries shows that the queries had median values ranging between 6.5 and 10 seconds.

The interface recorded poor speed for voice queries *isiXhosa* and *emzantsi afrika*, with median values of 10 and 9 respectively. This made sense as most participants were eager to find out more information about their language and country and, as such, spent more time assessing the returned voice results before making any further interaction with the interface.

In essence, the overall medians are not far apart for the 10 voice queries shown in the box-plot in Figure 5.2. This suggests that most of the time the voice interface took the same amount of time to provide participants with the voice results and the participants took about the same amount of time to interact with the voice interface. The speed of the voice interface is further corroborated by the positive comments of the participants when asked by the SUMI questionnaire, 'what do you think is the best aspect of this software, and why?':

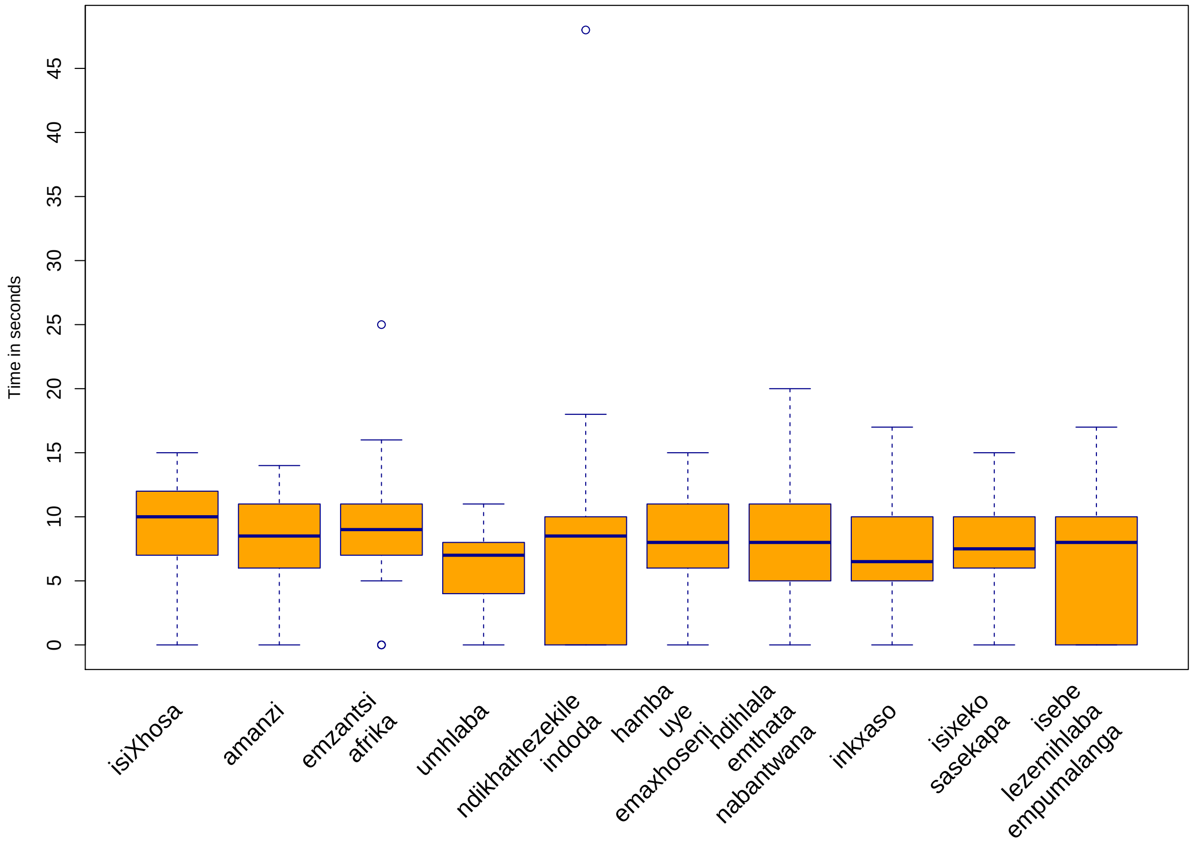


FIGURE 5.2. The distribution of speed per voice query on the mobile voice interface

'It is very fast'

'It works very fast and good'

'It works fast and says information fast'

'The speed is great'

5.4.8 Relevance Results

The relevance results for the mobile voice interface are presented in Figure 5.3. As can be noted from Figure 5.3, the mobile voice interface received a high total number of no clicks-throughs on the voice results with an overall total of 165. This shows that the participants did not find relevant information for most of the 10 voice queries.

In most cases, participants in the study did not find relevant information as the voice queries were not recognized by the voice interface. This could be mainly due to the fact that the participants in some cases were speaking faster than the interface could detect, as suggested by the participants' comments when asked, 'what do you think needs most improvement, and why?':

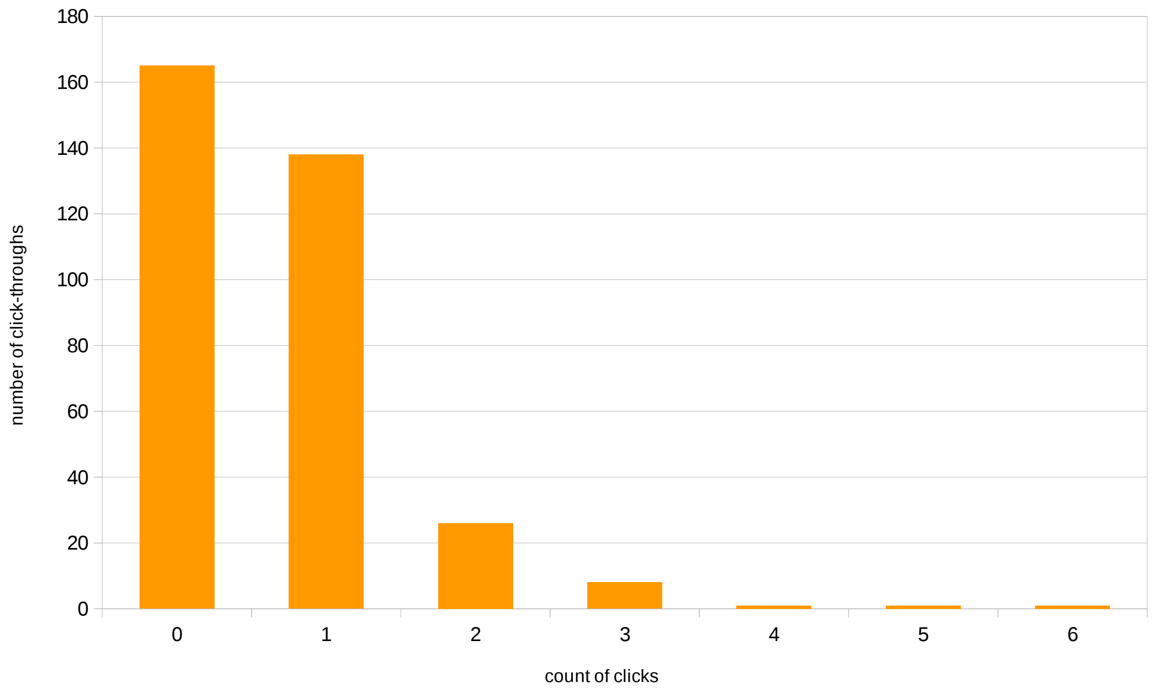


FIGURE 5.3. Distribution of clicks-throughs on the mobile voice interface

> 'The problem with it is not being able to identify what I have said because of the speed at which I am speaking'

Despite the high number (161) of no clicks-throughs on the voice results, the interface has proven to have recognized some voice queries and provided relevant results for such. This is shown in Figure 5.3, where 138 single clicks-throughs were made on the returned voice results. This suggests that, given the 10 voice queries, some participants found at least one relevant voice result per voice query. Additionally, the same figure shows that some participants were able to find 2 relevant voice results for some search queries. Better still, some participants found 3 to 6 relevant voice results for some voice queries. These statistics show that indeed the mobile voice interface was considered relevant by participants in the study as it was able to provide relevant results, and participants were able to identify these results and act on them.

5.5 Evaluation of Baseline Interface

5.5.1 Global

The baseline interface mostly falls within average and above average on the Global scale and the 5 subscales, as depicted in Figure 5.4. Additionally, the interface was considered to be average by 52.9% of the participants while 47.1% of the participants felt the interface was above average. The results in this case show that the usability of the interface was considered satisfactory by the participants in the study. Correspondingly, the statistical summary analysis of the results confirms the interface to be average, with a rating of 59.12 on the Global scale, with most ratings ranging between 46 and 70, as shown in Table 5.10. Further evidence provided by the strength and weakness analysis of the interface shows that the positive usability questionnaire statements received a 99% agree verdict while the negative usability questionnaire statements received a 99% disagree verdict across the 5 subscales, as shown in Appendix F.3.

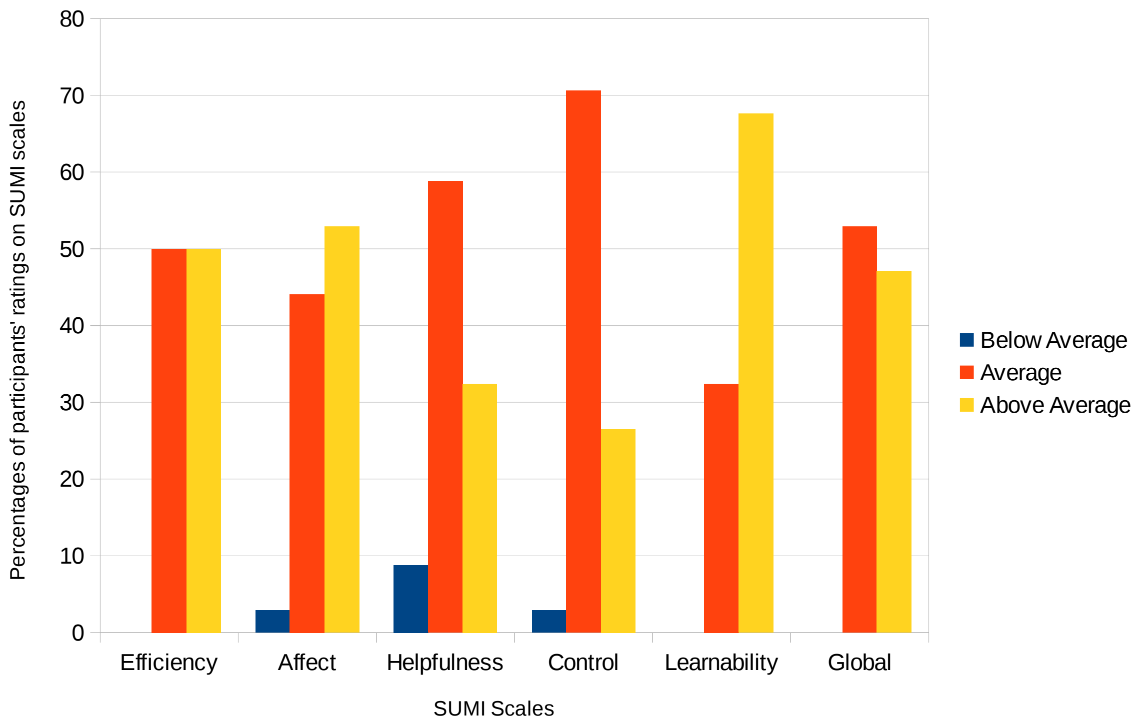


FIGURE 5.4. Distribution of responses on the Global scale and 5 SUMI subscales for the baseline desktop interface

Table 5.10: A statistical summary of the SUMI results of the baseline desktop interface

	Mean	St Dev	Median	IQR	Minimum	Maximum
Global	59.12	6.7	60	7	46	70
Efficiency	58.74	8.74	61.5	11	40	69
Affect	60.32	9.5	62	12	33	72
Helpfulness	55.74	8.33	57	9	31	70
Controllability	55.35	7.01	56.5	11	39	68
Learnability	61.68	6.13	64	7	44	71

5.5.2 Efficiency

50% of the participants rated the interface's Efficiency as average while 50% of them rated it as above average, as demonstrated in Figure 5.4. Similarly, the statistical analysis in Table 5.10 shows that the interface scored a mean value of 58.74 for the Efficiency subscale, with most of the participants' ratings ranging between 40 and 69. These statistics prove that most of the participants in the study were happy with the efficiency of the baseline interface. This sentiment is supported by the participants' exact responses when asked, 'what do you think is the best aspect of the software, and why?':

'Its efficiency, sleek, familiar and simple design as well as the ability to search for information in isiXhosa'

'The ability to give feedback fast'

'It is fast'

'The speed and ability to give results in isiXhosa.'

Further proof that the baseline interface was efficient is reflected by the high percentage number of responses to the following questionnaire statements:

a) This software responds too slowly to inputs; verdict: disagree, strength: 99.2%.

b) It is obvious that user needs have been fully taken into consideration; verdict: agree, strength: 99.6%.

c) There are too many steps required to get something to work; verdict: disagree, strength: 99.9%.

d) This software occasionally behaves in a way which can't be understood; verdict: disagree, strength: 99.9%.

5.5.3 Affect

According to the results in Figure 5.4, 52.9% of the participants rated the interface's Affect as above average while 44.1% felt the interface was average. Moreover, the summary statistics show that the participants rated the Affect subscale between 33 and 72, resulting in a mean value of 60.32, as shown in Table 5.10. These statistics show that the participants enjoyed the time they spent using the interface, and felt the interface was attractive and satisfying, as highlighted by the high percentage agree verdict on the following questionnaire statements:

a) I enjoy the time I spend using this software; verdict: agree, strength: 90.6%.

b) Working with this software is mentally stimulating; verdict: agree, strength: 95.5%.

c) The software presents itself in a very attractive way; verdict: agree, strength: 53.8%.

In spite of the above Affect subscale ratings, one participant (2.9%) felt the interface was below average. Although this number is relatively small, it still suggests that the interface needed improvement, as reflected by the participants' comments when asked, 'what do you think needs most improvement, and why?':

'It's aesthetic and general design of the initial search page'

'There is limited information and options that appear when you search. Some of them are not even relevant or related to the content covered in the word or work you are researching'

5.5.4 Helpfulness

The results shown in Figure 5.4 show that 58.8% of the participants rated the interface's Helpfulness as average while 32.4% rated the Helpfulness of the interface as above average. As outlined in the statistical analysis of the interface in Table 5.10, the Helpfulness subscale scored a mean value of 55.74, with most of the participants' ratings ranging between 31 and 70. The evidence these results provide is that the interface was found by the participants in the study to be self-explanatory with enough help facilities, and that the information provided was clear, understandable and helpful. The strength and weakness analysis of the interface further shows that, indeed, the participants were happy with the helpfulness of the interface, as shown by the high percentage number of responses to the following SUMI questionnaire statements:

a) The instructions and prompts are helpful; verdict: agree, strength: 95.3%.

b) The way that system information is presented is clear and understandable; verdict: agree, strength: 99.9%.

c) I find that the help information given by this software is not very useful; verdict: disagree, strength: 69.1%.

d) It is easy to see at a glance what the options are at each stage; verdict: agree, strength: 84.1%.

5.5.5 Control

As can be seen from Figure 5.4, 70.6% of the participants rated the interface's Control as average and 26.5% of the participants rated it as above average. From Table 5.10, it can be seen that the interface received a mean value of 55.35 for the Control subscale, with most of the participants' ratings ranging between 39 and 68. These statistics show that most of the participants in the study felt in command of the interface and had no problems making it do exactly what they wanted it to do. This is shown in the strength and weakness analysis of the interface, where the following questionnaire statements received high percentage responses:

a) I feel in command of this software when I am using it; verdict: agree, strength: 97.1%.

b) It is easy to make the software do exactly what you want; verdict: agree, strength: 96.9%.

c) It is relatively easy to move from one part of a task to another; verdict: agree, strength: 99.6%.

d) This software has at some time stopped unexpectedly; verdict: disagree, strength: 99.9%.

Even though most of the participants felt in control of the interface, one participant (2.9%) rated the interface's Control as below average. This signals that some improvements need to be made to the interface to help the participant to better command the interface. When asked, 'what do you think needs most improvement, and why?', the participant responded thus:

'Perhaps less key strokes'

The comment above reflects that the interface was somewhat not supporting the participant to be more economical with keystrokes, and therefore this needs some consideration. It is clear from this comment that the interface needed an auto-complete feature on the search box, which was disabled to maintain consistency between the three interfaces, as discussed in Chapter 3. This view is supported by a study that points out that auto-complete helps make the interface a better input feature as it provides information to the user as they type, thus allowing users to become economical with keystrokes [70].

5.5.6 Learnability

From the results in Figure 5.4, 67.6% of the participants rated the interface's Learnability as above average while 32.4% rated the interface's Learnability as average. Similarly, the summary statistics in Table 5.10 show that the interface had a mean value of 61.68 on the Learnability subscale. Indeed, these statistics demonstrate that most of the participants were able to learn how to use the interface features without hassle. As can be seen from the strength and weakness analysis in Appendix F.3, the following questionnaire statements received high positive percentage responses, proving the interface was learnable:

a) Learning to operate this software initially is full of problems; verdict: disagree, strength: 99.9%.

b) It takes too long to learn the software functions; verdict: disagree, strength: 97.8%.

c) Learning how to use new functions is difficult; verdict: disagree, strength: 94.4%.

d) I will never learn to use all that is offered in this software; verdict: disagree, strength: 90%.

5.5.7 Speed Results

Figure 5.5 shows the box-plot of the speed results for the baseline interface. The results show that the 10 text queries had median values ranging between 3.5 and 7.5 seconds. Furthermore, the queries *umhlaba* and *inkxaso* recorded the least time compared to other queries. This suggests that either the interface was very fast to process these queries or there were not many relevant results for these queries. Additionally, the results show that the text query *ndihlala emthatha nabantwana* had the lowest median value of 3.5 seconds while the text query *isiXhosa* had the highest median of 7.5 seconds.

The interface's time to provide the results and such results being acted on by the participants has proven to be consistent across all the 10 search queries. This is shown in Figure 5.5, where the medians are not far apart for most of the 10 search queries. Also, this shows that most of the time the baseline interface took the same amount of time to provide participants with the text results and the participants took about the same amount of time to interact with the baseline interface.

The speed of the baseline interface is further supported by the positive comments of the participants when asked by the SUMI questionnaire, 'what do you think is the best aspect of this software, and why?':

'It is fast.'

'The speed and ability to give results in isiXhosa.'

'The software picks up words faster. I did not have any problems. It is also easy to use.'

