An evaluation of conversational interfaces for pedestrian navigation

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Information Technology

Institute of Natural and Mathematical Sciences
Massey University, Albany
New Zealand

NATTAKAN LONGPRASERT
2017

An evaluation of conversational interfaces for pedestrian navigation

Abstract

The aim of this research was to compare the performance between the OsmAnd application and three types of conversational interface, to test whether the conversational interface is a more preferred navigation tool. We designed and tested four different navigation systems; the map with command interface, the conversational-only interface, the conversational with map interface, and the conversational with image interface. The research involved 100 participants who had different levels of experience when using navigation systems. Participants were divided into three groups and were given different navigation interfaces. This research was conducted with both quantitative and qualitative usability testing along a pre-defined route in Massey University campus, combined with a USE questionnaire to gain the user's feedback. The results indicated that both the OsmAnd and the conversational interface were good in different criteria. However, most participants preferred using the conversational interface more than the visual interface.

Key Words: conversational interface, pedestrian navigation system

Acknowledgments

I would like to thank many people who have supported me during my Masters thesis. I would like to thank my family and my partner for their support, love and understanding during my most difficult times. Without them I would not be myself now.

I would like to thank and express my gratitude to my supervisor Dr. Kristin Stock for her guidance and encouragement in many aspects of my Masters study. She has always been supportive, has answered all my questions and has cheered me up and re-built my confidence in difficult times.

I would like to thank Massey University for providing me the opportunity to complete this Masters thesis and enhance myself. I also thank Annette for providing me vouchers whenever I asked for them.

My thanks to all participants who helped me run the survey. Without them I would not have been able to finish my study.

I would like to thank my manager and colleagues in the Leaps and Bounds Early Learning Centre for giving me support and understanding during my study.

Lastly, I would like to thank all my friends who have cheered me up and helped me during my study. Especially, Frank who helped me to sort my coding out.

Dedication

My Father and Mother: Thanks for all your love and support you both have given me. I am so proud to be your daughter. Every time I feel down, I was thinking of both of you that made me want to move forward. Thanks for believing in me.

My Aunty Nee: Thanks for your big support and for guiding me about life and the importance of study. Without your support, I would not have been able to study overseas which I never expected to have this opportunity in my life.

My brother (Golf) and sister (Gift): Thanks both of you for being a lovely brother and sister. I am glad to be your big sister. Thanks, Golf for taking care of our parents and our sister while I am miles away.

My other half: Thanks for everything, Joey. Thanks for helping me settle down in Auckland. Thanks for your all support and always being by my side in my most difficult times. Thanks for wiping up my tears when I cried and for laughing with me when I was happy.

I love you all….

Contents

List of figures

List of tables

Chapter 1

Introduction

Navigation systems can help people with positioning, navigating and route planning. The number of people using navigation systems has increased dramatically as people use them for many purposes, including driving, walking, hiking and cycling (Park et al., 2012; Chou et al., 2006; Arentze et al., 2012). Navigation systems are a technology that make people's lives easier, especially when finding their way around unfamiliar places and when they need to arrive at a specific destination on time. Most navigation systems use visual modalities, containing maps, text and images (e.g. OsmAnd, OpenStreetMap, and Google Maps) which may be displayed in a readable format (text) and combined with voice commands to provide directional information to the user.

In the past few years, the new class of conversational interfaces includes speech-language technology, natural language processing. Mobile applications have high interest and are unique in supporting speech recognition between humans and computers (Darves et al., 2002). Our research has focused on conversational interfaces between the user and mobile devices, through an android application, in order to create an alternative method of finding directions for all classes of users when using their natural language. However, conversational technology for navigation systems still focuses on one-way communication. This means that only the commands of the speaker are voiced and the listener has no way to respond to the conversation or to ask questions. Many conversational systems have been investigated and deployed (Yoshino & Kawahara, 2015; Geutner et al., 1998; Zue et al., 2000), and some navigation systems have adopted conversational technology to improve the capability of the system and make it more useful to the user.