5.5.8 Relevance Results

The baseline interface received the most single clicks-throughs compared to as on the mobile interfaces on the search results, with an overall total of 198 single clicks-throughs, as shown

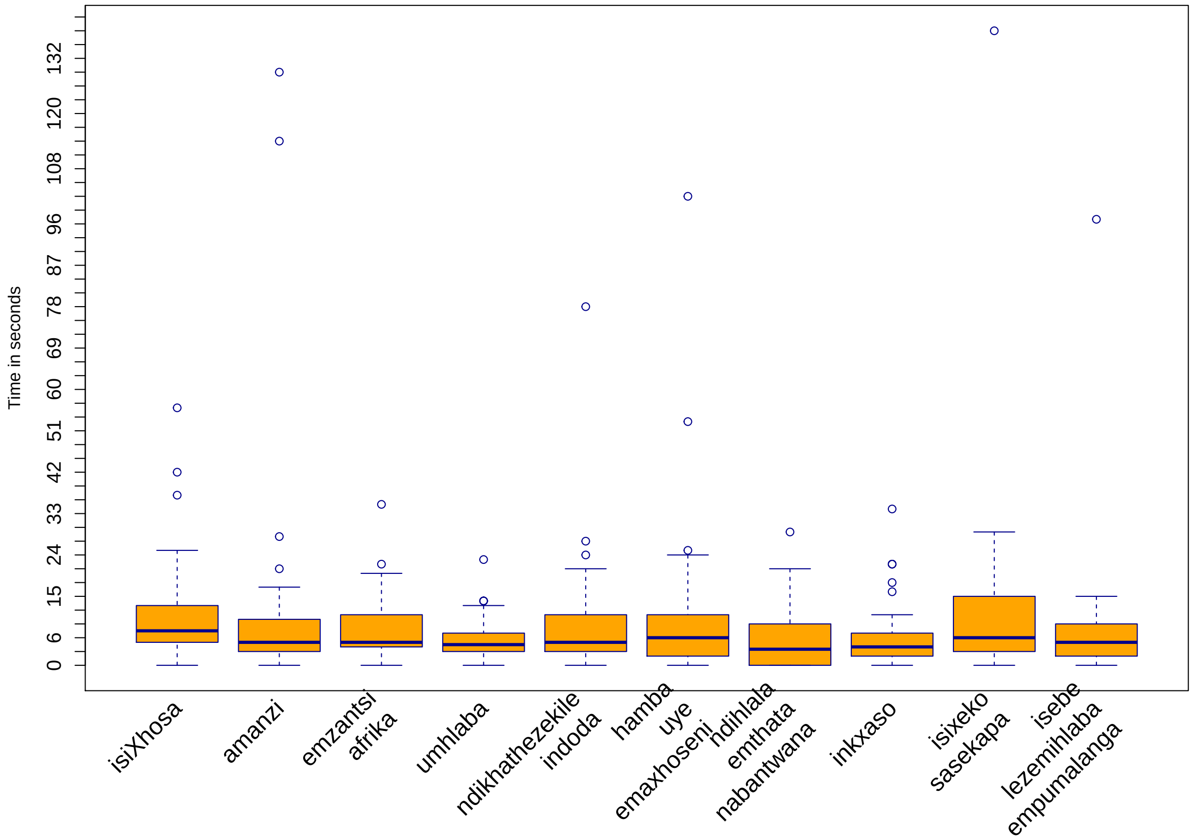


FIGURE 5.5. The dispersion of speed per text query for the baseline desktop interface

in Figure 5.6. This shows that, given the 10 search queries, most of the participants viewed a single result per search query. There were fewer zero clicks-throughs (62) compared to the mobile voice interface (165), which suggests that in most cases the interface provided relevant search results. However, these statistics can also suggest that due to a larger screen display the participants decided to explore more than one result as it was easy to navigate from one result to another. This is shown by the high number of double clicks-throughs, which is 47, as shown in Figure 5.6. Additionally, more participants viewed more than 3 results with a total number of 21 clicks-throughs, 4 results with a total number of 7 clicks-throughs and 5 results with a total number of 4 clicks-throughs.

The above statistics demonstrate that the baseline interface provided relevant results, which participants successfully accessed. Even so, some recommendations were made to help improve the baseline interface to give more relevant search results. This is apparent in the participants' verbatim responses when asked, 'what do you think needs most improvement, and why?':

> 'The information that is supplied needs to be relevant to what the user is searching.'

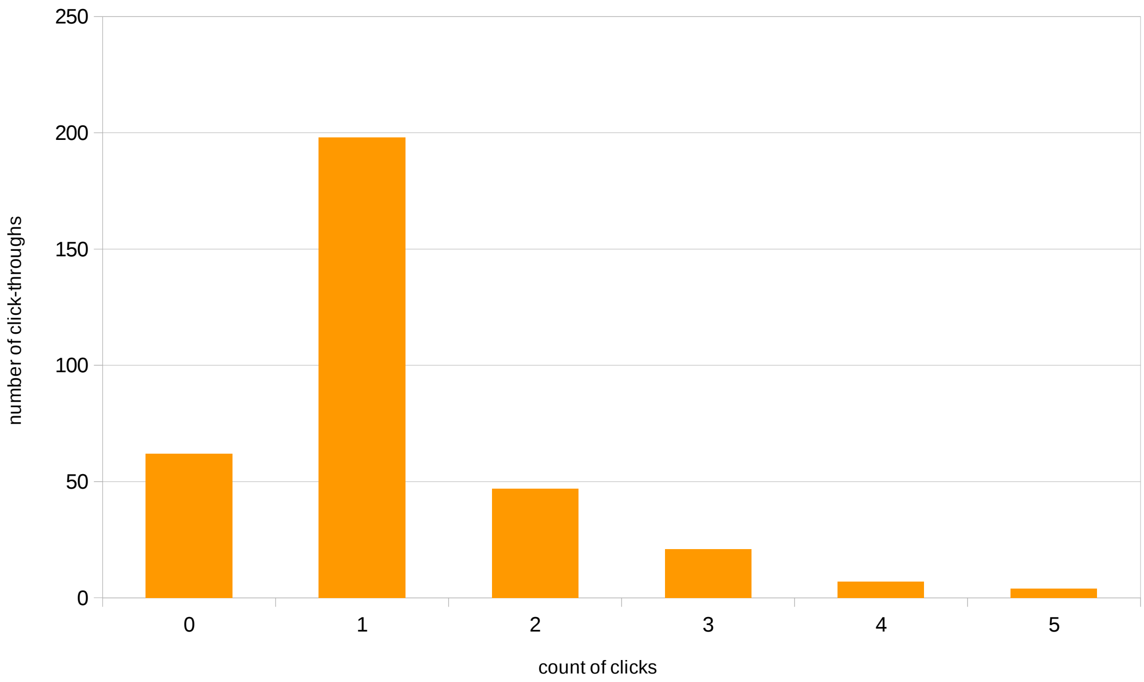


FIGURE 5.6. Distribution of clicks-throughs on the baseline desktop interface

'The accuracy of results. Because some of its results are irrelevant to the words which I am searching for.'

5.6 Evaluation of Mobile Text Interface

5.6.1 Global

The Global scale rating provides the general usability/satisfaction score (58.19) of the mobile text interface[1] which, according to the results in Table 5.11, falls within the average value range. Further, according to the results in Figure 5.7, the mobile text interface falls within average and above average on the Global scale and the 5 subscales. It is evident from Figure 5.7 that 56.3% of the participants felt the interface rated above average while 40.6% of the participants felt the interface was average. These figures show that indeed the usability of the mobile text interface was considered satisfactory by at least 96.9% of the participants who rated the interface as average and as above average. Furthermore, the strength and weakness analysis of the interface shows that the positive usability questionnaire statements received a 99% agree verdict from the participants in the study. The negative usability questionnaire statements received a 99%

[1] `http://sumi.ucc.ie/sumipapp.html#sumidev`

disagree verdict across all the 25 questionnaire statements that measure the Global scale, as shown in Appendix F.1.

A small percent of participants (3.1%) felt the interface was below average, as shown in Figure 5.7. Even so, the interface has proven satisfactory to the majority of the participants, given the previously discussed statistics. Furthermore, the 5 subscales that make up most of the Global scale show that indeed participants were satisfied with the interface, as discussed below.

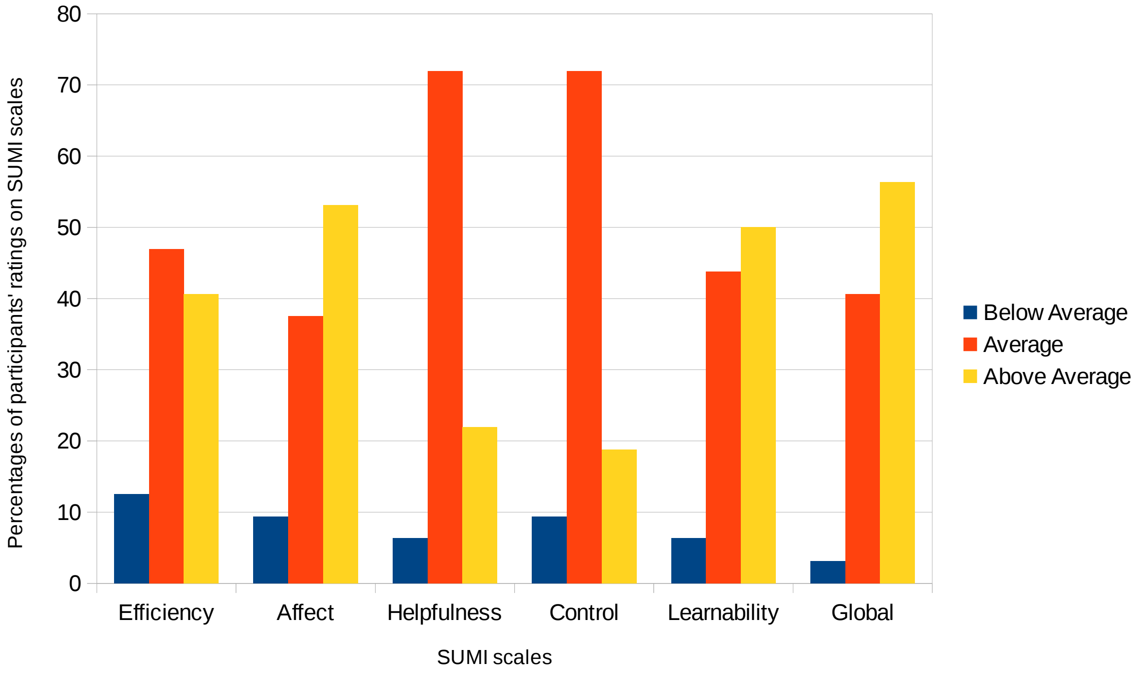


FIGURE 5.7. Distribution of responses on the Global scale and 5 SUMI subscales for the mobile text interface

Table 5.11: A statistical summary of the SUMI results for the mobile text interface

	Mean	St Dev	Median	IQR	Minimum	Maximum
Global	58.19	8.14	61.5	12.5	38	68
Efficiency	56.91	12.24	59.5	19.5	28	72
Affect	58.75	11.21	64	16	30	72
Helpfulness	54.22	8.31	57	11.5	29	64
Controllability	53.72	10	58	10.5	23	68
Learnability	57.91	9.21	60.5	9.5	32	71

5.6.2 Efficiency

As shown in Figure 5.7, 40.6% of the participants rated the interface's Efficiency as above average while 46.9% rated it as average. In like manner, the statistical analysis in Table 5.11 shows that the interface scored a mean value of 56.91 for the Efficiency subscale, with most of the participants' ratings ranging between 28 and 72. These statistics show that most of the participants in the study were happy with the extent to which the interface assisted them in their search. This is supported by the positive responses by the participants in the study to the SUMI questionnaire statement, 'what do you think is the best aspect of this software, and why?':

'It is very simple for everyone to use even the people who are not that technologically literate'

'It's portable and easy to use'

'The speed and ability to give results in isiXhosa'

'It is fast. when you are searching for something you want it to appear quickly so this software is fast enough'

'Ease of use and the interface is simple but also contains enough information to use effectively'

'The best aspect is that it is very simple and straight forward to use'

'It is very fast'

Despite most of the participants giving positive feedback regarding the ability of the interface to assist them in their search, 12.5% of the participants felt the interface's efficiency was below average. As such, some recommendations were made by the participants to help improve the efficiency of the interface:

'I think it needs more information and definitions of some words'

'There are duplicates of results'

'Just the content needs to improve, it runs rather smoothly'

It is clear from the above suggestions that the interface needed more content to better support participants' search needs. Indeed, this is correct as isiXhosa is a low resourced language, therefore it has few resources available online to utilize in a project such as of this study. In addition to this, it is evident that the search engine behind the interface needed improvement as some of the results were duplicates, thus affecting the participants' efficiency assessment of the interface. Because of limited isiXhosa documents online, the search engine crawled most documents from the South African government's isiXhosa parliament speeches and the isiXhosa Wikipedia. This resulted in duplicates as Wikipedia maintains several versions of one document, which the search engine crawled.

5.6.3 Affect

For this subscale, the interface scored a mean rating of 58.75, as shown in Table 5.11. Additionally, 53.1% of the participants rated the interface's Affect as above average, 37.5% rated it as average and 9.4% rated it as below average. The results for the Affect subscale were strongly influenced by the views of the 53.1% and 37.5% of the participants, who rated the interface's Affect as above average and as average respectively. In addition to this, the Affect subscale rating was influenced by the large percentage of responses to the following SUMI questionnaire statements:

a) I think this software has sometimes given me a headache; verdict: disagree. strength: 98.7%.

b) This software is awkward when I want to do something which is not standard; verdict: disagree. strength: 98.8%.

c) There have been times in using this software when I have felt quite tense; verdict: disagree, strength: 93.7%.

d) Using this software is frustrating: verdict: disagree, strength: 93.4%.

Some suggestions were made, which aimed to help improve the interface in regard to the Affect subscale. The participants, when asked 'what do you think needs most improvement, and why?', responded to the question by suggesting that the interface should include a clear back button to the search menu. This was suggested as the interface seemed inconsistent when the participant moved from the results menu back to the search menu. This is shown by the participants' verbatim responses:

> 'I think there should be a 'back to search' button. There is something inconsistent about the interface when one moves from the 'search' page'

Indeed, this suggestion is valid as the back button was disabled in order to force participants to use the device's back button in order to track the amount of time the participant spent on a single result, as discussed in Chapter 3. The design to exclude a clear back button therefore proved to make the interface inconsistent with the common mobile application designs, and therefore inappropriate.

5.6.4 Helpfulness

From Figure 5.7, 71.9% of the participants rated the interface's Helpfulness as average while 21.9% of the participants rated it as above average. The statistical analysis in Table 5.11 show that the interface scored a mean value of 54.22, with most of the ratings ranging between 29 and 64. These statistics demonstrate that the interface was considered by 93.8% of the participants to be helpful. Nonetheless, 6.3% of the participants felt the interface was somewhat not helpful.

Even so, it is clear from the statistics that the interface's final rating was influenced by the majority of the participants who found it to be helpful. Further evidence, which proves the interface was helpful to the participants in the study, is displayed by the high percentage agree verdict from the participants in regard to the following questionnaire statements:

a) It is obvious that user needs have been fully taken into consideration; verdict: agree, strength: 99.9%.

b) I keep having to go back to look at the guides; verdict: disagree, strength: 99.6%.

c) The instructions and prompts are helpful: verdict: agree, strength: 98.3%.

d) The way that system information is presented is clear and understandable; verdict: agree, strength: 98.2%.

e) The software has helped me overcome any problems I have had in using it; verdict: agree, strength: 93.5%.

5.6.5 Control

As shown in Figure 5.7, 71.9% of the participants rated the interface's ability to be controlled as average, 18.8% of the participants rated it as above average and 9.4% of the participants rated it as below average. From Table 5.11, the average Control rating is 53.72. These statistics show that most of the participants in the study felt in command of the interface. This is further represented by the strength and weakness analysis of the interface, where the following questionnaire statements received a high percentage of positive responses:

a) I feel in command of this software when I am using it; verdict: agree, strength: 83.8%.

b) It is easy to make the software do exactly what you want; verdict: agree, strength: 99.9%.

c) This software is awkward when I want to do something which is not standard; verdict: disagree, strength: 98.8%.

d) I feel safer if I use only a few familiar functions; verdict: agree, strength: 99.9%.

e) I prefer to stick to the functions that I know best; verdict: agree, strength: 99.9%.

5.6.6 Learnability

The results from Figure 5.7 show that 50% of the participants rated the interface's Learnability as above average, 43.8% rated it as average and 6.3% rated it as below average. Identically, the results from Table 5.11 show that most of the participants were able to learn how to use the interface features, with an average value of 57.91 on the Learnability subscale. These statistics demonstrate that the interface accommodated most of the participants as they were able to learn how to use most of its features. Furthermore, the strength and weakness analysis shows that most participants felt the interface was straight-forward, as shown by the high percentage of positive responses to the following questionnaire statements:

a) I will never learn to use all that is offered in this software; verdict: disagree, strength: 99%.

b) I keep having to go back to look at the guides: verdict; disagree, strength: 99%.

c) Learning to operate this software initially is full of problems; verdict: disagree, strength: 81.8%.

In addition to the above, most of the participants when asked by the SUMI questionnaire, 'what do you think is the best aspect of this software, and why?', responded that the interface was straight-forward and easy to use:

'It is very simple for everyone to use even the people who are not that technologically smart'

'The best aspect is that it is very simple and straight forward to use.'

5.6.7 Speed Results

Figure 5.8 shows a box-plot of the speed results for the mobile text interface. According to the results, the text queries had median values ranging between 3.5 to 5.5 seconds. Additionally, the results from the box-plot suggest that the text interface's time to provide search results and such results being acted on by the participants was consistent across all the 10 search queries as the medians are not far apart, as shown in Figure 5.8.

Furthermore, the results also show that the text query *isiXhosa* recorded the most time, with its time ranging between 0 and 18 seconds, excluding the 3 outliers. This made sense as most participants were eager to find out more information about their language and culture, so spent time assessing the returned results before making any further interaction with the interface. Interesting enough, similar findings were discovered on the mobile voice interface, as discussed in Section 5.4.7.

The results also suggest that the interface and, in turn, the participants took less time on the text queries *ndihlala emthatha nabantwana* and *isixeko sasekapa*. This suggests that the interface did not provide sufficiently many relevant results for the participants to explore. This is reflected by the participants' responses to the SUMI questionnaire statement, 'what do you think needs most improvement, and why?', where participants responded:

'Search options should be improved, information is limited and sometimes irrelevant'

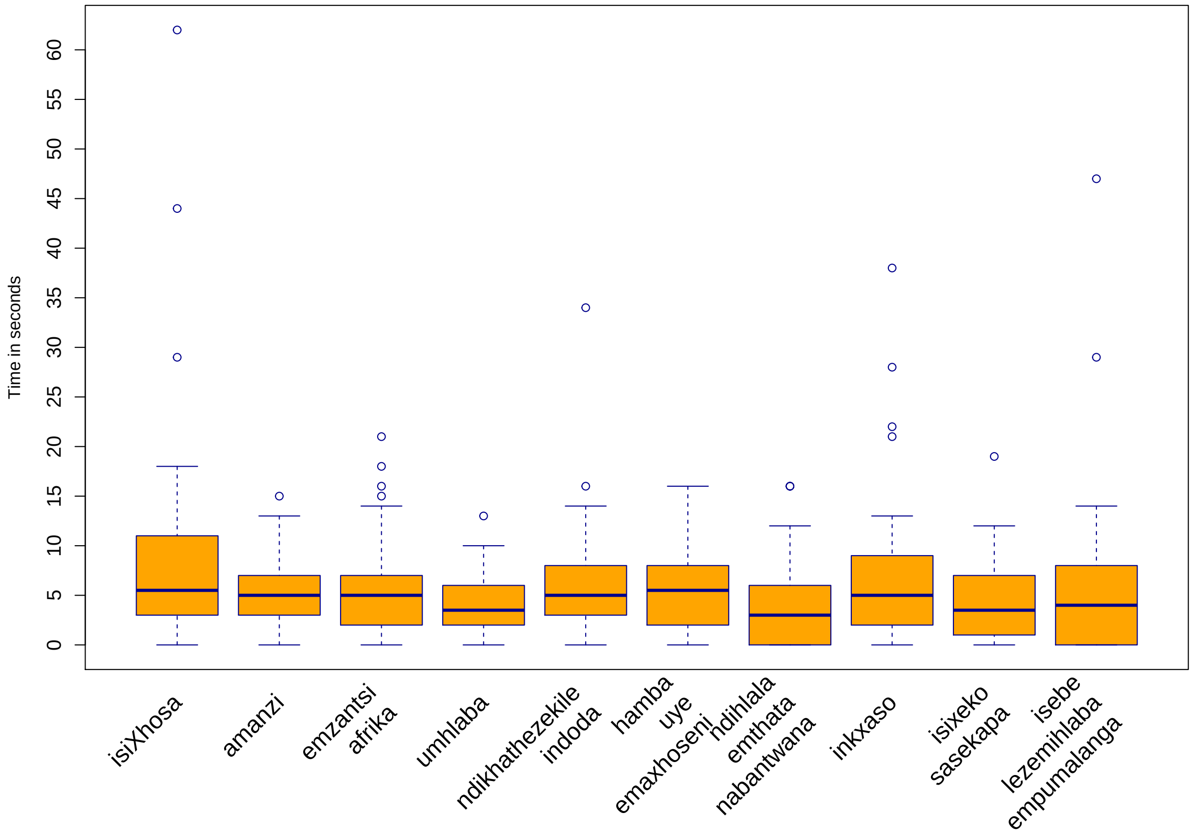


FIGURE 5.8. The distribution of speed per text query on the mobile text interface

> 'The accuracy of results. Because some of its results are irrelevant to the words which I am searching for.'
>
> 'Quality and the relevance of the search results'

Additional analysis of the speed results show that most participants were happy with the pace at which the interface presented itself to assist them on the search process. Most of the participants, when asked about the best aspect of the interface by the SUMI questionnaire, stated that the interface was fast, some with more emphasis. This can be seen by their precise comments on the speed of the interface when asked by the SUMI questionnaire, 'what do you think is the best aspect of this software, and why?':

> 'It was very fast and easy to use. It loads quickly which isn't always the case for mobile apps. Given how familiar I already am to popular search engines like Google, I adapted very quickly to this interface'
>
> 'Mobile app = accessible and convenient. Works fast without hassles'
>
> 'It is very fast'

'The speed, it is nice to get the results as soon as possible and this software does exactly that as it is fast enough to retrieve information'

5.6.8 Relevance Results

Figure 5.9 shows the relevance results for the mobile text interface. According to these results, most participants viewed only one result per given query, with a total count of 171 clicks. These statistics suggest that the interface provided relevant results given the 10 text queries as participants viewed one result and then changed the search query. However, this could also suggest that the interface might have not provided sufficient relevant results per query for the participants to explore.

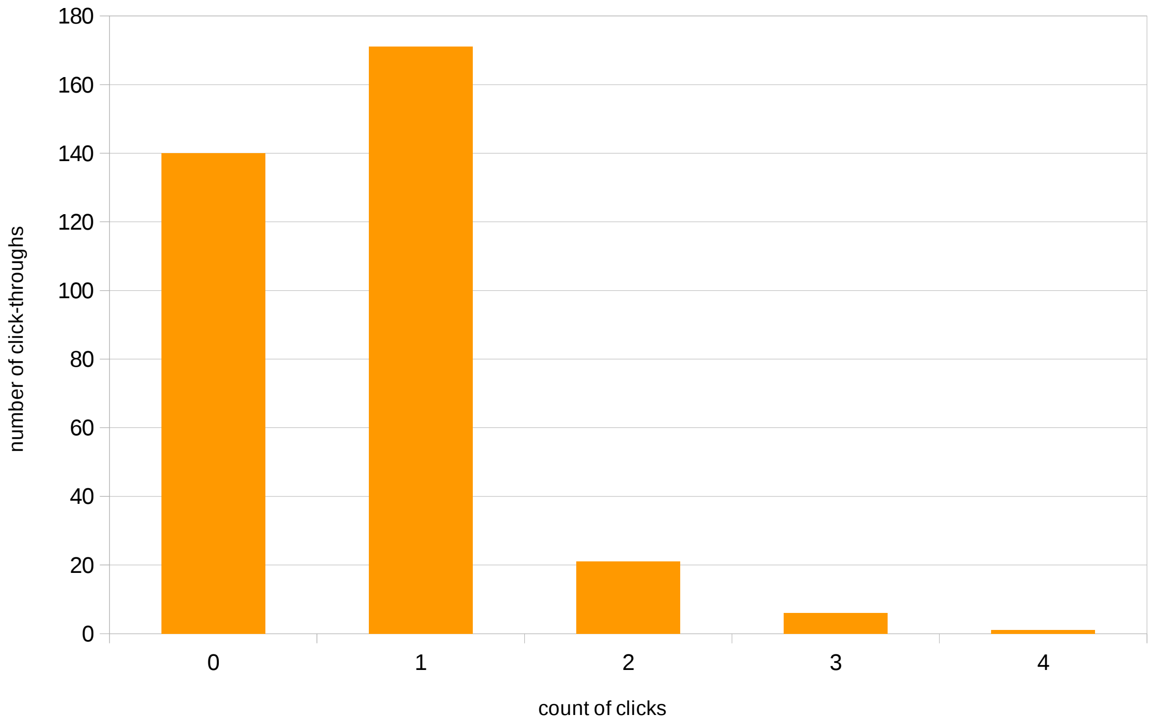


FIGURE 5.9. Distribution of clicks-throughs on the mobile text interface

There are cases where the mobile text interface is assumed to have not provided relevant results. This is shown by a total count of 140 of no clicks on the returned search results. However, this can also mean the participants in the study did not find any relevant information for the given text queries on the returned results. When compared to the baseline interface in Section 5.5.8, the mobile text interface had 140 no clicks-throughs while the baseline interface had 62. These statistics suggest that, due to a smaller screen display, the participants did not explore as many results as possible on the mobile text interface as compared to the larger display of the baseline interface. Nonetheless, recommendations were made to help improve the interface to give more relevant search results. The participants, when asked, 'what do you think needs most improvement, and why?', responded:

> 'The accuracy of results. Because some of its results are irrelevant to the words which I am searching for.'
>
> 'Just the accuracy of search results'
>
> 'The information that is supplied needs to be relevant to what the user is searching'
>
> 'There are duplicates of results'

These comments suggest that the interface had some duplicate and inaccurate results in some cases. However, the accuracy of the search results was the responsibility of the search engine, which was not the main focus of this study. Even so, the above recommendations suggest that in order for the interface to be considered to provide more relevant results the search engine has to be improved. From Figure 5.9, it can also be seen that some participants viewed more than 2 results for some queries, with a click-through count of 21, and some viewed 3 and 4 results, with click-through counts of 6 and 1, respectively. These statistics demonstrate that the mobile text interface was considered to have provided relevant results by most of the participants in the study, which they successfully viewed.

5.7 Test for Normality

This section of the chapter discusses the results from the Shapiro-Wilk test for normality. The Shapiro-Wilk test was computed using the R script shown in Appendix K. The tests were conducted in order to determine whether the data distribution is normal. This is important, as it will help in deciding, which test to apply to check statistical significant difference in the SUMI subscales' mean values for each interface, as discussed in Section 4.5. As seen from Tables 5.12, 5.13 and 5.14, the p-values for most of the SUMI subscales is less than 0.05, as such the null hypothesis is rejected. This is further confirmed by the Q-Q plots shown in Appendix J. Consequently, most (not all) SUMI subscales data for the interfaces is not normal and, as such, the Wilcoxon test will be used for testing significant difference in the subscale means, as the test is applicable whether the data distribution is normal or not.

Table 5.12: Baseline Interface: Shapiro-Wilk test for normality

SUMI Subscales	Statistic	p-Value
Global	0.941	0.065
Efficiency	0.893	0.003
Affect	0.929	0.030
Helpfulness	0.911	0.009
Controllability	0.974	0.575
Learnability	0.899	0.004

Table 5.13: Mobile Text Interface: Shapiro-Wilk test for normality

SUMI Subscales	Statistic	p-Value
Global	0.897	0.005
Efficiency	0.935	0.055
Affect	0.905	0.008
Helpfulness	0.895	0.005
Controllability	0.897	0.005
Learnability	0.897	0.005

Table 5.14: Mobile Voice Interface: Shapiro-Wilk test for normality

SUMI Subscales	Statistic	p-Value
Global	0.973	0.559
Efficiency	0.941	0.071
Affect	0.971	0.497
Helpfulness	0.954	0.173
Controllability	0.959	0.244
Learnability	0.924	0.022

5.8 Baseline Interface vs Mobile Text Interface

In this section of the chapter, the baseline interface is statistically compared to the mobile text interface to determine the most effective interface between the two, as determined by the participants in the study. This will result in answering the first research question of the study, which reads:

a) Is a mobile text interface more effective than a non-mobile text interface for isiXhosa speakers when searching for information on the Web?

The consolidated SUMI results for the baseline interface and mobile text interface are shown in Table 5.15. The Wilcoxon rank sum tests were computed using the R script shown in Appendix I. As seen from Table 5.15, there was no significant difference between the Global scale mean for the baseline interface and the mobile text interface. The p-value of 0.974 suggests that the null hypothesis cannot be rejected. That is, the mean value of the Global scale for the baseline interface is not different from the mean value for the mobile text interface. This means that the baseline interface was not considered more effective than the mobile text interface. The p-values for the subscales Efficiency, Affect, Helpfulness, Learnability and Controllability attest to this, as they are greater than the significance level of 0.05, as shown in Table 5.15.

Table 5.15: Wilcoxon rank sum test of the baseline interface against the mobile text interface on the SUMI scales

	Baseline Interface	Mobile Text Interface	Wilcoxon Signed-Rank Test	
SUMI Scales	Mean	Mean	Wilcoxon Value	p-Value
Global	59.12	58.19	541.0	0.974
Efficiency	58.74	56.91	566.0	0.782
Affect	60.32	58.75	564.5	0.797
Helpfulness	55.74	54.22	600.0	0.476
Controllability	55.35	53.72	554.0	0.903
Learnability	61.68	57.91	686.0	0.068

5.9 Baseline Interface vs Mobile Voice Interface

In this section of the chapter, the baseline interface is statistically compared to the mobile voice interface to determine the most effective interface between the two, as determined by the participants in the study. This will result in answering the second research question for the study, which reads:

b) Is a mobile voice interface more effective than a non-mobile text interface for isiXhosa speakers when searching for information on the Web?

Table 5.16 presents the consolidated SUMI results for the baseline interface and the mobile voice interface. As seen from the results, there was a significant difference between the Global scale mean for the baseline interface and the mobile voice interface. The p-value of 0.001 suggests that the null hypothesis is rejected in favor of the alternate hypothesis. As such, the mean value of the Global scale for the baseline interface is higher than the mean value for the mobile voice interface. This means that the baseline interface was considered more effective than the mobile voice interface. The p-values for the subscales Efficiency, Affect and Learnability attest to this, with Efficiency and Learnability subscales having p-values less than 0.001, as shown in Table 5.16.

Table 5.16: Wilcoxon rank sum test of the baseline interface against the mobile voice interface on the SUMI scales

	Baseline Interface	Mobile Voice Interface	Wilcoxon Signed-Rank Test	
SUMI Scales	Mean	Mean	Wilcoxon Value	p-Value
Global	59.12	49.70	831.5	0.001
Efficiency	58.74	45.39	881.5	< 0.001
Affect	60.32	50.06	826.5	0.001
Helpfulness	55.74	52.12	687.5	0.113
Controllability	55.35	49.58	782.0	0.006
Learnability	61.68	51.27	874.5	< 0.001

These statistics show that the participants in the study were more pleased with the baseline interface, they found it easy to assist them in their search and it was easy to learn. However, there was no significant difference between the Controllability (p-value of 0.006) and Helpfulness (p-value of 0.113) mean values for the baseline interface and the mobile voice interface.

5.10 Mobile Text Interface vs Mobile Voice Interface

In this section of the chapter, the mobile text interface is statistically compared to the mobile voice interface to determine the preferred interface between the two by the participants in the study. This will result in answering the third research question of the study. As discussed in Chapter 4, the third research question for the study reads:

c) Is a mobile voice interface preferred over a mobile text interface for isiXhosa speakers when searching for information on the Web?

The consolidated SUMI results for the mobile text interface and mobile voice interface are presented in Table 5.17. The R script used to compute the Wilcoxon signed-rank test is shown in Appendix I. According to the results, there was a significant difference between the Global scale mean for the mobile text interface and the mobile voice interface. The p-value of 0.002 means that the null hypothesis can be rejected in favor of the alternate hypothesis. Consequently, the mean for the Global scale for the mobile text interface is higher than the mean for the mobile voice interface. Therefore, the mobile text interface was preferred over the mobile voice interface. The p-values for Efficiency, Affect, Controllability and Learnability confirm this, as shown in Table 5.17, where the p-values for these subscales are less than 0.05. However, the p-value of 0.281 for the Helpfulness subscale suggests that there was no significant difference in the helpfulness of the mobile text and voice interfaces.

Table 5.17: Wilcoxon rank sum test of the mobile text interface against the mobile voice interface on the SUMI scales

	Mobile Voice Interface	Mobile Text Interface	Wilcoxon Signed-Rank Test	
SUMI Scales	Mean	Mean	Wilcoxon Value	p-Value
Global	49.70	58.19	762.0	0.002
Efficiency	45.39	56.91	772.5	0.001
Affect	50.06	58.75	744.0	0.005
Helpfulness	52.12	54.22	610.5	0.281
Controllability	49.58	53.72	710.0	0.017
Learnability	51.27	57.91	696.0	0.028

5.11 SUMI Scales vs Speed vs Relevance

The results in Sections 5.8 and 5.9 show that the mobile voice interface was considered by most participants to be less effective when compared to the text and baseline interfaces. The mobile text interface was considered more satisfying and overall preferred over the mobile voice interface, as discussed in Section 5.10. Nonetheless, the mobile text interface and baseline interface were equal with regards to their effectiveness.

The speed results for the 3 interfaces show that the baseline interface was the fastest to provide participants with search results and that the participants took the shortest time to interact with the provided results. The box-plot for this interface shows that for the 10 search queries the interface provided results and participants acted on them in between 3.5 and 7.5 seconds. On the other hand, the mobile text interface was the second fastest (3.5 to 5.5 seconds) while the mobile voice interface was the slowest (6.5 and 10 seconds). Despite this, the interfaces were consistent in providing participants with the results and participants took about the same amount of time to interact with the results as well.

The results in Section 5.4.8 suggest that the mobile voice interface did not provide many relevant results for most of the 10 search queries compared to the other interfaces. However, participants did not find relevant results as the voice queries were mostly not recognized by the voice interface and not necessarily that the interface did not provide relevant results, as discussed in Section 5.4.8. As it could perhaps been expected, the baseline interface was considered by participants to provide more relevant results than the mobile interfaces, as discussed in Section 5.5.8. This suggests that due to a larger screen display the participants were able to explore more results as it was easy to navigate from one result to another, which is not the case on a smaller mobile display.

5.12 Chapter Summary

A pilot study was conducted with 3 students from the first year Computer Science class. In this pilot study, the participants evaluated the mobile interfaces and the baseline interface and at the end of the evaluation they completed SUMI questionnaires for each of the interfaces. Additionally, application log files were used in conjunction with the SUMI questionnaires.

The results from the pilot study indicated that the experiment design was feasible. As such, experiments were conducted to collect data while participants evaluated the two mobile interfaces and the baseline interface. The results from the experiments indicate that the participants were able to successfully evaluate the interfaces. Furthermore, the results show that the participants were satisfied with the usability of the interfaces. In addition to this, the participants were pleased with the speed and relevance of the interfaces in supporting their search process. However, some improvements were suggested to help improve the usability, relevance and speed of the interfaces.

The Global scale rating measured the general usability satisfaction of the interfaces by the participants in the study. From the results, there was no significant difference between the Global scale mean for the baseline interface and the mobile text interface. This demonstrates that the baseline interface was considered equivalently satisfying as the mobile text interface by the participants in the study. There was a significant difference between the Global scale mean for the baseline interface and the mobile voice interface. This suggests that the baseline interface was considered more satisfying than the mobile voice interface by the participants in the study. Further, the results indicate that there was a significant difference between the Global scale mean for the mobile text interface and the mobile voice interface. Therefore, the mobile text interface is considered preferable to the mobile voice interface.

The speed and relevance assessment of the interfaces shows that the text interface was faster than the voice interface, and provided more relevant results. The results also suggest that due to the inability of the voice interface to recognize some queries, it could not provide participants with the expected relevant results.

Chapter 6

Conclusions

This chapter discusses the conclusions drawn from the results of the experiments. How the experimental findings addressed the research questions for the study is discussed in Section 6.1. In Section 6.2 a discussion of the implications of the study is outlined. Thereafter, recommendations for future research are discussed in Section 6.3.

6.1 Research Questions

The focus point of this study was to compare mobile search interfaces that had been adapted to the requirements of isiXhosa speakers to help them search for information on the Web. Three research questions were posed to address this (see Section 1.3).

The results from the experiment (discussed in Section 5.8) show that the effectiveness of the baseline desktop interface was equivalent to that of the mobile text interface for isiXhosa speakers when searching for information on the Web. The two interfaces were equivalent in most aspects, such as: assisting isiXhosa speakers to complete their search process; self-explanatory help facilities; the ability to allow isiXhosa speakers to command the interface; and the ability for isiXhosa speakers to learn how to use the interfaces. Additionally, both interfaces provided search results in consistent time and the search results were considered relevant by most of the isiXhosa speakers.

The second research question compared the baseline desktop interface and the mobile voice interface to determine which is effective, noting that recent research and mobile phone innovations suggest that the mobile voice interface could be more effective as users do not have to type search queries but rather simply speak them [68][55][38]. However, on the contrary the results presented in Section 5.9 indicate that isiXhosa speakers found the baseline interface to be more effective than the mobile voice interface. This could be because the participants who were students had greater familiarity with PCs, as they use them on a daily basis in the university laboratories. It would be interesting to see how a change in the population sample could affect the results of the experiments. In particular, a technology illiterate sample of isiXhosa speakers, especially those in rural settings, where access to a PC is not common, who might not find the baseline interface effective as they are not used to it. Nonetheless, the results also indicate that isiXhosa speakers found the mobile voice interface to be usable and efficient in assisting them to complete their search process. This is an important finding that suggests the mobile voice interface could work if it was the only option in searching for information while mobile.

In the last research question, the mobile text and voice interfaces were compared to determine which is preferable, taking into consideration that recent research and mobile phone innovations suggest that the mobile voice interface could be the preferred choice [68][55]. In contrast to what was expected, participants preferred the text interface in general, and according to most SUMI subscales, speed assessment and the ability to provide relevant results, as shown in Sections 5.10 and 5.11 respectively. This could be because of greater familiarity with text search interfaces or because of the relative scarcity of voice interfaces in African (Bantu) languages or because participants were students hence were used to typing in platforms like WhatsApp, Twitter and Facebook. Where users are not literate, the voice interface may be the only option, so the fact that it was deemed usable is an important independent finding.

6.2 Implications of the Study

6.2.1 Technology Acceptance

In this section of the chapter, the Technology Acceptance Model (TAM) is discussed, with regard to the results presented in Chapter 5, particularly focusing on why participants in the study did not prefer the mobile voice interface over the mobile text interface. TAM explains the determinants of users' acceptance of a technology [32]. It proposes that users' attitude towards a technology is determined by the users' perceived usefulness (PU) and perceived ease of use (PEOU) of the given technology [32]. TAM follows the illustration shown in Figure 6.1, and the variables are as explained below:

External Variables: these are factors that could influence the beliefs of a user towards a system such as system design characteristics, user participation in design, system implementation and user training.

Perceived Usefulness (PU) : this refers to the extent to which a user believes that using a system would boost their task performance.

Perceived Ease of Use (PEOU) : this refers to the extent to which a user believes that using a particular system would be effortless.