1.1 Background

This section introduces all kinds of navigation systems that we have used in the research and explanations of the features and reasons of selecting. The last part is also pointed out advantages and disadvantages of navigation devices in order to show the strengths and the weaknesses of navigation systems to the users

1.1.1 The OsmAnd Map

Figure 1 The OsmAnd application

OsmAnd (Figure 1) is an application that uses visual navigation modalities, combined with map and text on the screen, providing a number of necessary functions for the user on Android and iOS platforms. For instance, turn-by-turn directions, traffic reports (e.g. stop signs, pedestrian crosswalks, and speed reminders), optional lanes and estimated time of arrival. It also provides navigation modes for car, bicycle and pedestrian and displays public transport stops containing line names. OsmAnd is an open source application actively being developed and all data contributes to the OpenStreetMap database. For this research, we used the OsmAnd application as a visual modality to compare with a non-visual modality (the conversational-only interface). The reasons for this application selection are:

- This is an open source application which everyone can contribute to by reporting bugs, improving translations and coding for new features. On this point, it might be useful for future developers to add the conversational interface to this application to make it more flexible and easier to use.
- There is no cost for basic functions which are easy for everyone to download.

- All map data can be stored on the user's device memory card for offline use.

The application can be launched on both Android and iOS platforms and works with OpenStreetMap. This provides the user the ability to edit and future developers the abilities to access and improve the application.

1.1.2 The conversational interfaces

Three distinct kinds of conversational interfaces have been created by the researcher during this research which allowed the user to obtain information about the sites and navigate around Massey University. This was achieved simply by pressing the blue microphone and talking to the system through their smartphone (Figure 2). Included were the conversational-only interface (Group B), the conversational with map interface (Group C), and the conversational with image interface (Group D). The user's current location is displayed in the highlighted blue box, and the user is able to ask for more detailed directions, and a number of other clarifications (for example, 'give me directions', 'I am here', 'then what?', 'tell me more', etc.). This is a basic prototype of the conversational interface. For this research, we selected a pre-defined route within Massey University to test the performance, usability and usefulness between visual and non-visual navigation systems. All conversation is based on data collected from human conversations.

Figure 2 The conversational interface

1.1.3 Advantages and disadvantages of using navigation systems

Often technology has positive and negative aspects to it, and navigation technology is no exception. As discussed above, this is a new technology which has brought many new functions to communication technology and various other spheres of our lives. Navigation systems have many advantages with useful tools for guiding directions that can be used for many purposes including walking, hiking, driving and even finding things. For example, a new student may use the technology to find their classroom on the first day. This can assist in preventing them from getting lost and saving the time taken to get to the classroom. As reported in Smith, (2015) 88 percent of people who live in America use smartphones as their map or navigation devices. There are various ways the user can utilize their phone as a GPS device - either through its built-in map features or via other downloadable applications. Another example is if a car is stolen and it has GPS technology, this could have assisted in locating the vehicle and returning it to the rightful owner. Despite this technology providing many benefits to users, there are a few drawbacks that have been identified. An example of this is when the GPS system moves out of range and is unable to be used as a locational device.

There are a number of existing navigation systems and these systems can be very useful. Some researches have revealed some advantages and disadvantages of these systems. For instance, Rose India (n.d.) and other researchers have reviewed the advantages and disadvantages of using navigation systems and these are detailed below.

Advantages:

As reported in Rose India (n.d.), navigation systems are easier to use and more readily available via the user smartphone compared to the traditional paper map and compass. It can also be used by all groups of people such as students, the elderly, and people with disabilities. It helps users save travelling time, especially when on unfamiliar routes as it has turn by turn instructions making it easy for the user to follow the instructions and arrive at their destination (Rose India, n.d.). Navigation systems also work well in any kind of weather and anywhere around the world where there is mobile coverage (Rose India, n.d.). Using navigation systems are extremely cost competitive when compared to other available tools (Rose India, n.d.). There are many mobile navigation applications for IOS, Android, and Windows which the user can download for free and some of them that require a small payment for extra features. For example, searching points of interest (POI) such as hotels, restaurants, petrol station, shopping malls, and even traffic reports are provided in some navigation systems. Tracking features are available through some systems which they can monitor the user safety by providing speed alerts when driving, or low fuel warnings. (Vonk et al., 2007; Minami, & Madnick, 2010).