Behavioral Intention to Use: a user's actual behaviour can be determined by considering their prior intention along with beliefs that the user might have for the given behaviour. Therefore, the intent that a user has prior to an actual behaviour is considered behavioural intent of the user to use a system [11].

Actual System Use: this refers to the act of using a system by a user.

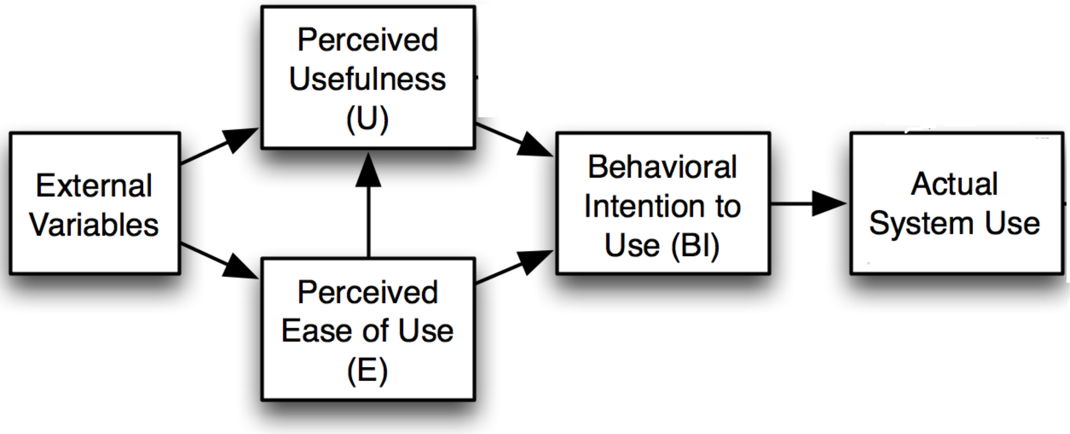


FIGURE 6.1. Technology Acceptance Model Overview, from Chuttur [11]

6.2.1.1 TAM and Results

According to the study's findings presented in Section 5.10, there was a significant difference in the Global scale mean for the mobile text interface and the mobile voice interface. Therefore, the mobile text interface was preferred over the mobile voice interface. Furthermore, there was a significant difference in the Efficiency, Affect, Controllability and Learnability subscales for the mobile text interface and the mobile voice interface.

In regard to the Efficiency subscale, there were more participants who found the mobile voice interface to be less efficient than the mobile text interface in supporting them in their search, with below average ratings of 30.3% for the mobile voice interface and 12.5% for the mobile text interface, as discussed in Sections 5.4.2 and 5.6.2. Therefore, given the TAM, the participants perceived the mobile voice interface to be less useful as it was less efficient than the mobile text interface in supporting them in their search. This in turn negatively influenced the participants' behavioural intention to use the mobile voice interface compared to the mobile text interface, which they deemed more efficient. Consequently, efficiency as discussed above has contributed to the overall preference of the mobile text interface by participants in the study. The study's findings also suggest that the participants found the mobile voice interface to be inefficient as in most cases it had failed to recognize the queries they had spoken. These findings are similar to previous research by Lee and Grice [36], who found out that the efficiency of voice systems and their interfaces are dependent on their ability to recognize human speech.

The study's findings on the Affect subscale show that there were more participants who did not like the mobile voice interface compared to the mobile text interface. This is shown by 21.2% of the participants who rated the mobile voice interface's likability/affect as below average, and

9.4% who rated the mobile text interface's likability/affect as below average. From Section 5.4.3, the study's findings demonstrate that the participants did not like the mobile voice interface as it was awkward and frustrating, leading to a tense feeling while interacting with it. The participants' frustration was caused by the way the interface presented the spoken feedback information to them, which resulted in the participants perceiving the mobile voice interface uneasy to use, when compared to the mobile text interface. The participants' PEOU, therefore, negatively influenced their behavioural intention to use the mobile voice interface. These findings also suggest that the participants in the study found it easy to process written information compared to spoken information, similar to previous work by Trippas et al. [66], who found out that because speech is linear it makes it difficult to present complex information structure, which humans can easily process without overloading the short-term memory.

According to the study's findings in Section 5.10, there was a significant difference in the Controllability subscale mean of the mobile text interface and the mobile voice interface. This implies that there were more participants who felt they could not control the mobile voice interface to get it to do what they wanted, as compared to those who felt they could control the mobile text interface. These findings suggests that because participants were not able to control the mobile voice interface to successfully complete their search, they perceived the mobile voice interface uneasy to use, when compared to the mobile text interface. As such, this perception negatively influenced the participants' behavioural intention to use the mobile voice interface.

There was a significant difference in the Learnability subscale mean of the mobile text interface and the mobile voice interface, which suggests that most participants found it easier to learn the functions of the mobile text interface than those of the mobile voice interface. According to the results in Section 5.4.6, 24.2% of the participants rated the mobile voice interface's Learnability as below average compared to 6.3% who rated the mobile text interface as below average. These statistics suggest that because the participants were not able to easily learn the functions of the mobile voice interface, as compared to the mobile text interface, they perceived the ease-of-use of the mobile voice interface to be less than that of the mobile text interface. Thus, this negatively influenced the participants' intention to use the mobile voice interface, as compared to the mobile text interface. However, these findings also suggest that the participants' past experience with text interfaces, that they have been using to search for information on project assignments as part of their curriculum, could have contributed to the overall familiarity with the mobile text interface.

Although the mobile voice interface was considered by participants to be fast, majority of the participants felt the mobile text interface was faster, as discussed in Section 5.11. The fact that the voice interface was perceived slower than the text interface could have resulted in participants perceiving it less useful compared to the text interface, thus negatively influencing their behavioural intention to use the mobile voice interface. Also, the mobile text interface was considered by most participants to have provided most relevant results compared to the mobile voice interface. This outcome has negatively influenced the participants' behavioural intention to use the voice interface. Nonetheless, the participants were unable to retrieve relevant voice results as in most cases the interface failed to recognize voice queries as participants were speaking faster than the interface could detect, as suggested in Section 5.4.8.

The study's findings have shown that the TAM has a satisfactory fit with the analyzed data. The Efficiency, Affect, Controllability and Learnability subscales were found to influence participants' PU and PEOU of the mobile voice interface, which in turn affected their intention to use the interface to search for information on the Web. Similarly, the voice interface's speed and ability to provide relevant results were found to influence participants' PU and PEOU of the mobile voice interface, which also affected their intention to use the interface. These findings show that, generally, users would not prefer to use a voice interface that they cannot have control over, that is inefficient, provides less relevant results and less obvious to use from the initial point of interaction. In addition, the findings have shown that external variables such as past experience and interface design, generally, affect how users perceive the usefulness and ease of use of a search interface, that in turn affect their intention to use the interface. With greater familiarity, users tend to be more accepting of a user interface.

6.2.2 Designing for Search in African (Bantu) Languages

In Chapter 3, the design process followed while developing the mobile interfaces is discussed. The design process was mainly focused on addressing the gaps that had been identified in the literature, as discussed in Chapter 2. Furthermore, the design process was guided by the search needs of isiXhosa speakers, the limited amount of speech corpora available when compared to languages such as English, the limited amount of isiXhosa content on the Web, the limitations of mobile phones and best design practices as recommended in the literature. The limitations of mobile phones were the universal limitations as identified by other researchers in different studies [18][69]. These include smaller display, insufficient battery life and Internet connection costs. The search needs of isiXhosa speakers were mainly to search for information on the Web in isiXhosa. However, this can be applicable to other communities that speak African (Bantu) languages, which are mainly found in developing countries, where English literacy is considerably low. As such, this study confirms that search interfaces in African (Bantu) languages are possible.

Open-source software was used to develop the prototype of this study and it has proven to work successfully. This goes to show that where budget is limited or non-existent, open-source can be useful given its zero licensing costs. Additionally, the ASR system that compliments the voice interface was built to work offline thus reducing Internet connection costs as the voice interface did not have to connect to an online ASR system that could incur Internet costs for detecting and interpreting voice queries. This design choice is particularly important given that Internet connection costs are high in South Africa.

6.2.3 Content and African (Bantu) Languages

As noted in Chapter 3, the isiXhosa Wikipedia and the South African government's isiXhosa parliament speech websites were used to provide content for all the interfaces. These were the only sites that had informative isiXhosa documents that could better support the motive of this study, which was to provide search interfaces that could enable access to isiXhosa information on the Web. Although the sites provided documents for participants to interact with, the content of the documents was inadequate, as it was not up to date. Furthermore, participants in the study felt the information provided by the interfaces was irrelevant.

The lack of availability of relevant information in IR systems affects usage of such systems [72]. Users feel empowered to use IR systems that can provide useful, current and relevant information [72]. Therefore, the lack of relevant information in isiXhosa is a factor that can lead to users not opting to use the developed interfaces. Thus, the issue of insufficient digital content in isiXhosa and other African (Bantu) languages needs to be improved by translating resource-rich languages' documents to isiXhosa or generating original content.

6.3 Future Work

6.3.1 Prototype

The Android mobile application developed as part of this study was a prototype that had the two mobile interfaces. Therefore, the prototype could be expanded into a fully functioning mobile application. The speech recognizer could be expanded by using the NCHLT speech corpora to extend the ability of the ASR to recognize more words. This speech data was not used because it did work at an acceptable level of accuracy at the time of training. However, with more time this speech data could potentially result in good accuracies, as it was targeted for mobile use. Further, collaboration with ASR systems developers such as those identified in Chapter 2 is needed, particularly those working with the isiXhosa language. Also, the search engine the interfaces were reliant on could also be improved to include as many isiXhosa Web resources as possible.

6.3.2 Other African (Bantu) Languages

The fact that it was possible to design and develop usable mobile search interfaces for isiXhosa speakers means that similar interfaces could be extended to other African (Bantu) languages, for example, isiZulu, Setswana, Sesotho, Tshivenda, Xitsonga and Sepedi.

6.3.3 Other Types of Search

In addition to text and voice search, the search interfaces could be extended to include other types of search, for example image, video, and audio search. In addition to this, the interfaces could be extended to include searching Web Directories for isiXhosa, which currently do not exist.

6.3.4 More Experiments

Additional experiments could be conducted with the improved prototype to see how the new changes will affect how isiXhosa speakers search for information on the Web. Additionally, the participants in the study were native isiXhosa speaking university students. Therefore, the experiments could be extended to include the general public of native isiXhosa speakers, more especially the low literate to illiterate speakers. This will be helpful to determine the impact of mobile search interfaces in supporting information access by isiXhosa speakers who are not English or technology literate.

The experiments were conducted in a controlled environment and, as such, it was easy for the mobile voice interface to detect spoken queries without noise interference from the surroundings. Although this approach has proven to work, it might not be a true representation of a real life usage of voice search interfaces. As such, more experiments could be conducted in an environment that is relatively a true representation of usage of mobile voice interfaces, for example, while driving or walking in public.

Relevance and speed results were analysed over all trials from all 34 participants in the experiments. In future, the results could be analysed on a per-query basis to do a more fine-grained analysis of user behaviour as well as the interface speed.

BIBLIOGRAPHY

[1] C. ALLAUZEN, E. BENSON, C. CHELBA, M. RILEY, AND J. SCHALKWYK, *Voice Query Refinement.*, in 13th Annual Conference of the International Speech Communication Association 2012 (INTERSPEECH 2012), vol. 3, Portland, Oregon, USA, 2012, International Speech Communication Association (ISCA), pp. 1–4.

[2] J. Q. ANDERSON AND L. RAINIE, *The Future of the Internet III*, tech. rep., Pew Internet & American Life Project, Washington, USA, 2008.

[3] A. AULA, *Query formulation in web information search*, in Proceedings of IADIS International Conference WWW/Internet 2003, IADIS Press, Algarve, Portugal, 2003, ICWI, pp. 403–410.

[4] J. BADENHORST, C. VAN HEERDEN, M. DAVEL, AND E. BARNARD, *Collecting and evaluating speech recognition corpora for nine Southern Bantu languages*, in Proceedings of the First Workshop on Language Technologies for African Languages, Athens, Greece, 2009, Association for Computational Linguistics, pp. 1–8.

[5] E. BARNARD, M. H. DAVEL, C. V. HEERDEN, F. D. WET, AND J. BADENHORST, *The NCHLT Speech Corpus of the South African languages*, in SLTU 2014, no. 01, St. Petersburg, Russia, 2014, pp. 194–200.

[6] E. BARNARD, J. SCHALKWYK, C. VAN HEERDEN, AND P. J. MORENO, *Voice search for development*, Journal of the Human Language Technologies General Science, Engineering and Technology, (2010), pp. 282–285.

[7] M. J. BATES, *Where should the person stop and the information search interface start?*, Journal of the Information Processing and Management, 26 (1990), pp. 575–591.

[8] M. CHAU, X. FANG, AND C. C. YANG, *Web searching in Chinese: A study of a search engine in Hong Kong*, Journal of the American Society for Information Science and Technology, 58 (2007), pp. 1044–1054.

[9] E. CHENG, F. JING, L. ZHANG, AND H. JIN, *Scalable relevance feedback using click-through data for web image retrieval*, in Proceedings of the 14th annual ACM international conference on Multimedia, California, USA, 2006, ACM, pp. 173–176.

[10] K. CHURCH, B. SMYTH, P. COTTER, AND K. BRADLEY, *Mobile information access: A study of Emerging Search Behavior on the Mobile Internet*, ACM Transcations on the Web, 1 (2007), pp. 1–38.

[11] M. Chuttur, *Overview of the Technology Acceptance Model: Origins, Developments and Future Directions.*, Sprouts Working Papers on Information Systems, 9 (2009), pp. 9–37.

[12] P. Clarkson and R. Rosenfield, *Statistical Language Modeling Using the CMU-Cambridge Toolkit*, Eurospeech, 97 (1997), pp. 2707–2710.

[13] G. Cordasco, M. Esposito, F. Masucci, M. T. Riviello, A. Esposito, P. Milhorat, and G. Pelosi, *Assessing Voice User Interfaces: The vAssist System Prototype*, in 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom), Vietri sul Mare, Italy, 2014, IEEE, pp. 91–96.

[14] W. B. Croft, D. Metzler, and T. Strohman, *Search Engines Information Retrieval in Practice*, Pearson Education, Inc, 2 ed., 2015.

[15] L. Dalvit, A. Terzoli, and F. Wolff, *Open-Source Software and Localisation in Indigenous South African Languages with Pootle*, in Southern Africa Telecommunication Network and Application Conference, Durban, South Africa, 2008, SATNAC, pp. 1–6.

[16] M. Davel and O. Martirosian, *Pronunciation dictionary development in resource-scarce environments*, in Proceedings of the Annual Conference of the International Speech Communication Association (NTERSPEECH), Brighton, United Kingdom, 2009, International Speech Communication Association, pp. 2851–2854.

[17] N. J. De Vries, M. H. Davel, J. Badenhorst, W. D. Basson, F. De Wet, E. Barnard, and A. De Waal, *A smartphone-based ASR data collection tool for under-resourced languages*, Speech Communication, 56 (2014), pp. 119–131.

[18] M. D. Dunlop and S. a. Brewster, *The Challenge Of Mobile Devices For Human Computer Interaction*, Personal and Ubiquitous Computing, 6 (2002), pp. 235–236.

[19] R. Eiselen and M. J. Puttkammer, *Developing Text Resources for Ten South African Languages*, in International Conference on Language Resources and Evaluation (IREC), Reykjavik, Iceland, 2014, IREC, pp. 3698–3703.

[20] J. E. Gilbert and D. Hall, *Speech User Interfaces for Information Retrieval*, in Proceedings of the twelfth international conference on Information and knowledge management, New York, USA, 2003, ACM, pp. 77–82.

[21] J. Goodman, *A Bit of Progress in Language Modeling*, Computer Speech & Language, 15 (2001), pp. 403–434.

[22] Google and Nielsen, *Google/Nielsen Life360 Mobile Search Moments Q4 2012*, Tech. Rep. March, Google and Nielsen, 2012.

[23] M. Hattab, B. Haddad, M. Yaseen, S. Duraidi, Asem, and A. Abu, *Addaall Arabic Search Engine: Improving Search based on Combination of Morphological Analysis and Generation Considering Semantic Patterns*, in The second International Conference on Arabic Language Resources and Tools, Cairo, Egypt, 2009, Mediterranean Arabic Language and Speech Technology, pp. 1–4.

[24] M. A. Hearst, *Search User Interfaces*, Cambridge University Press, 1st ed., 2009.

[25] M. A. Hearst, *User Interfaces for Search*, in Modern Information Retrieval: The Concepts and Technology behind Search, Addison-Wesley, 2011, ch. 2, pp. 21–55.

[26] C. V. Heerden, M. H. Davel, and E. Barnard, *Performance analysis of a multilingual directory enquiries application*, in Pattern Recognition Association of South Africa (PRASA), South Africa, 2015, PRASA, pp. 1–6.

[27] P. Ingwersen, *Information Retrieval Interaction*, Taylor Graham Publishing, 1992.

[28] M. Iwata, T. Hara, K. Shimatani, T. Mashita, K. Kiyokawa, S. Nishio, and H. Takemura, *A location-based content search system considering situations of mobile users*, Procedia Computer Science, 5 (2011), pp. 426–433.

[29] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay, *Accurately interpreting clickthrough data as implicit feedback*, in Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '05, New York, USA, 2005, ACM, pp. 154–161.

[30] M. Jones, G. Marsden, N. Mohd-Nasir, K. Boone, and G. Buchanan, *Improving Web interaction on small displays*, in Proceedings of the Eighth International Conference on World Wide Web, vol. 31 of WWW '99, New York, USA, 1999, Elsevier North-Holland, Inc., pp. 1129–1137.

[31] P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, *Usability Evaluation In Industry*, CRC Press, London, 1996.

[32] Y. G. Kim and E. Woo, *Consumer acceptance of a quick response (QR) code for the food traceability system: Application of an extended technology acceptance model (TAM)*, Food Research International, 85 (2016), pp. 266–272.

[33] J. Kirakowski and M. Corbett, *SUMI: The software usability measurement inventory*, vol. 24, John Wiley & Sons, Inc., 1993.

[34] J. Koenemann and N. J. Belkin, *A case for interaction: A study of interactive information retrieval behavior and effectiveness*, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '96, New York, USA, 1996, ACM, pp. 205–212.

[35] M. KYEYUNE, *IsiXhosa Search Engine Development Report*, tech. rep., University of Cape Town, Cape Town, 2015.

[36] K. LEE AND R. GRICE, *The Design and Development of User Interfaces for Voice Application in Mobile Devices*, in 2006 IEEE International Professional Communication Conference, New York, USA, 2006, IEEE, pp. 308–320.

[37] K.-F. LEE, H.-W. HON, AND M.-Y. HWANG, *Recent progress and future outlook of the SPHINX speech recognition system*, Computer Speech and Language, (1990), pp. 1–13.

[38] L. S. LEE AND Y. C. PAN, *Voice-based information retrieval - How far are we from the text-based information retrieval?*, in 2009 IEEE Workshop on Automatic Speech Recognition and Understanding, Merano, Italy, 2009, IEEE, pp. 26–43.

[39] C. LIU, P.-L. P. RAU, AND F. GAO, *Mobile information search for location-based information*, Computers in Industry, 61 (2010), pp. 364–371.

[40] B. LONG AND Y. CHANG, *Relevance Ranking for Vertical Search Engines*, Elsevier, 2014.

[41] M. LOT, R. S. MIZZARO, AND U. VIA, *Evaluating User Interfaces to Information Retrieval Systems : A Case Study on User Support*, in In 19th International Conference on Research and Developments in Information Retrieval (SIGIR-96), Zurich, Switzerland, 1996, Association for Computing Machinery, Inc., pp. 128–136.

[42] G. MARCHIONINI, *Interfaces for end-user information seeking*, Journal of the American society for information science, 43 (1992), pp. 156–163.

[43] G. MARCHIONINI, *User Interfaces for Information Retrieval on the WWW*, INFORUM 2005: 11th Conference on Professional Information Resources, (2005).

[44] G. MARCHIONINI, *Toward human-computer information retrieval*, Bulletin of the American Society for Information Science and Technology, 32 (2007), pp. 20–22.

[45] I. MEDHI THIES, *User Interface Design for Low-literate and Novice Users: Past, Present and Future*, Foundations and Trends ® Human–Computer Interaction, 8 (2015), pp. 1–72.

[46] T. MENEMENIS, S. PAPADOPOULOS, B. BRATU, S. WADDINGTON, AND Y. KOMPATSIARIS, *AQUAM : Automatic Query Formulation Architecture for Mobile Applications*, in Proceedings of the 7th international Conference on Mobile and Ubiquitous Multimedia, New York, USA, 2008, ACM, pp. 32–39.

[47] T. MISHRA, P. AVE, F. PARK, P. AVE, AND F. PARK, *Qme ! : A Speech-based Question-Answering system on Mobile Devices*, in Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational

Linguistics, Los Angeles, USA, 2010, Association for Computational Linguistics, pp. 55–63.

[48] J. NIELSEN, *Iterative User-Interface Design*, IEEE Computer Society Press, 26 (1993), pp. 32–41.

[49] S. PACKHAM AND H. SULEMAN, *Crowdsourcing a Text Corpus is not a Game*, Springer International Publishing, Cham, 2015, pp. 225–234.

[50] N. RAJPUT AND A. A. NANAVATI, *Speech in Mobile and Pervasive Environments*, Wiley Publishing, 1st ed., feb 2012.

[51] B.-Y. RICARDO, G. DUPRET, AND J. VELASCO, *A study of mobile search queries in Japan*, in Proceedings of the International World Wide Web Conference, New York, USA, 2007, ACM.

[52] J. C. ROUX, P. H. LOUW, AND T. R. NIESLER, *The African Speech Technology Project : An Assessment*, in LREC, Lisbon, Portugal, 2004, pp. pp. 93–96.

[53] T. RUSSELL-ROSE AND T. TATE, *Designing the Search Experience*, Elsevier, 2013.

[54] N. SAMBASIVAN, M. YOUSSEFMIR, A. T. LARSSEN, P. LEE, G. HECHT, P. M. AOKI, M.-I. CARRERA, J. CHEN, D. P. COHN, P. KRUSKALL, AND E. WETCHLER, *Chale, how much it cost to browse?: Results from a mobile data price transparency trial in Ghana*, in Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1, New York, USA, 2013, ACM, pp. 13–23.

[55] J. SCHALKWYK, D. BEEFERMAN, B. FRANCOISE, B. BYRNE, C. CHELBA, AND M. COHEN, *Advances in Speech Recognition*, Springer Publishing Company, Inc, 1st ed., 2010.

[56] B. N. SCHILIT, *Mobile Computing: Looking to the Future*, Computer, 44 (2011), pp. 28–29.

[57] M. SCHUSTER AND K. NAKAJIMA, *Japanese and korean voice search*, International Conference on Acoustics, Speech, and Signal Processing, 1 (2012), pp. 5149–5152.

[58] J. SHAN, G. WU, Z. HU, AND X. TANG, *Search by Voice in Mandarin Chinese*, in Interspeech, Chiba, Japan, 2010, pp. 354–357.

[59] A. A. SHIRI, C. W. REVIE, AND G. CHOWDHURY, *Thesaurus-enhanced search interfaces*, Journal of Information Science, 28 (2001), pp. 111–122.

[60] B. SHNEIDERMAN, D. BYRD, AND W. B. CROFT, *Clarifying search: A user-interface framework for text searches*, D-Lib Magazine, 3 (1997), pp. 3–20.

[61] STATISTICS SOUTH AFRICA, *Census in Brief*, 2011.

[62] B. SUHM, *Towards Best Practices for Speech User Interface Design*, in 8th European Conference on Speech Communication and Technology, Eurospeech 2003 - Interspeech 2003, Geneva, Switzerland, 2003.

[63] Y. SUNG, M. JANSCHE, AND P. MORENO, *Deploying Google Search by Voice in Cantonese*, in Interspeech, Florence, Italy, 2011, ISCA, pp. 2865–2868.

[64] S. SWEENEY AND F. CRESTANI, *Effective search results summary size and device screen size: Is there a relationship?*, Information Processing & Management, 42 (2006), pp. 1056–1074.

[65] M. TAIT, A. VISAGIE, AND T. R. NIESLER, *Development of a Spoken Dialogue System Operating in Afrikaans, South African English and Xhosa*, Transactions of the SAIEE, 96 (2005), pp. 224–233.

[66] J. R. TRIPPAS, M. SANDERSON, D. SPINA, AND L. CAVEDON, *Results Presentation Methods for a Spoken Conversational Search System*, in Proceedings of the First International Workshop on Novel Web Search Interfaces and Systems, New York, USA, 2015, ACM, pp. 13–15.

[67] U. UMMUHANYSIFA AND P. NIZAR BANU, *Voice Based Search Engine And Web Page Reader*, International Journal of Computational Engineering Research, (2013), pp. 1–5.

[68] Y. Y. WANG, D. YU, Y. C. JU, AND A. ACERO, *An introduction to voice search*, IEEE Signal Processing Magazine, 25 (2008), pp. 29–38.

[69] R. WANT, *The power of smartphones*, IEEE Pervasive Computing, 13 (2014), pp. 76–79.

[70] M. L. WILSON, *Search User Interface Design*, Morgan & Claypool, 2012.

[71] M. WOLTERS, K. GEORGILA, J. D. MOORE, R. H. LOGIE, S. E. MACPHERSON, AND M. WATSON, *Reducing working memory load in spoken dialogue systems*, Interacting with Computers, 21 (2009), pp. 276–287.

[72] I. XIE AND K. K. MATUSIAK, *Discover Digital Libraries: Theory and Practice*, Elsevier Inc., 2016.

[73] N. YANKELOVICH AND J. LAI, *Designing speech user interfaces*, ACM, 1998.

[74] N. YANKELOVICH, G.-A. L. LEVOW, AND M. MARX, *Designing SpeechActs: Issues in speech user interfaces*, in Proceedings of the SIGCHI conference on Human factors in computing systems, Los Angeles, USA, 1995, ACM Press/Addison-Wesley Publishing Co., pp. 369–376.

[75] J. ZHANG AND S. LIN, *Multiple language supports in search engines*, Online Information Review, 31 (2007), pp. 516–532.

[76] M. ZIEFLE, *Information presentation in small screen devices: The trade-off between visual density and menu foresight*, Applied Ergonomics, 41 (2010), pp. 719–730.

APPENDIX A

Appendix A: Ethical Clearances

A.1 Science Faculty

Ethical clearance from the Science Faculty, University of Cape Town.

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Faculty of Science
University of Cape Town
RONDEBOSCH 7701 South Africa
E-mail: richard.hill@uct.ac.za
Telephone: + 27 21 650 2786
Fax: + 27 21 650 3456

7 September 2015

Mr Morebodi Modise
Department of Computer Science

A comparison of Mobile Search Interfaces for isiXhosa Speakers

Dear Mr Morebodi Modise

I am pleased to inform you that the Faculty of Science Research Ethics Committee has approved the above-named application for research ethics clearance, subject to the conditions listed below. You are required to:

- Implement the measures described in your application to ensure that the process of your research is ethically sound;
- Uphold ethical principles throughout all stages of the research, responding appropriately to unanticipated issues: please contact me if you need advice on ethical issues that arise; and
- Ensure that you get written consent of research participants whether or not you take videos or pictures of them.

Your approval code is: FSREC 49 – 2015

I wish you success in your research.

Yours sincerely

RCHill

Dr Richard Hill
Chair: Faculty of Science Research Ethics Committee

Cc: A/Prof Hussein Suleman, Supervisor

A.2 Student Affairs

Permission to access University students from the Department of Student Affairs

	RESEARCH ACCESS TO STUDENTS	DSA 100

NOTES

1. This form must be **FULLY** completed by all applicants requesting access to UCT students for the purpose of research/surveys.
2. Return the fully completed **(a) DSA 100** application form **by email**, in **the same word format**, together with **your: (b) research proposal inclusive of your survey, (c) copy of your ethics approval letter / proof (d) informed consent letter** to: Moonira.Khan@uct.ac.za. You application will be attended to by the Executive Director, Department of Student Affairs (DSA), UCT.
3. The turnaround time for a reply is **approximately 10 working days**.
4. NB: It is the responsibility of the researcher/s to apply for and to obtain **ethics approval and to comply with amendments that may be requested;** as well as **to obtain** approval to access UCT staff and/or UCT students, from the following, at UCT, respectively: (a) **Ethics:** Chairperson, Faculty Research Ethics Committee' (FREC) for ethics approval, (b) **Staff access:** Executive Director: HR for approval to access UCT staff, and (c) **Student access:** Executive Director: Student Affairs for approval to access UCT students.
5. **Note:** UCT Senate Research Protocols requires compliance to the above, **even if prior approval has been obtained from any other institution/agency. UCT's research protocol requirements applies to *all* persons, institutions and agencies from UCT and external to UCT who want to conduct research on human subjects for academic, marketing or service related reasons at UCT.**
6. **Should approval be granted to access UCT students for this research study, such approval is effective for a period of one year from the date of approval (as stated in Section D of this form), and the approval expires automatically on the last day.**
7. **The approving authority reserves the right to revoke an approval based on reasonable grounds and/or new information.**

SECTION A: RESEARCH APPLICANT/S DETAILS

Position	Staff / Student No	Title and Name	Contact Details (Email / Cell / land line)
A.1 Student Number	MDSMOR003	Mr. Morebodi Modise	MDSMOR003@myuct.ac.za/ +27631798461
A.2 Academic / PASS Staff No.			
A.3 Visitor/ Researcher ID No.			
A.4 University at which a student or employee	University of Cape Town	Address if *not* UCT:	
A.5 Faculty/ Department/School	Science/Computer Science/University of Cape Town		
A.6 APPLICANTS DETAILS If different from above	Title and Name	Tel.	Email

SECTION B: RESEARCHER/S SUPERVISOR/S DETAILS

Position	Title and Name	Tel.	Email
B.1 Supervisor	Assoc Prof. Hussein Suleman	+27216505106	hussein@cs.uct.ac.za
B.2 Co-Supervisor/s			

SECTION C: APPLICANT'S RESEARCH STUDY FIELD AND APPROVAL STATUS

C.1 Degree – if applicable	MSc in Computer Science (Coursework and dissertation)
C.2 Research Project Title	A Comparison of Mobile Search Interfaces for isiXhosa Speakers
C.3 Research Proposal	Attached: Yes [X] No [ ]
C.4 Target population	First Language Xhosa Speaking Students
C.5 Lead Researcher details	If different from applicant:
C6. Will use research assistant/s	Yes [ ] No [X] If yes- provide a list of names, contact details and ID no.
C.7 Research Methodology and Informed consent:	Research methodology: focus group discussion, and post evaluation survey interviews Informed consent: will be obtained, with voluntary participation, and ensuring confidentiality and anonymity of data collection and reporting.
C.8 Ethics clearance status from UCT's Faculty Ethics Research Committee (FREC)	Approved by the FREC Yes [X] With amendments: Yes [ ] No [X] (a) Attach copy of your ethics approval. Attached: Yes (b) State date and reference no. of ethics approval: 07/09/2015 Ref./Faculty: FSREC 49 – 2015/Science

SECTION D: APPLICANT/S APPROVAL STATUS FOR ACCESS TO STUDENTS FOR RESEARCH PURPOSE

(*To be completed by the ED, DSA or Nominee)*

	Approved / With Terms / Not	* Conditional approval with terms	Applicant/s Ref. No.:
D.1 APPROVAL STATUS	(i) Approved ~~(ii) With terms~~ ~~(iii) Not approved~~	**(a)** Access to students for this research study must only be undertaken after written ethics approval has been obtained. **(b)** In event any ethics conditions are attached, these must be complied with before access to students.	MDSMOR003/ Mr. Morebodi Modise

	Designation	Name	Signature	Date of Approval
D.2 APPROVED BY:	*Executive Director Department of Student Affairs*	*Dr Moonira Khan*		*24 September 2015*

Version. 2015 | Page 1 of 1 | DSA 100

APPENDIX

B

Appendix B: Informed Consent Form

Informed consent form signed by participants to show participation was on voluntary basis in the design focus group, pilot study and final experiments.

Informed Consent

Project Title
A Comparison of Mobile Search Interfaces for isiXhosa Speakers

Project Summary
I am a Masters (coursework and dissertation) student in the Department of Computer Science at the University of Cape Town. As part of my research, I am studying different mobile search interfaces, text and voice, and how these can be adapted to the requirements of isiXhosa speakers to help them search for information digitally using isiXhosa as a discovery language.

I would appreciate it if you could help by taking part in this survey.

Important Information

1. Your personal details in the informed consent form will not be associated with your survey responses
2. You may decide to stop being a part of the research study at any time without explanation
3. There are no known risks associated with this survey
4. The data collected will be kept confidential and used solely for the purposes of this research
5. Ethics approval has been obtained from the Science Faculty (approval code **FSREC 49 – 2015**) and from the Department of Student Affairs (REF. No. **MDSMOR003/Mr. Morebodi Modise**)
6. For more information contact me at MDSMOR003@MYUCT.AC.ZA

Informed Consent
By signing below, you are agreeing that:

1. You understood the information provided above
2. Questions about your participation in this study have been answered satisfactorily
3. You are aware that there are no potential risks
4. You are taking part in this research study voluntarily (without coercion)

Participant Full name: ___________________________

Participant Signature: ________________________

Date:_____________________________________

APPENDIX C

Appendix C: SUMI Questionnaire

SUMI questionnaire used as post-evaluation questionnaire

Software Usability Measurement Inventory

SUMI

NB The information you provide is kept completely confidential, and no information is stored on computer media that could identify you as a person.

This questionnaire has 50 statements. Please answer them all. After each statement there are three boxes.

- Check the first box if you generally AGREE with the statement.
- Check the middle box if you are UNDECIDED, or if the statement has no relevance to your software or to your situation.
- Check the right box if you generally DISAGREE with the statement.

In checking the left or right box you are not necessarily indicating strong agreement or disagreement but just your general feeling most of the time.

There are also five general questions at the end.

Password: •••••

Statements 1 - 10 of 50.	**Agree**	**Undecided**	**Disagree**
This software responds too slowly to inputs.	○	○	○
I would recommend this software to my colleagues.	○	○	○
The instructions and prompts are helpful.	○	○	○
This software has at some time stopped unexpectedly.	○	○	○
Learning to operate this software initially is full of problems.	○	○	○
I sometimes don't know what to do next with this software.	○	○	○
I enjoy the time I spend using this software.	○	○	○
I find that the help information given by this software is not very useful.	○	○	○
If this software stops it is not easy to restart it.	○	○	○
It takes too long to learn the software functions.	○	○	○

Statements 11 - 20 of 50.	**Agree**	**Undecided**	**Disagree**
I sometimes wonder if I am using the right function.	○	○	○
Working with this software is satisfying.	○	○	○
The way that system information is presented is clear and understandable.	○	○	○
I feel safer if I use only a few familiar functions.	○	○	○
The software documentation is very informative.	○	○	○
This software seems to disrupt the way I normally like to arrange my work.	○	○	○
Working with this software is mentally stimulating.	○	○	○
There is never enough information on the screen when it's needed.	○	○	○
I feel in command of this software when I am using it.	○	○	○
I prefer to stick to the functions that I know best.	○	○	○

Statements 21 - 30 of 50.	**Agree**	**Undecided**	**Disagree**
I think this software is inconsistent.	○	○	○
I would not like to use this software every day.	○	○	○
I can understand and act on the information provided by this software.	○	○	○
This software is awkward when I want to do something which is not standard.	○	○	○
There is too much to read before you can use the software.	○	○	○
Tasks can be performed in a straight forward manner using this software.	○	○	○
Using this software is frustrating.	○	○	○
The software has helped me overcome any problems I have had in using it.	○	○	○
The speed of this software is fast enough.	○	○	○
I keep having to go back to look at the guides.	○	○	○

Statements 31 - 40 of 50.	**Agree**	**Undecided**	**Disagree**
It is obvious that user needs have been fully taken into consideration.	○	○	○
There have been times in using this software when I have felt quite tense.	○	○	○
The organisation of the menus seems quite logical.	○	○	○
The software allows the user to be economic of keystrokes.	○	○	○
Learning how to use new functions is difficult.	○	○	○
There are too many steps required to get something to work.	○	○	○
I think this software has sometimes given me a headache.	○	○	○
Error messages are not adequate.	○	○	○
It is easy to make the software do exactly what you want.	○	○	○
I will never learn to use all that is offered in this software.	○	○	○

Statements 41 - 50 of 50.	**Agree**	**Undecided**	**Disagree**
The software hasn't always done what I was expecting.	○	○	○
The software presents itself in a very attractive way.	○	○	○
Either the amount or quality of the help information varies across the system.	○	○	○
It is relatively easy to move from one part of a task to another.	○	○	○
It is easy to forget how to do things with this software.	○	○	○
This software occasionally behaves in a way which can't be understood.	○	○	○
This software is really very awkward.	○	○	○
It is easy to see at a glance what the options are at each stage.	○	○	○
Getting data files in and out of the system is not easy.	○	○	○
I have to look for assistance most times when I use this software.	○	○	○

What, in general, do you use this software for?

How important for you is the kind of software you have just been rating?

- Extremely important
- Important
- Not very important
- Not important at all

How would you rate your software skills and knowledge?

- Very experienced and technical
- I'm good but not very technical
- I can cope with most software
- I find most software difficult to use

What do you think is the best aspect of this software, and why?

What do you think needs most improvement, and why?

When you've answered all the questions, please click the 'Send' button.

Send

APPENDIX D

APPENDIX D: SEARCH TASKS

Search tasks given to participants during evaluation

1. isixhosa
2. amanzi
3. emzantsi afrika
4. umhlaba
5. ndikhathezekile indoda
6. hamba uye emaxhoseni
7. ndihlala emthatha nabantwana
8. inkxaso
9. isixeko sasekapa
10. isebe lezemihlaba empumalanga

APPENDIX E

APPENDIX E: SUMI RECORDS

E.1 Mobile Text Interface

A breakdown of ratings by the participants per the SUMI Global scale and the 5 SUMI subscales for the mobile text interface.

Participant	Global	Efficiency	Affect	Helpfulness	Control	Learnability
1	68	65	66	63	67	66
2	67	66	58	58	68	69
3	67	72	72	62	58	63
4	66	72	69	59	62	62
5	66	72	66	62	57	59
6	66	72	72	59	64	62
7	66	68	66	64	60	71
8	65	72	58	60	53	64
9	65	58	69	58	58	55
10	64	55	50	53	59	53
11	63	69	69	64	52	60
12	63	63	64	63	59	63
13	63	69	65	60	51	58
14	63	68	66	47	60	61
15	63	59	69	59	62	69
16	62	60	72	61	59	58
17	61	48	54	58	50	64
18	61	63	69	56	58	64
19	58	60	56	51	48	61
20	58	47	55	49	62	67
21	53	50	64	48	59	55
22	53	48	53	46	60	64
23	53	38	46	54	50	**34**
24	53	50	64	48	59	55
25	52	37	64	56	**33**	56
26	51	53	38	58	46	64
27	50	52	36	37	41	45
28	49	42	44	52	49	50
29	47	49	57	40	52	45
30	45	36	50	47	43	50
31	43	28	49	54	37	**32**
32	38	60	30	**29**	**23**	54

E.2 Mobile Voice Interface

A breakdown of ratings by the participants per the SUMI Global scale and the 5 SUMI subscales for the mobile voice interface.