Disadvantages:

There are also some disadvantages behind this technology, battery life can be an issue depending upon the length of time required for each use. (Rose India, n.d.). For instance, mobile devices can work for a maximum of 10 hours (depending upon the applications being used on the devices, and may require the user to carry an external power source. Nowadays, there are so many new technologies that have been released to meet users needs, that it sometimes becomes difficult to learn and use the new technology. For example, many new vehicles provide built-in navigation systems which the driver can search for the destinations from a small monitor, however they are required to input information into the system. This can sometimes be problematic if the user does not know the exact address (D'Este et al., 1999). A further drawback of this technology is keeping the map data up to date. Roads and other navigational inputs are constantly changing and require regular updating of the background system to ensure it remains valid.

The user derives many benefits from using navigation systems. For instance, they help to reduce distances, times and can alert to low fuel while travelling around new places. However, this technology also has some limitations, which were described above, that need to be addressed or improved in future to make navigation systems use more widespread. The factors for improving navigation system efficiency is important and developers need to understand what these improvements are.

1.2 Aim and Objectives

The purpose of this research was to compare the performance between the OsmAnd application and three types of conversational interfaces and to seek out whether the conversational interface would be easier to use and more useful for the user when navigating. The process includes defining a way to select an existing visual map and designing a new conversational interface. The objectives are outlined below:

1) Select a suitable existing visual navigation system and design a prototype of conversational interface for comparison.

2.) Investigate the best kind of navigation system for pedestrians through corresponding surveys by testing four different interfaces (e.g. the map with command interface (OsmAnd), the conversational-only interface, the conversational with map interface and the conversational with image interface).

1.3 Research questions

The following questions were asked when undertaking the research:

1.) What kind of navigation interface is the best for pedestrians?
2.) Does a conversational interface allow users to navigate more effectively and efficiently than a more conventional map interface?
3.) Does a conversational interface provide better usability than a more conventional map interface?
4.) How should a conversational navigation interface be designed to optimise the user experience?

This research is organized as follows:

- In the subsequent sessions, we review the literature.
- Next, we present our research methodology.
- Then we discuss our results and analyse the results.
- Finally, we conclude our research.

Chapter 2

Literature review

This section reviews the literature related to this research. We first consider current navigation systems for driving and walking, and then we step into the topic of improvement of navigation systems. The gap in current systems is discussed. Following this, speech recognition technology for navigation and other fields, the conversational interfaces and analysis are reviewed.

2.1 Existing navigation systems for vehicles

Vehicle navigation systems on either mobile devices or dedicated GPS devices (‘satnavs’) are one of the most successful types of systems available (Park et al., 2013; Skog, & Handel, 2009; Park, & Kim, 2014). Lin, and Da Young (2015) demonstrated the percentage of use of in-car navigation systems from Qualcomm was 18% at least twice a week It was found the user usually turns the navigation system on while they are travelling. There are many existing in-car navigation systems that come with useful functionalities. These include searching points of interest, helping the driver to find common places to go and popular activities to do.

AR-Navi is a new in-car navigation system developed by Akaho et al., (2012). Their system is based on augmented reality technology (AI) that enhances the quality of user-friendliness and safety. Some previous studies also compared the performance of displays to improve the quality for in-car navigation systems. For example, Medenica et al., (2010) examined two displays between AI and street view navigation, both with personal navigation devices, to find out which one had the best performance in presenting navigation instructions. The result indicated that AI with personal navigation devices is the most suitable display in vehicles. It was shown through results that Medenica et al., (2011) street view with personal navigation devices (SV-PNDs) using head-down displays (HDDs) to present images of the world, is worse when driving compared to augmented reality. Using AI with a navigation system can help the user feel more secure while travelling to unfamiliar places as the user can see pictures of the directions. Like our research’s experiment, one group of participants used the conversational with image interface (Group D). We put photos of the main landmarks around Massey University which made it easy for them to recognize. Most participants in this group were visitors and had never been on the campus and the other three groups which used different kinds of navigation systems. All commented that they would like to see this navigation system exist in the future.

The University maps could become one of useful navigation systems particularly for freshmen and visitors. Due to in the complex nature of the University, many students and visitors lose their way when moving around the campus. One of the participants is a Massey University Librarian, who commented that “there are many students that come to the counter to ask for directions to their classrooms. She estimated this was more than 10 people daily, and felt it would be greatly beneficial if there was a Massey map provided for them rather than only paper maps”. As shown in Chou et al., (2006) a campus navigation and parking assistant system (CaNPAs) is one navigation system for students which provides car parking information and directions within a university. This system contains two parts: First is the Parking Assistant System (Pas) for providing information regarding available parking in the driver’s areas and suggestions to the driver of suitable parking spaces. The parking assistant could be another helpful system for avoiding traffic within the university in peak times. This feature could potentially be included into the pedestrian navigation system in the future. The second part is a campus navigation (CaN), designed for students who need to walk to unfamiliar places within the campus. Their results revealed that the campus navigation system is one of the most useful tools for new students who are unfamiliar with buildings and places in the University and require assistance to guide them to destinations easily and with clear instructions (Chou et al., 2006). This research has created a navigation for everyone who comes to take university’s services. However, the system can only communicate with the user via a special device for which students are required to register for by providing their ID card. This step is inconvenient for the user as it is not able to be used as freely as with mobile applications on their smartphones. Adding to the difficulty is the fact that this system is an offline navigation system which does not rely on GPS and therefore does not allow updating of information. This could pose problems in the future when some areas around the university change, such as buildings, classrooms and roads.

There is some commentary regarding large and small vehicles on the University roads and a number of serious incidences and damaged roads that occur because of large vehicles (Du, & Aultman-Hall, 2007). Some prior studies highlighted this issue and evaluated the navigation system especially for heavy-cars to help drivers drive safer, alternative routes and reduce travel distances (Du et al., 2007; Kawamura, 2000). For instance, Arentze et al., (2012) investigated a new navigation system for trucks by using GPS data and diary data recorded by 100 truckers. They evaluated the design of a system based on user-friendly and safe methods, which were easy to use. This was designed to help the group of users who drive large vehicles to avoid uncomfortable conditions while driving. For example, help to avoid small streets or traffic congestions. Their study indicated that a new navigation system not only helps to avoid bad conditions on the road but also helps reduce travel times and distances. The traffic report function is a new technology in vehicle navigations that enhance the system's performance and is useful for the user to know the traffic while travelling. A study done by Dembczynski et al., (2012) presented "the NaviExpert's Community Traffic technology" (p.867) to report traffic data and recommend the best route due to traffic delays and traffic jams. It can even present unexpected roadworks, road accidents or diversions. It is a helpful functionality which assists the user in reducing their travelling time (Elizabeth, 2013). A few navigation systems use this technology such as Waze, Google Maps, Here Maps, TomTom UK & Ireland. These systems can be downloaded free on Android and IOS operating systems (Britta, 2015). Another study by Dong, (2011) mentioned "Intelligent Transportation System (ITS)" (Dong, 2011, p.1), a new in-car technology navigation systems that report traffic for drivers. This is a similar concept to "NaviExpert's Community Traffic technology" (Dembczynski et al., 2012, p.867). These two systems are similar as they both report on road user information via the system. Advantages of these technologies are that they can help reduce travel times and distances for users in unfamiliar locations (Adler, 2001; Ding et al., 2010). However, ITS is still new and not commonly used in small navigation systems mainly due to the cost. They remain valid in higher-priced vehicles. The Ph.D. dissertation by Kaparias, (2008) argued that recently many things and situations have developed and navigation systems have now become more important in routine life for the new generation and that this technology makes people's lives easier and more convenient on the road.

As shown above, previous research has focused on in-car navigation. Most were command systems which did not allow users to ask questions or interact with the systems, despite some of them having developed speech recognition. For instance, SIRI on iOS and Google Voice. Many researchers have tried to find ways to make in-car navigation systems more efficient, more flexible and to better meet users' needs by deploying and creating useful new technologies in their systems and identifying how these systems could be made more usable, and user friendly. This research aims to present the idea of a navigation system that can speak to the user, responding to questions.

2.2 Existing navigation systems for pedestrians

More recently, systems that support pedestrian navigation have been developed, but it is still rare to see walking navigation itself compared to other kinds of navigation. However, many navigation systems were developed for two modes, driving and walking; users can select the mode and then the directions will be changed to reflect that selection. In reality, people often travel by walking. Thus, some developers have begun to think more about implementing stand-alone navigation systems especially for people walking on the street or in areas that cars cannot enter. For example, shopping areas, indoor areas and university areas.