Participant	Global	Efficiency	Affect	Helpfulness	Control	Learnability
1	70	66	63	67	57	58
2	67	59	66	64	54	59
3	66	61	64	64	58	66
4	64	55	64	64	55	56
5	63	57	52	57	56	64
6	61	46	72	55	58	65
7	61	60	54	44	49	50
8	61	44	66	70	55	54
9	60	60	59	55	54	58
10	57	62	72	56	48	61
11	57	62	56	51	58	63
12	56	61	52	39	65	66
13	55	54	41	58	53	64
14	55	59	55	58	61	64
15	52	54	55	46	49	57
16	51	36	60	57	54	46
17	50	34	54	53	48	63
18	50	48	50	50	46	39
19	49	41	52	35	57	52
20	46	42	33	61	36	54
21	45	44	40	37	49	61
22	43	48	22	58	41	36
23	43	23	53	58	33	36
24	43	46	52	53	46	44
25	42	35	38	55	43	56
26	41	49	41	30	59	50
27	40	29	38	53	48	38
28	40	31	40	49	51	38
29	38	49	46	51	40	41
30	36	26	35	47	37	39
31	31	17	40	53	38	23
32	29	28	38	35	50	42
33	18	12	29	37	30	29

E.3 Baseline Desktop Interface

A breakdown of ratings by the participants per the SUMI Global scale and the 5 SUMI subscales for the baseline desktop interface.

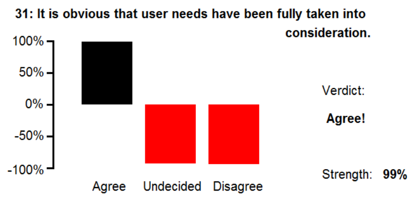
Participant	Global	Efficiency	Affect	Helpfulness	Control	Learnability
1	70	67	72	64	62	71
2	69	65	72	70	55	64
3	68	68	59	63	61	61
4	68	68	72	64	63	65
5	66	57	60	63	57	66
6	65	69	72	57	60	68
7	65	57	66	64	61	57
8	64	65	48	66	68	69
9	63	65	62	58	64	69
10	63	68	64	56	51	63
11	63	65	69	61	56	64
12	63	69	64	51	58	69
13	62	51	66	57	55	54
14	62	66	72	60	61	64
15	62	45	58	61	68	59
16	61	60	64	50	62	64
17	60	66	56	53	55	61
18	60	57	49	62	50	66
19	59	63	45	56	57	64
20	59	63	60	52	60	61
21	59	65	57	55	54	66
22	58	54	54	61	51	53
23	58	65	72	47	57	64
24	57	57	64	50	47	66
25	57	56	63	52	56	56
26	56	54	64	53	49	64
27	55	59	52	52	44	64
28	51	63	45	60	39	59
29	49	45	62	57	42	59
30	49	52	50	**31**	57	61
31	49	42	60	57	47	**44**
32	48	40	72	57	49	**45**
33	46	51	**33**	**38**	58	59
34	46	40	53	**37**	48	58

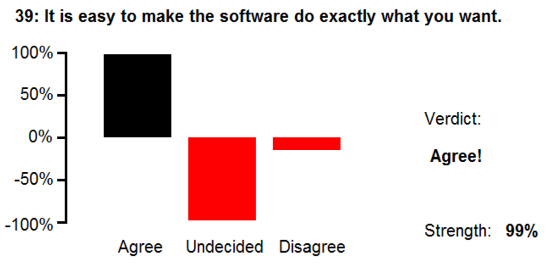
APPENDIX F

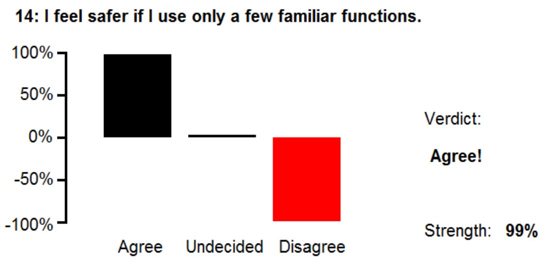
Appendix F: Strength and Weakness Analysis

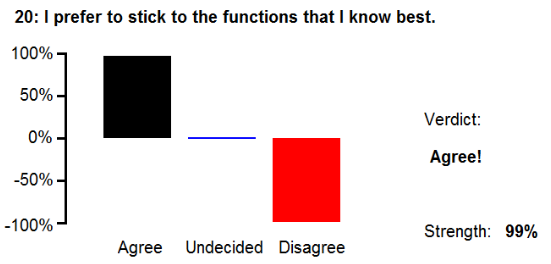
F.1 Mobile Text Interface

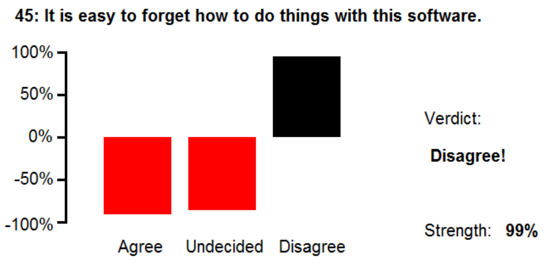
The strength and weakness analysis of the mobile voice interface as carried out by SUMI.

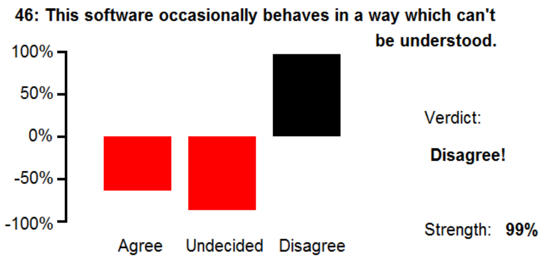


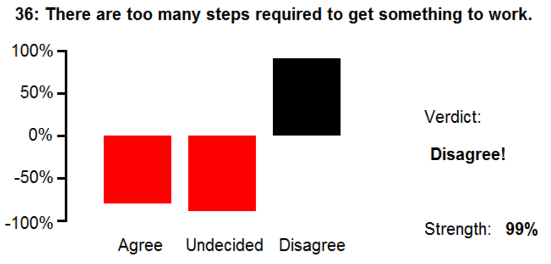


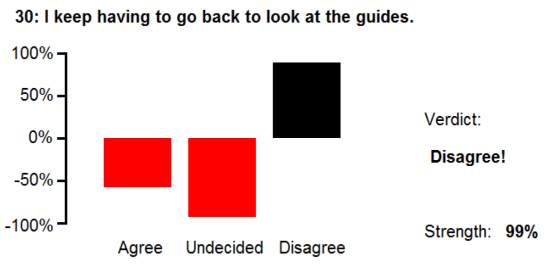


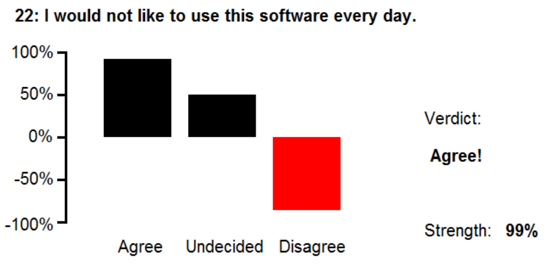

20: I prefer to stick to the functions that I know best.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

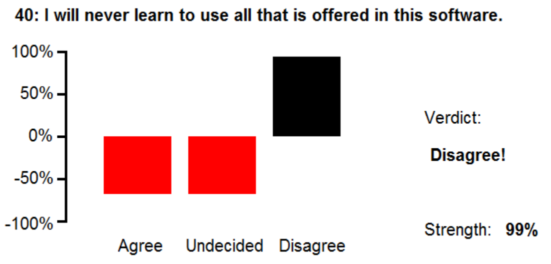

45: It is easy to forget how to do things with this software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

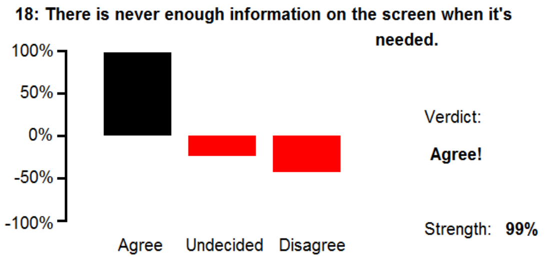

46: This software occasionally behaves in a way which can't be understood.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

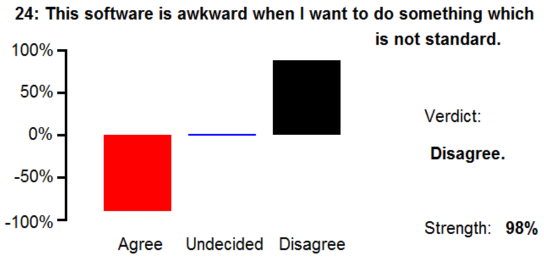

36: There are too many steps required to get something to work.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

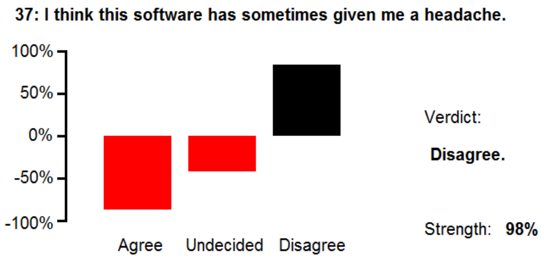

30: I keep having to go back to look at the guides.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

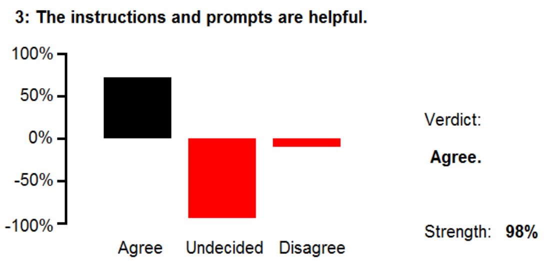

22: I would not like to use this software every day.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

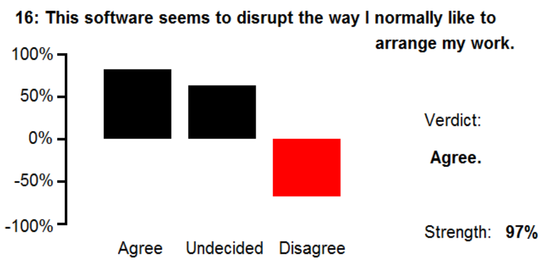

40: I will never learn to use all that is offered in this software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

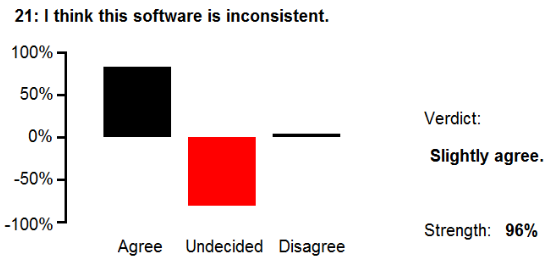

18: There is never enough information on the screen when it's needed.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

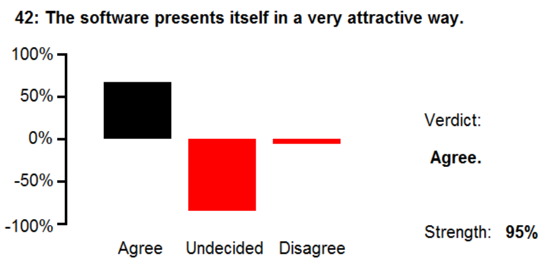

24: This software is awkward when I want to do something which is not standard.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree.
Strength: 98%


37: I think this software has sometimes given me a headache.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree.
Strength: 98%


3: The instructions and prompts are helpful.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 98%

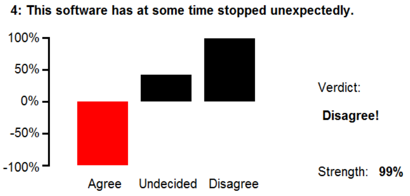

16: This software seems to disrupt the way I normally like to arrange my work.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 97%

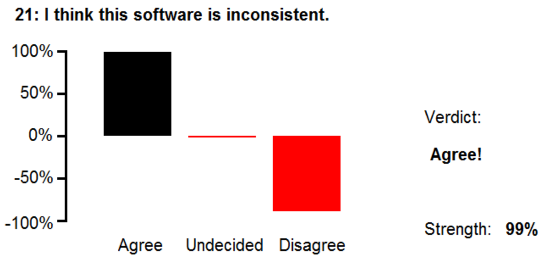

21: I think this software is inconsistent.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Slightly agree.
Strength: 96%

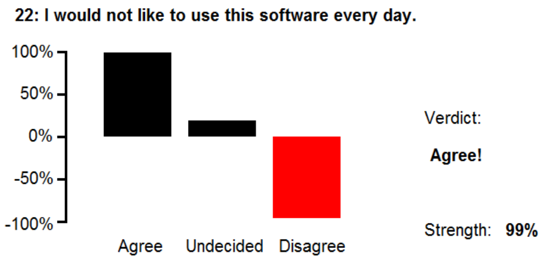

42: The software presents itself in a very attractive way.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 95%

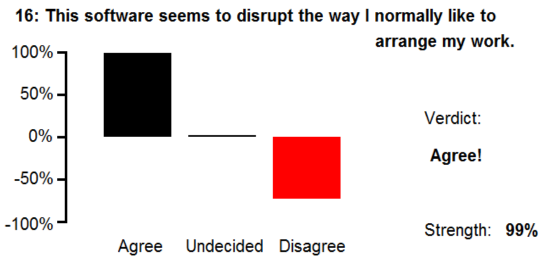
F.2 Mobile Voice Interface

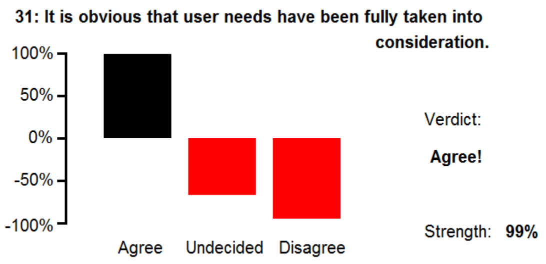
The strength and weakness analysis of the mobile voice interface as carried out by SUMI.

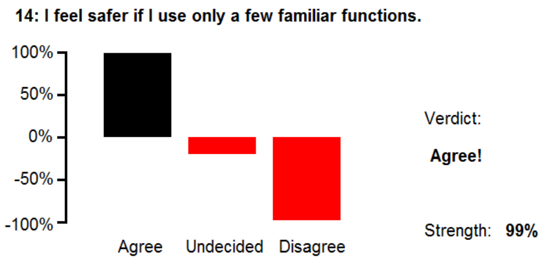




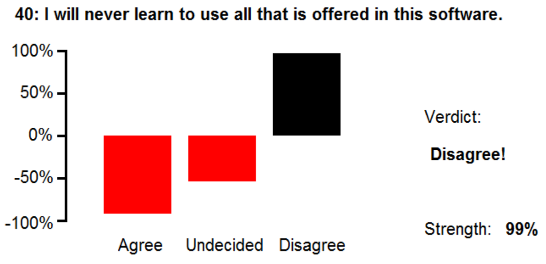


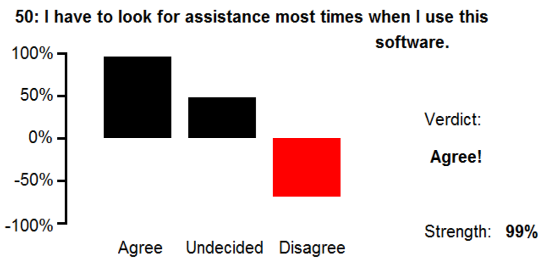


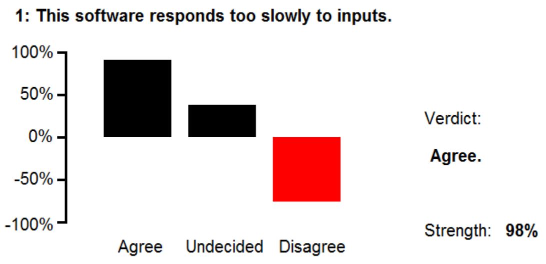

31: It is obvious that user needs have been fully taken into consideration.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

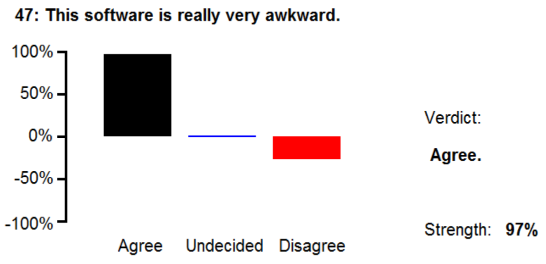

14: I feel safer if I use only a few familiar functions.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

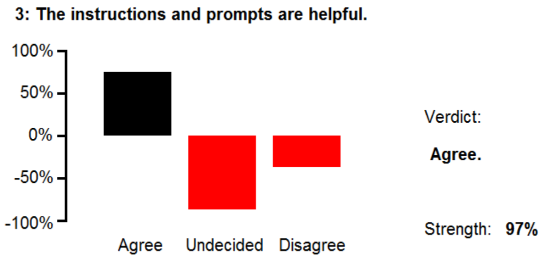

36: There are too many steps required to get something to work.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%


40: I will never learn to use all that is offered in this software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

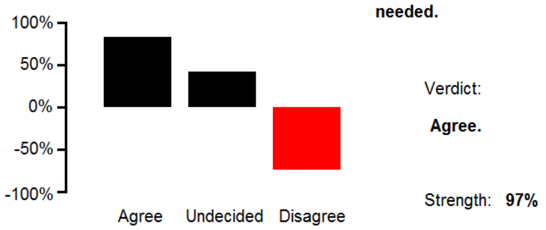

50: I have to look for assistance most times when I use this software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%


1: This software responds too slowly to inputs.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 98%

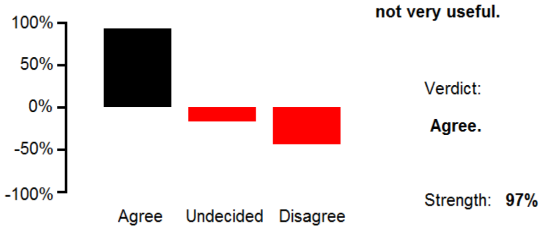

47: This software is really very awkward.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 97%


3: The instructions and prompts are helpful.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 97%

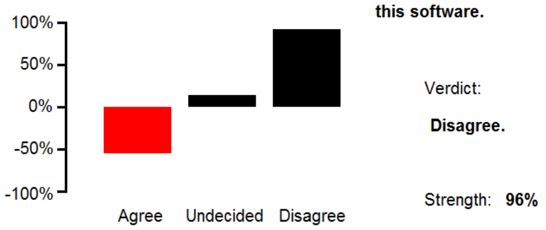
18: There is never enough information on the screen when it's needed.



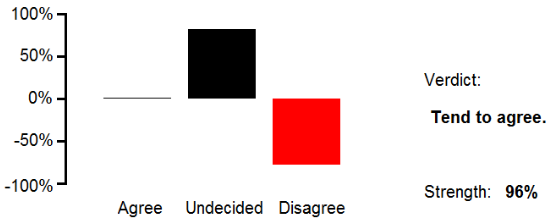
8: I find that the help information given by this software is not very useful.

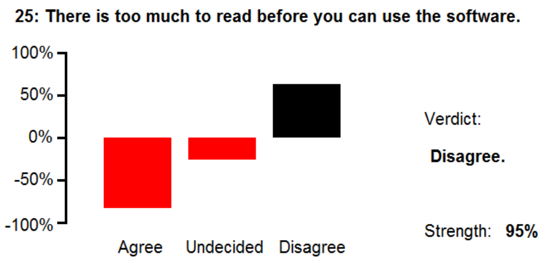


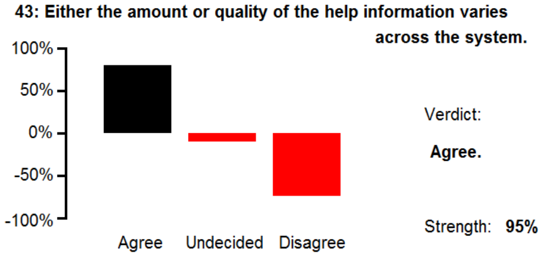
26: Tasks can be performed in a straight forward manner using this software.



49: Getting data files in and out of the system is not easy.

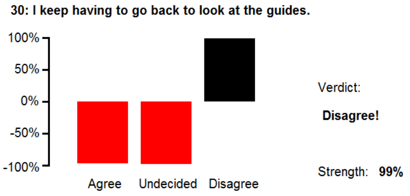


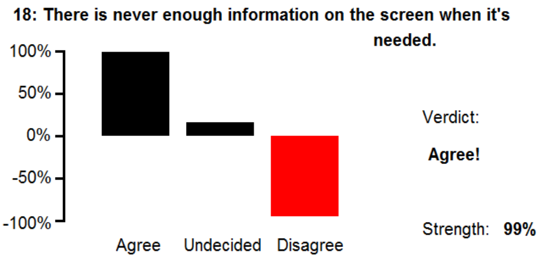

25: There is too much to read before you can use the software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree.
Strength: 95%

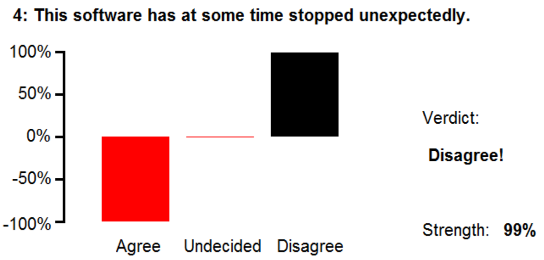

43: Either the amount or quality of the help information varies across the system.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree.
Strength: 95%

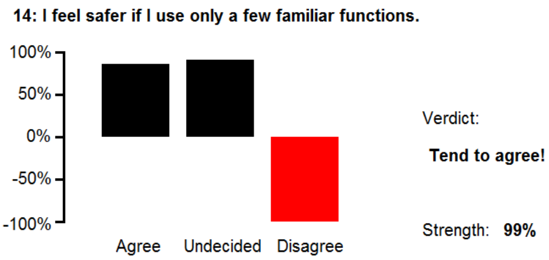
F.3 Baseline Desktop Interface

The strength and weakness analysis of the baseline desktop interface as carried out by SUMI.

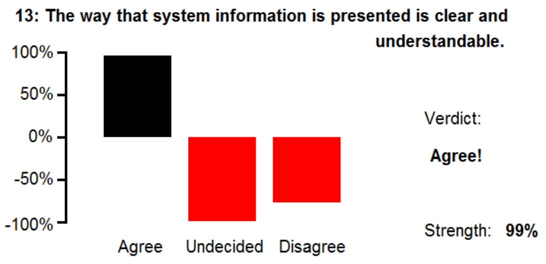


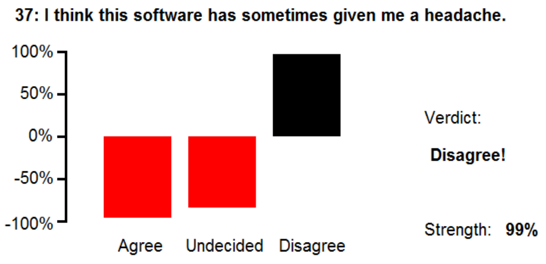


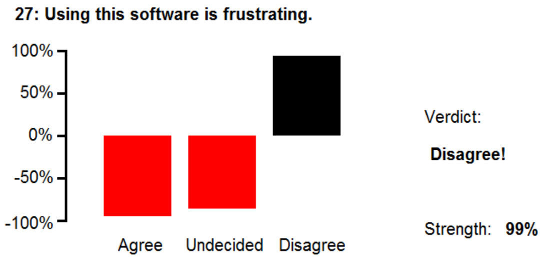


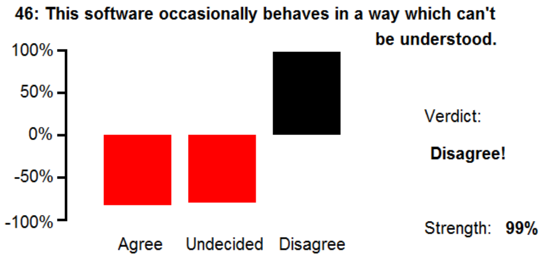


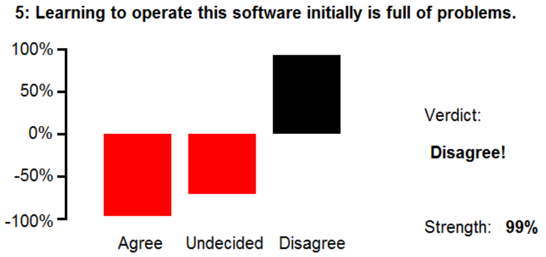

36: There are too many steps required to get something to work.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

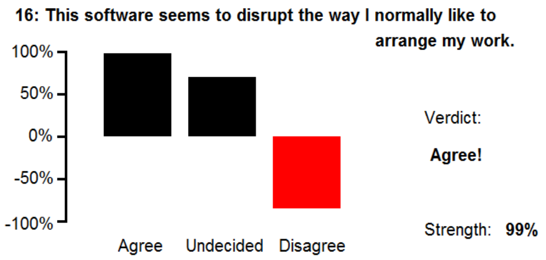

13: The way that system information is presented is clear and understandable.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

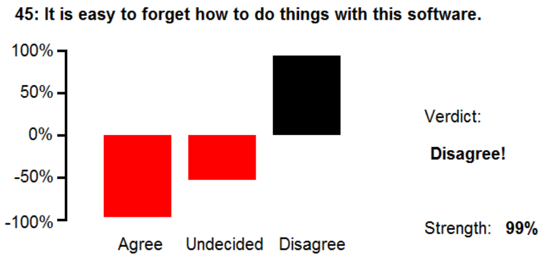

37: I think this software has sometimes given me a headache.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

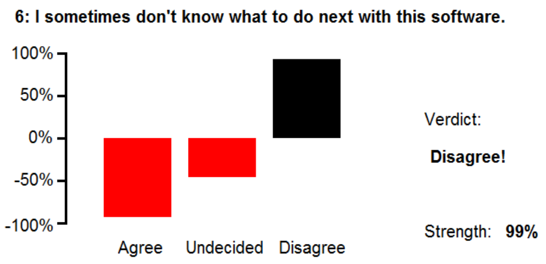

27: Using this software is frustrating.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

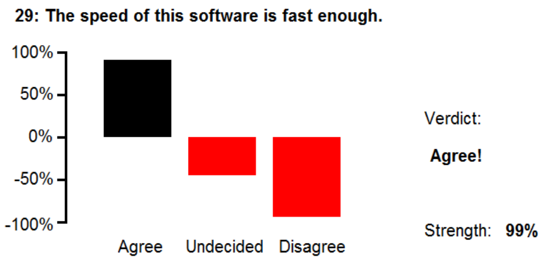

46: This software occasionally behaves in a way which can't be understood.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

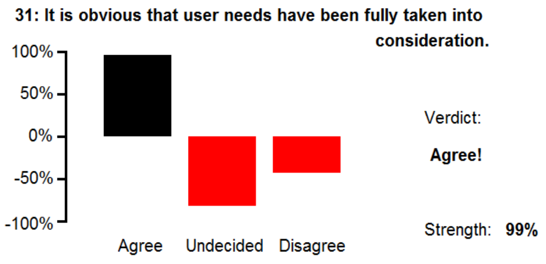

5: Learning to operate this software initially is full of problems.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

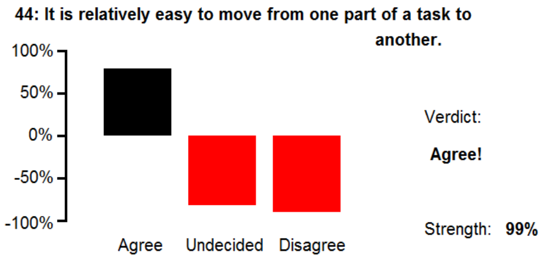

16: This software seems to disrupt the way I normally like to arrange my work.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%


45: It is easy to forget how to do things with this software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%

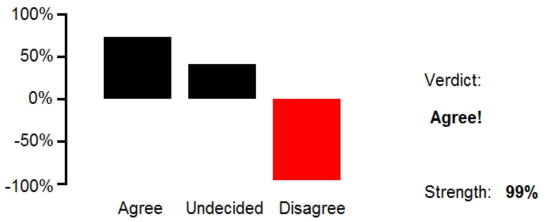

6: I sometimes don't know what to do next with this software.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Disagree!
Strength: 99%


29: The speed of this software is fast enough.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

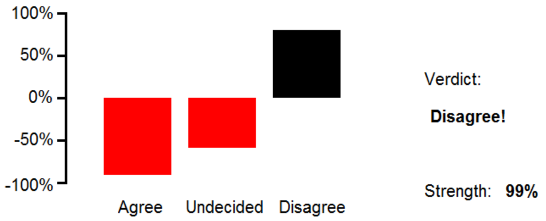

31: It is obvious that user needs have been fully taken into consideration.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%


44: It is relatively easy to move from one part of a task to another.
100%
50%
0%
-50%
-100%
Agree
Undecided
Disagree
Verdict:
Agree!
Strength: 99%

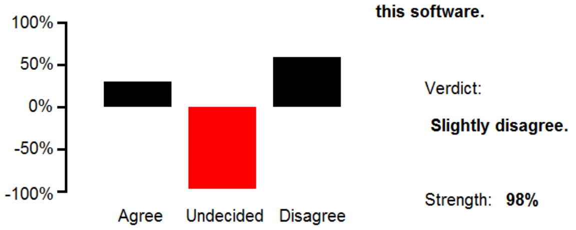
20: I prefer to stick to the functions that I know best.



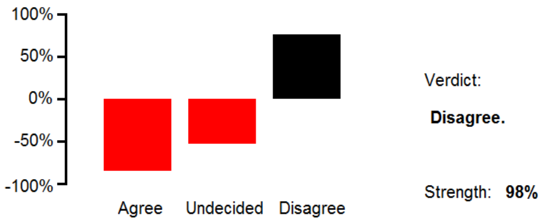
1: This software responds too slowly to inputs.



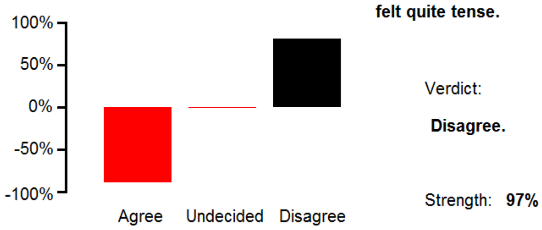
23: I can understand and act on the information provided by this software.



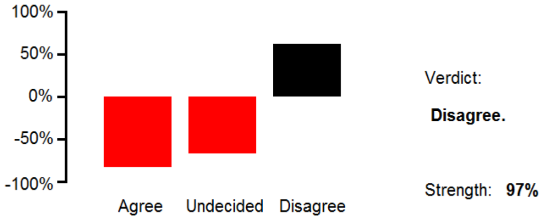
25: There is too much to read before you can use the software.



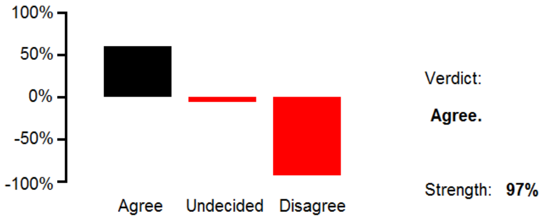
32: There have been times in using this software when I have felt quite tense.



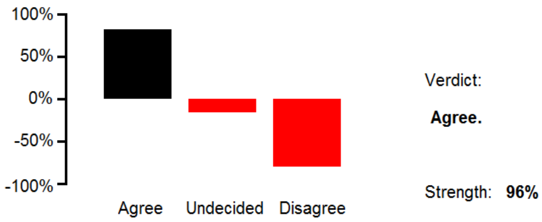
10: It takes too long to learn the software functions.



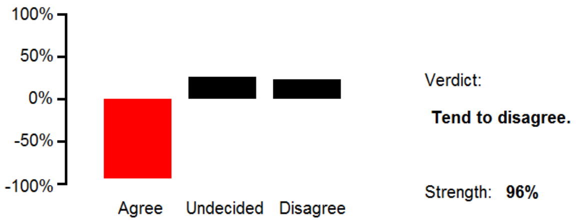
19: I feel in command of this software when I am using it.



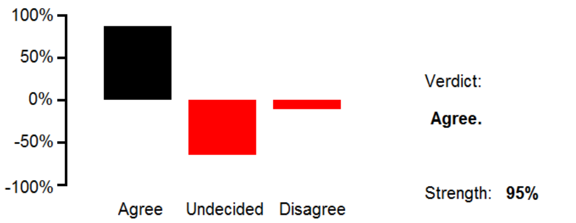
39: It is easy to make the software do exactly what you want.



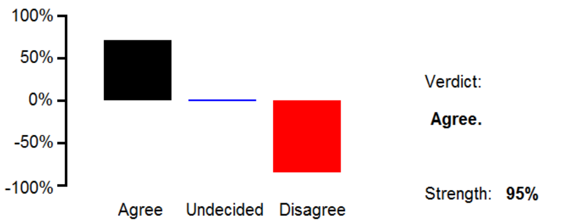
9: If this software stops it is not easy to restart it.



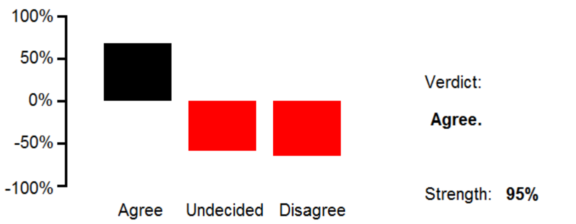
15: The software documentation is very informative.



17: Working with this software is mentally stimulating.



3: The instructions and prompts are helpful.



APPENDIX G

APPENDIX G: SPEED DATA

G.1 Mobile Text Interface

Speed data for the mobile text interface.

Participant	isiXhosa	amanzi	emzantsi afrika	umhlaba	ndikhathezekil e indoda	hamba uye emaxhoseni	ndihlala emthata nabantwana	inkxaso	isixeko sasekapa	isebe lezemihlaba empumalanga
1	3	13	4	5	11	5	5	5	5	8
2	3	2	3	3	0	8	0	7	9	0
3	3	7	0	10	14	8	16	2	0	0
4	0	4	5	4	5	3	0	7	8	8
5	6	5	6	3	3	5	8	5	2	8
6	0	0	0	0	0	0	0	0	0	47
7	18	7	5	2	5	8	3	5	3	5
8	9	10	4	0	0	0	0	0	0	0
9	12	4	16	6	5	6	9	9	11	14
10	7	5	8	13	4	10	9	8	4	0
11	3	2	2	2	6	1	2	3	1	29
12	4	3	7	6	8	4	12	5	3	13
13	0	7	0	0	12	11	0	22	7	0
14	14	4	5	2	4	10	0	8	9	4
15	3	2	5	2	2	5	5	7	2	0
16	5	3	5	4	5	6	6	9	10	3
17	5	10	3	5	5	4	5	4	2	4
18	5	3	7	3	16	6	3	5	4	4
19	10	3	2	4	3	3	6	5	4	4
20	12	6	9	8	5	7	16	5	5	14
21	5	5	3	2	5	6	3	3	1	5
22	6	3	0	0	0	10	9	0	0	0
23	29	8	21	6	34	9	0	38	19	7
24	8	2	2	3	3	2	1	2	1	2
25	62	6	18	5	3	9	4	28	6	12
26	12	15	15	7	10	5	3	2	5	8
27	0	0	0	0	0	0	0	0	0	0
28	8	5	0	10	8	0	3	10	11	4
29	5	6	9	7	5	11	0	21	7	0
30	0	0	0	0	0	0	0	0	0	0
31	0	5	0	0	0	0	0	0	0	0
32	44	8	14	6	4	6	0	13	3	4
33	6	4	6	4	4	0	0	0	0	0
34	11	6	7	0	14	16	0	13	12	9

R script used to generate boxplot from mobile text interface speed data above.

```
baseline.speed = read.csv("/home/mrmodise/Desktop/SpeedResults-Mobile-Text.csv",
header = TRUE)
boxplot(baseline.speed[, 2:11],yaxp = c(0,62,62),cex.axis=1, las=3,ylab = "Time
in seconds", xaxt="n",
        col="orange", border="darkblue")

lablist <-
  c(
    "isiXhosa",
    "amanzi",
    "emzantsi\nafrika",
    "umhlaba",
    "ndikhathezekile\nindoda",
    "hamba\nuye\nemaxhoseni",
    "ndihlala\nemthatha\nnabantwana",
    "inkxaso",
    "isixeko\nsasekapa",
    "isebe\nlezemihlaba\nempumalanga"
  )

labels <- paste(c(lablist))

text(x = seq_along(labels),y = par("usr")[3] - 0.1,srt = 45,adj =
1,cex=.85,labels = labels,xpd = TRUE)
```

G.2 Mobile Voice Interface

Speed data for the mobile voice interface with related R script used to generate boxplot.

Participant	isiXhosa	amanzi	emzantsi afrika	umhlaba	ndikhathezekile indoda	hamba uye emaxhoseni	ndihlala emthata nabantwana	inkxaso	isixeko sasekapa	isebe lezemihlaba empumalanga
1	13	7	5	8	2	5	20	5	10	15
2	6	6	6	6	8	8	7	4	5	9
3	12	0	16	7	0	9	9	0	12	11
4	11	5	12	0	0	10	18	11	6	6
5	10	10	8	0	10	11	14	13	11	9
6	9	10	6	4	0	5	0	7	5	8
7	4	4	7	5	9	5	6	3	6	0
8	0	14	0	6	0	6	7	6	10	9
9	7	7	10	7	0	8	0	10	8	0
10	13	4	13	0	0	7	9	8	0	0
11	0	12	12	8	10	15	9	9	7	10
12	12	10	15	7	13	9	7	8	7	9
13	12	6	0	0	0	0	0	12	0	0
14	10	5	11	7	11	9	9	6	8	8
15	5	10	8	0	0	0	0	0	0	0
16	11	12	10	8	11	9	9	3	8	0
17	9	13	7	6	9	12	11	6	10	12
18	10	13	9	11	9	7	7	7	6	9
19	15	10	7	0	0	7	8	6	7	0
20	5	4	7	5	9	11	7	5	6	7
21	10	12	25	9	0	0	0	12	11	0
22	12	11	11	9	9	13	8	4	14	8
23	14	9	9	11	18	14	15	17	15	16
24	10	11	9	8	12	7	10	16	10	9
25	8	0	13	0	12	12	12	13	12	13
26	8	9	7	10	0	0	0	11	0	0
27	13	8	11	7	48	12	11	9	8	0
28	8	8	8	5	8	8	0	9	9	11
29	12	5	5	8	0	5	7	3	7	7
30	7	8	9	3	8	8	5	6	6	6
31	7	7	7	8	10	7	9	6	5	9
32	12	12	9	9	11	8	14	0	12	17
33	7	7	8	8	9	11	11	9	6	10
34	8	14	12	10	8	6	0	5	8	5

```
baseline.speed = read.csv("/home/mrmodise/Desktop/SpeedResults-Mobile-
Voice.csv", header = TRUE)
boxplot(baseline.speed[, 2:11],yaxp = c(0,48,48),cex.axis=1, las=3,ylab = "Time
in seconds", xaxt="n",
        col="orange", border="darkblue")

lablist <-
  c(
    "isiXhosa",
    "amanzi",
    "emzantsi\nafrika",
    "umhlaba",
    "ndikhathezekile\nindoda",
    "hamba\nuye\nemaxhoseni",
    "ndihlala\nemthatha\nnabantwana",
    "inkxaso",
    "isixeko\nsasekapa",
    "isebe\nlezemihlaba\nempumalanga"
  )

labels <- paste(c(lablist))

text(x = seq_along(labels),y = par("usr")[3] - 0.1,srt = 45,adj =
1,cex=.85,labels = labels,xpd = TRUE)
```

G.3 Baseline Desktop Interface

Speed data for the baseline desktop interface with related R script used to generate boxplot.