Many researchers have addressed different issues. Some of them focused on the designing of a navigation aid for disability (Begampure et al., 2016; Dodson et al., 1999; Helal et al., 2001; Strothotte et al., 1995), managing details of turn by turn on mobile devices to make systems reliable and accurate for route guidance (Tarkiainen et al., 2001). The research by May et al., (2003). claimed that the user makes their decision of wayfinding based on a previously acquired spatial understanding; called a ‘‘cognitive map’’. “Wayfinding is defined as building and maintaining a cognitive map and is used to determine how to get from one location to another” (Peck et al., 2012, p.1053). Understanding the information requirements and the nature of the navigation task of pedestrians is very necessary and important for developers, including consideration of factors such as landmarks and street names to understand what pedestrians need and how it is used (May et al., 2003). This was similar to our conversational navigations where we considered the details of landmarks as the first factor to provide for students and visitors who are unfamiliar with the University campus. We have designed clear University landmarks that are easily recognized. For example, Group B used the conversational-only interface, where the user sees only the name of landmarks on the screen. We added further details of the landmarks as much as we could such as stairs, shops, and cafes. The result was that most of the participants understood and recognized places faster and easier. May et al. (2003) identified requirements for implementing the navigation system for walking that all developers need to know. There are several requirements were mentioned. For example, types of turn by turn instructions, as the explanations of routes are different because there are tons of locations and so many details about it. Another one is identifying how to design the information to be suitable for mobile applications. Thus, Landmarks are essential information that needs to be clear in every single detail (May et. al., 2003), regardless of the size of the areas. The user could get lost easily if the landmark details are unclear. For example, in Michon and Denis (2001), the city is the best area for pedestrians to learn, all participants were not familiar with this environment. There are three types of information used to create the directions within the city area which matches our conversational interfaces. Firstly, street or road information. Secondly the position of a landmark such as "in front of….", "next to…", "walk toward to the third stairs", "continue walking about 100 meters, you will see another stair". And lastly, all other objects which can be found along the way while walking such as trees, bridges, and public benches. Walkers require more detail of the areas and the names of shops or small streets are crucial to consider when providing directions. It is different from driving, where names are visible and the name of landmarks should be clear and specific (Begampure et al., 2016). For example, we added a number of detailed directions into our applications, including small cafes, small shops and stairs which we considered important details for describing landmarks. Normally, when walkers search for the place while walking, they read the names of small areas that they can see or are near them instead of the names of main areas.

Another function is a sensor to help users be aware of where the location is or if they are on the right track. A vibration or sound that alerts them to their destination is one of the more popular recommendations from participants. Previous studies have also adopted sensing tools into new navigation systems to help walkers and the blind. For example, Cavallo, Sabatini, and Genovese, (2005) implemented a new walking navigation system with a foot sensor to signal the user while walking by using "a Digital Signal Processor (DSP) -based control box" (p.1) to control data from a "Global Positioning System (GPS) and the Inertial Measurement Unit (IMU)" (p.1). This technology has been deployed in many systems as it helps to people who use buses, trains or any other kinds of public transports determine the arrival time for disembarking (Jacob et al., 2011).

Once again, most navigation systems designed for walking are mainly command-based systems which means the system gives instructions and the user follows these instructions to reach their destination. System commands can sometimes be incorrect. and when the user gets lost or misses out instructions, they have no one to confirm, clarify, and ask questions to get back on track. Most navigation systems do not support this function. This is the gap that our research will try to fill to improve the usability of pedestrian navigation systems for all kinds of users. The SIRI function in IOS is an example of a conversational interface that has the capability to interact with the user. Our research differs because SIRI can only help the user find out information. For instance, the user says "I want to go to ….." and SIRI can find information and link itself to the iPhone map. It cannot talk or give instructions directly. At this point, the user still needs to rely on the map; whereas our research tries to find a way to allow the user to have a conversation with the system instead of merely seeing the information on the screen. We also aim for a human-like conversational experience, so that users feel like they are talking to a person while using the system, making them feel more comfortable.

2.3 Elements for improving navigation systems

The factors for improving navigation systems are necessary to discuss as it is important developers are aware of these important factors before creating the systems. Our conversational navigation systems were also based on these factors. Previous research has revealed ways to improve navigation system capabilities and efficiencies for the user. Park et al. (2015) examined four factors for improving car navigation systems which should be considered in developmental stages: “locational accuracy, system reliability, service & display quality, and satisfaction” (p.742). These factors are essential.