Participant	isiXhosa	amanzi	emzantsi afrika	umhlaba	ndikhathezekil e indoda	hamba uye emaxhoseni	ndihlala emthata nabantwana	inkxaso	isixeko sasekapa	isebe lezemihlaba empumalanga
1	11	10	7	4	1	2	2	1	2	2
2	11	13	14	7	5	23	9	9	10	4
3	3	3	0	0	4	11	9	6	5	9
4	0	7	0	0	0	0	13	5	25	7
5	56	4	6	4	2	10	4	3	12	9
6	37	0	5	5	7	0	0	0	0	0
7	5	5	2	4	3	4	6	2	15	2
8	13	5	6	7	4	6	3	3	2	2
9	11	5	13	4	11	9	0	0	29	8
10	8	5	3	3	3	0	0	0	0	6
11	13	3	4	3	5	4	5	5	20	13
12	0	0	0	4	27	4	0	2	17	4
13	6	3	5	14	4	6	0	7	21	14
14	2	2	3	5	0	0	0	0	0	0
15	42	5	11	6	5	6	4	4	11	9
16	6	3	5	9	17	25	0	8	6	15
17	6	6	10	14	3	2	2	2	3	2
18	15	3	5	7	24	13	9	4	4	5
19	7	17	5	3	12	11	15	18	8	5
20	7	7	0	0	5	2	21	34	16	8
21	17	28	6	10	5	15	4	5	5	10
22	7	3	22	4	3	0	0	22	14	0
23	25	14	35	23	78	102	29	22	138	97
24	8	7	4	0	21	5	0	4	0	14
25	13	4	20	6	19	53	0	11	0	0
26	14	3	7	9	8	5	8	16	5	8
27	3	5	5	3	10	7	5	5	4	2
28	5	7	11	0	0	0	12	0	0	5
29	15	17	18	3	14	8	0	4	17	0
30	7	5	12	4	8	24	0	6	5	0
31	4	21	4	13	10	3	12	2	6	11
32	0	129	8	9	5	12	0	4	6	7
33	4	6	5	7	3	6	8	5	8	4
34	10	114	0	6	11	0	0	0	11	0

```
baseline.speed = read.csv("/home/mrmodise/Desktop/SpeedResults-Desktop.csv",
header = TRUE)
boxplot(baseline.speed[, 2:11],yaxp = c(0,140,70),cex.axis=1, las=3,ylab = "Time
in seconds", xaxt="n",
        col="orange", border="darkblue")

lablist <-
  c(
    "isiXhosa",
    "amanzi",
    "emzantsi\nafrika",
    "umhlaba",
    "ndikhathezekile\nindoda",
    "hamba\nuye\nemaxhoseni",
    "ndihlala\nemthatha\nnabantwana",
    "inkxaso",
    "isixeko\nsasekapa",
    "isebe\nlezemihlaba\nempumalanga"
  )

labels <- paste(c(lablist))

text(x = seq_along(labels),y = par("usr")[3] - 0.1,srt = 45,adj =
1,cex=.85,labels = labels,xpd = TRUE)
```

APPENDIX H

APPENDIX H: RELEVANCE DATA

H.1 Mobile Text Interface

Relevance data for the mobile text interface

Participant	isiXhosa	amanzi	emzantsi afrika	umhlaba	ndikhathezekile indoda	hamba uye emaxhoseni	ndihlala emthata nabantwana	inkxaso	isixeko sasekapa	isebe lezemihlaba empumalanga
1	1	1	1	1	1	0	0	0	0	1
2	1	1	1	1	0	0	0	1	0	0
3	1	2	0	1	1	1	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0
5	2	1	1	4	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	1
7	1	1	1	2	1	1	1	1	1	1
8	0	0	1	0	0	0	0	0	0	0
9	0	1	0	0	0	0	0	0	0	0
10	1	1	0	0	0	0	0	0	0	0
11	2	3	3	2	2	2	2	2	2	2
12	2	1	1	1	1	1	1	1	1	1
13	0	1	0	0	1	0	0	1	1	0
14	1	1	1	1	3	1	0	1	1	1
15	3	1	0	1	1	0	1	0	1	0
16	1	1	1	1	1	0	1	1	1	1
17	2	2	1	1	1	1	7	1	1	1
18	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	2	0
21	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	0	0	0	0	0	0
23	1	1	1	2	1	0	0	1	1	1
24	0	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1
26	0	0	0	0	0	1	1	1	1	3
27	0	0	0	0	0	0	0	0	0	0
28	1	2	0	1	0	0	0	0	0	3
29	2	1	2	1	2	1	0	0	1	0
30	0	0	0	0	0	0	0	0	0	0
31	0	1	0	0	0	0	0	0	0	0
32	1	1	0	1	0	1	0	0	1	1
33	1	1	1	1	1	0	0	0	0	0
34	1	2	0	0	0	0	0	0	0	0

H.2 Mobile Voice Interface

Relevance data for the mobile voice interface

Participant	isiXhosa	amanzi	emzantsi afrika	umhlaba	ndikhathezekil e indoda	hamba uye emaxhoseni	ndihlala emthata nabantwana	inkxaso	isixeko sasekapa	isebe lezemihlaba empumalanga
1	0	0	0	0	0	0	0	0	0	0
2	1	0	2	1	0	1	1	1	1	1
3	1	0	1	0	0	1	1	0	1	1
4	1	1	0	0	0	0	3	1	0	0
5	2	1	2	0	0	1	2	1	1	2
6	0	0	2	3	0	2	0	1	1	0
7	1	3	1	0	0	1	1	1	1	0
8	0	1	0	0	0	0	0	1	0	0
9	1	0	1	1	0	1	0	1	1	0
10	2	1	1	0	0	1	1	1	0	0
11	0	1	1	1	0	0	1	1	1	2
12	0	1	0	1	0	1	1	1	1	1
13	3	0	0	0	0	0	0	0	0	0
14	2	1	1	2	0	1	1	1	1	2
15	2	2	6	0	0	0	0	0	0	0
16	2	4	1	2	0	3	0	0	1	0
17	1	1	1	1	0	1	1	1	0	1
18	1	0	0	0	0	0	0	0	0	0
19	1	0	3	0	0	1	1	0	1	0
20	1	1	1	1	0	1	1	1	1	1
21	1	0	0	0	0	0	0	0	0	0
22	2	1	0	2	0	0	0	0	0	0
23	1	3	2	1	0	0	1	1	1	0
24	0	0	0	0	0	0	0	1	0	0
25	1	0	1	0	0	1	1	0	1	1
26	1	1	2	1	0	0	0	1	0	0
27	1	1	1	0	0	1	2	1	1	0
28	3	0	2	0	0	1	0	1	0	0
29	1	1	1	1	0	1	0	1	1	1
30	1	1	0	1	0	0	0	0	1	0
31	1	1	1	2	0	1	1	0	0	1
32	0	0	0	0	0	0	0	0	0	2
33	2	1	0	1	0	1	1	1	1	0
34	1	0	0	2	0	5	0	1	1	1

H.3 Baseline Desktop Interface

Relevance data for the baseline desktop interface

Participant	isiXhosa	amanzi	emzantsi afrika	umhlaba	ndikhathezekil e indoda	hamba uye emaxhoseni	ndihlala emthata nabantwana	inkxaso	isixeko sasekapa	isebe lezemihlaba empumalanga
1	1	1	1	1	1	1	1	1	1	1
2	1	1	2	1	1	1	2	1	5	2
3	1	1	0	1	1	1	1	1	1	1
4	0	1	0	0	0	0	2	1	1	1
5	1	1	2	1	1	1	1	2	3	3
6	1	0	2	1	1	0	0	0	0	0
7	1	1	1	1	1	0	1	1	1	1
8	4	2	4	4	1	4	3	3	1	3
9	1	1	1	1	1	1	0	1	1	1
10	2	1	2	1	1	0	0	0	0	1
11	2	3	3	2	2	3	2	3	4	5
12	0	0	0	1	1	1	0	1	1	1
13	0	0	0	0	1	1	0	2	1	2
14	2	2	1	1	0	0	0	0	0	0
15	3	4	2	3	2	2	2	1	1	1
16	2	3	2	2	1	1	0	1	1	1
17	1	1	2	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	2
19	1	1	1	1	1	1	1	1	2	1
20	1	2	0	0	1	5	2		3	3
21	2	1	2	1	1	1	1	2	1	1
22	2	1	2	1	1	0	0	1	1	0
23	1	1	1	1	1	1	2	1	1	1
24	1	1	1	0	1	1	0	1	0	1
25	1	1	1	1	1	1	0	1	0	0
26	1	1	1	2	1	1	1	1	1	3
27	2	1	1	1	2	1	1	1	1	1
28	1	1	1	0	0	0	1	0	0	1
29	1	1	2	1	1	1	0	1	1	0
30	1	1	3	1	1	1	0	1	1	0
31	1	1	5	0	1	3	3	3	4	3
32	0	3	2	1	1	2	0	1	1	1
33	1	1	2	1	1	1	1	1	2	2
34	2	1	0	2	1	0	0	0	0	1

APPENDIX I

APPENDIX I: WILCOXON TEST R SCRIPT

The R script used to perform Wilcoxon sum rank tests on the three interfaces.

```r
# Input from file on PC
textResults = read.csv("/home/mrmodise/Desktop/Analysed-Mobile-Text-Results.csv", header = TRUE)
desktopResults = read.csv("/home/mrmodise/Desktop/Analysed-Desktop-Results.csv", header = TRUE)
voiceResults = read.csv("/home/mrmodise/Desktop/Analysed-Mobile-Voice-Results.csv", header = TRUE)

# Summarized data
summary(desktopResults)
summary(voiceResults)
summary(textResults)

# Desktop Vs Mobile Voice Interface
wilcox.test(desktopResults$Global, voiceResults$Global, exact = FALSE)
wilcox.test(desktopResults$Efficiency, voiceResults$Efficiency, exact = FALSE)
wilcox.test(desktopResults$Affect, voiceResults$Affect, exact = FALSE)
wilcox.test(desktopResults$Helpfulness, voiceResults$Helpfulness, exact = FALSE)
wilcox.test(desktopResults$Control, voiceResults$Control, exact = FALSE)
wilcox.test(desktopResults$Learnability, voiceResults$Learnability, exact = FALSE)

# Desktop Vs Mobile Text Interface
wilcox.test(desktopResults$Global, textResults$Global, exact = FALSE)
wilcox.test(desktopResults$Efficiency, textResults$Efficiency, exact = FALSE)
wilcox.test(desktopResults$Affect, textResults$Affect, exact = FALSE)
wilcox.test(desktopResults$Helpfulness, textResults$Helpfulness, exact = FALSE)
wilcox.test(desktopResults$Control, textResults$Control, exact = FALSE)
wilcox.test(desktopResults$Learnability, textResults$Learnability, exact = FALSE)

# Mobile Voice Interface Vs Mobile Text Interface
wilcox.test(voiceResults$Global, textResults$Global, exact = FALSE)
wilcox.test(voiceResults$Efficiency, textResults$Efficiency, exact = FALSE)
wilcox.test(voiceResults$Affect, textResults$Affect, exact = FALSE)
wilcox.test(voiceResults$Helpfulness, textResults$Helpfulness, exact = FALSE)
wilcox.test(voiceResults$Control, textResults$Control, exact = FALSE)
wilcox.test(voiceResults$Learnability, textResults$Learnability, exact = FALSE)
```

APPENDIX J

APPENDIX J: NORMALITY Q-Q PLOTS

The Normality Q-Q Plots for different SUMI attributes on the three interfaces.

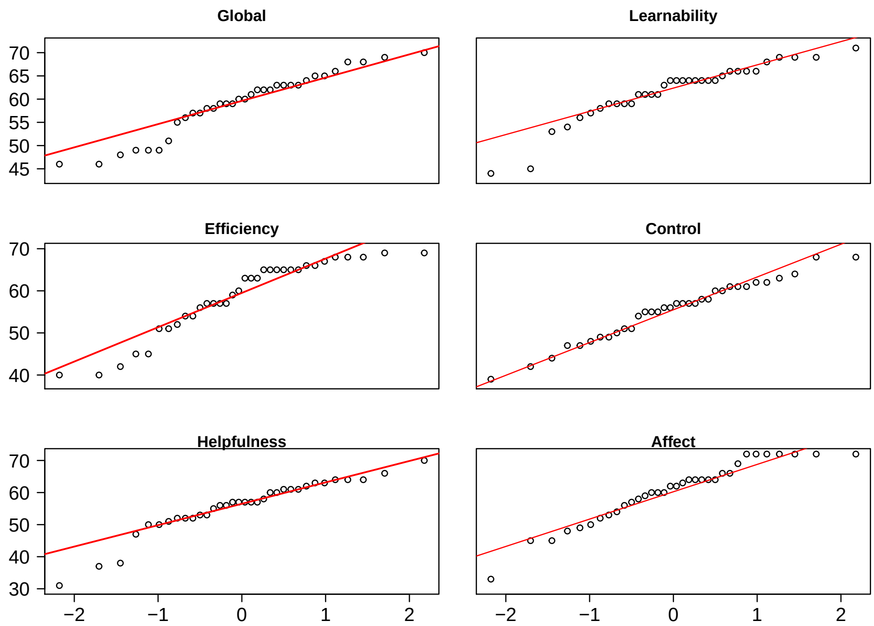


Figure J.1: Baseline Desktop Interface

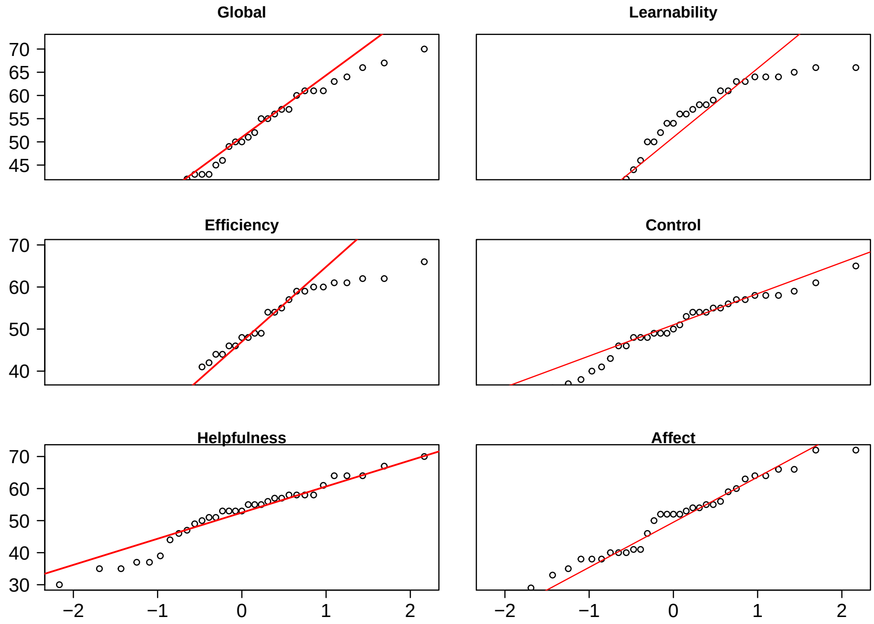


Figure J.2: Mobile Voice Interface

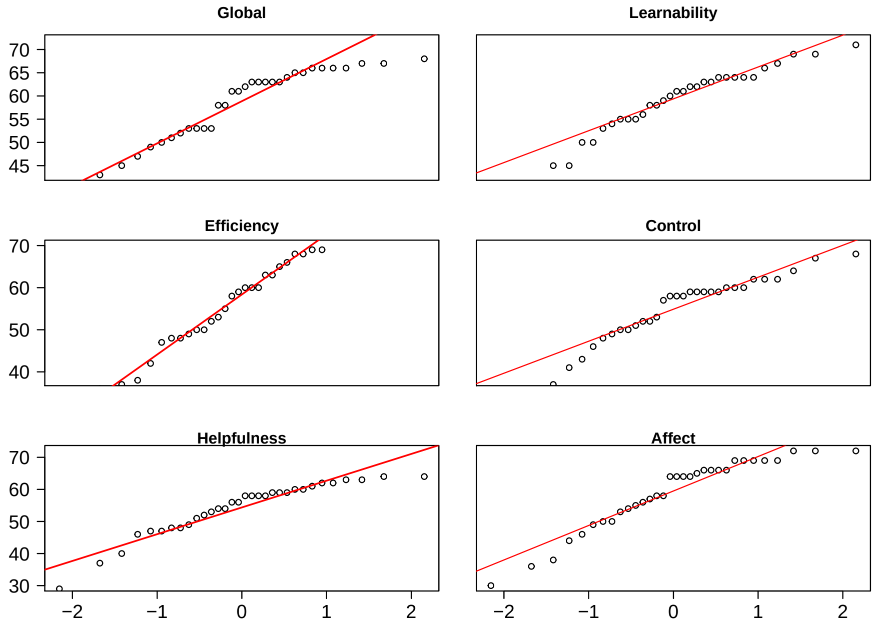


Figure J.3: Mobile Text Interface

APPENDIX

K

APPENDIX K: SHAPIRO-WILK TEST R SCRIPT

The Shapiro-Wilk test for normality used to test the interfaces' data for normality.

```r
# read CSV files from PC drive
textResults = read.csv("/home/mrmodise/Desktop/Analysed-Mobile-Text-Results.csv", header = TRUE)
desktopResults = read.csv("/home/mrmodise/Desktop/Analysed-Desktop-Results.csv", header = TRUE)
voiceResults = read.csv("/home/mrmodise/Desktop/Analysed-Mobile-Voice-Results.csv", header = TRUE)

# provide summary of each CSV read
summary(desktopResults)
summary(voiceResults)
summary(textResults)

# run shapiro tests on each interface data
lshapDesktop <- lapply(desktopResults, shapiro.test)
lshapVoice <- lapply(voiceResults, shapiro.test)
lshapText <- lapply(textResults, shapiro.test)

# retrieve the statistic and p-value for each of the SUMI subscales
lresText <- sapply(lshapText, `[`, c("statistic", "p.value"))
lresVoice <- sapply(lshapVoice, `[`, c("statistic", "p.value"))
lresDesktop <- sapply(lshapDesktop, `[`, c("statistic", "p.value"))

# neatly format output
t(lresText);
t(lresVoice);
t(lresDesktop);
```