Firstly, car navigation systems should have accurate and up-to-date information of locations such as numbers, street names and routes as these aspects are often changing. The purpose of using navigation systems is to guide people to unfamiliar places and people need to be able to entirely rely on for system for instructions of timing and routing (Kawasaki et al., 2001; Chou, et al., 2006). Chou et al., (2006) presented that when new students arrive at university they often found it difficult to move around the campus because of the complex nature of the area. Navigation systems should provide the user with accurate directions that will take them to the place where they want to go. Like our conversational navigation systems, the user has to communicate with the system. Adding the names of buildings in its varied descriptive forms was necessary as many of the buildings have more than one name. As an example, “Quadrangle Building A” is sometimes called “Quad A”, or “QA”. These common names were added into the conversations so as to be flexible for all users. Buildings names often change, so updates were required, during the experiment.

Secondly, systems should be reliable and flexible, using techniques and services (Park et al., 2015) which can help the users achieve their purpose of using navigation systems. Some driving and walking navigation systems have both online and offline modes which is convenient for drivers who are unable to connect with a network or signal. The system will still be able to guide them to their destination. This function can be used on Google maps where drivers can download specific areas. Once the system has the driver data, it will go into offline mode automatically and when the connectivity improves, it will change back to online mode (Ian, 2015). In our experiment, we also focused on the user's flexibility. We have four kinds of conversational navigation interfaces, including the map only, the conversational only, the conversational with map, and the conversational with photos to enable the user to test and determine which interface has the best potential for future conversational navigation systems.

Thirdly, the quality of service - the overall performance of a telephone or computer network (Cardoso, Sheth, & Miller, 2003) that influences performance of the use of navigation systems as along with users' prospections. For example, Lin and Hsieh. (2006) reported the quality of display affects the user's attitudes about using the systems as much as the service. This is because navigation devices for drivers and walkers have both a service factor - such as a navigation system which is an actual GPS - and an operating system (OS) which is an application launched on mobiles. The user will choose the service that they feel the most comfortable with and which displays the most benefit to them. The display factor is mainly related to user interfaces where the user is considered when the system is designed (Park et al., 2015). In today's technology environment, everyone has a mobile phone and it has quickly become part of our routine lives. We use mobile phones for everything from calling, reading books or news, searching for jobs or information to navigating from place to place. As reported by Smith (2015), a high percentage of all ages of people use mobile devices use for navigating purposes. This is especially evident in young adults. There are many options for smartphone platforms on both IOS and Android. Many navigation applications can be installed for little or no cost and time, and many of the systems perform differently. First are visual modalities include maps, texts, and images. Second are non-visual modalities such as voice and tactile. Previous research by Rehrl et al., (2012) compared the performance of three distinct types of pedestrian navigations, including augmented reality, speech and maps by providing equal route information for walkers. They revealed maps and conversational-only interfaces had significantly better navigation performance compared to AR interfaces. The user feels easier viewing and more comfortable using maps and the conversational-only interfaces.

Lastly, research done by Park et al., (2012) and Park et al., (2013) evaluated a system by using usability testing to gain the users' satisfaction results. It is one of the best tools for understanding the users experience of websites or applications (Battleson, Booth, & Weintrop, 2001). The researchers take notes, and record all participant's movements from the beginning to the end of the experiment in order to evaluate the results for future development. The aim of this testing is to investigate any usability issues, collect qualitative and quantitative data and gain the participant's feedback (Battleson et al. 2001).

Overall, developers should understand basic factors to improve their system's accuracy, reliability, functionality and flexibility. Navigation systems could become the best assistant navigation for users in the future.

2.4 Speech recognition technology

Speech recognition technology is being adopted in many different systems. Speech recognition systems work on the human voice to understand and carry out spoken commands (Brett, 2008; Margaret, 2007). It is used for operating devices, writing without a hand (Jim, n.d.). The use of voice recognition systems gives the user a feeling of realism as they can communicate with the system as if talking with a real human being. However, using voices as an input to a computer can be complicated as the human voice and the traditional form of computer are totally different (Turk & Robertson, 2000). In particular, all human voices are different and it can be complexed for the system to recognize meaning and tone.

One of computer systems called Automatic Speech Recognition (ASR) allows the user to train their speech so that the system can recognize it and be accurate when the system converts the user's speech to text (Margaret, 2007). For instance, our research used API.AI to build conversations between the user and our interfaces. API.AI has ASR as part of system in order to help to train the user's expressions. It is easier to recognize the user's speech compared to other tools. Many participants gave positive feedback about API.AI being able to recognize their voices so well, even though their English was not so good.

Some ASR systems display a word or a sentence after the user speaks the words or sentences into a microphone (Jim, n.d.). The system will then process words into vocabulary data. This type of system is called "speaker-dependent" (Figure 3) voice recognition which works on the order of vocabularies being about 98 percent successfully recognized (Yannakoudakis, & Hutton, 1987; Dharun, & Karnan, 2012). Another form of voice recognition is called "speaker-independent" (Figure 3) which processes the spoken input using Linear Predictive Coding (LPC) (Bradbury, 2000; Yannakoudakis, & Hutton, 1987; Dharun, & Karnan, 2012). Here the system will find similar patterns between the expected inputs and the user's spoken input (Yannakoudakis, & Hutton, 1987).

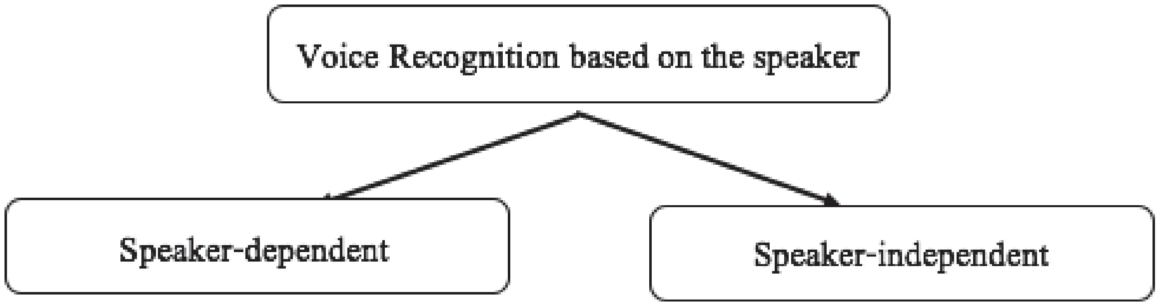


Figure 3 Voice Recognition based on the speaker

2.5 Existing speech interfaces in other fields

Speech interfaces are one way in which the spoken word can be used as an input to a computer system (Yannakoudakis, & Hutton, 1987). For the past few years, voice recognition has been improved and become an active area in researche (Dharun, & Karnan, 2012). Systems have tended to be developed by adopting voice recognition to help some specific people such as children, the elderly, disabled people and foreign language learners. Voice recognition and speech-to-text are widely used in many systems in order to help the user to improve their writing ability and can be an alternative way to access a computer or other devices (Bruce et al., 2003). Dellea, (2011) implemented an assistive technology clinic by using voice-recognition to help people who have difficulty using a keyboard and mouse for a variety of reasons. For example, for someone with no computer experience, it is hard to learn computer tasks and software, especially the elderly (Colman, 2016). Modern technology is difficult for some elderly people to learn, so speech recognition can be one of the most helpful technologies and it is convenient for them to access a computer or other devices. For example, Asano, (2008), a smart home navigation system, is a useful technology particularly for the elderly and disabled people as the system is mainly used by the user's speech. The system consists of a wheelchair, navigation module and voice module. It has pre-defined speech commands relating to different rooms in the house, and pre-defined routes relating to navigating around rooms. Thus, the user can use this technology anywhere they are without movement or touch, but simply by voice.

Speech recognition technology can be used for enhancing students learning ability as well as navigation. Some researchers reported that this technology can help with improving reading and writing skill for struggling students by seeing vocabularies on the screen, helping students remember. The system includes related sound and words as the student speaks out loud and see words at the same time (Silver-Pacuilla, 2006). Another benefit of using speech recognition technology in the education area is to help student correct and check words and sentences by comparing sounds and structures that look alike. It will also help them to select the most suitable word for the appropriate situation (Higgins, & Raskind, 2004). Hence, this technology has the potential to take a major role in the education area and improve student's performance. It is an active way for them to study by using only their speech and touch (Brainline.org, n.d.) Speech recognition proves that navigation technology has multiple uses and benefits for students of all types.

2.6 Existing speech interfaces with navigation systems

Some navigation systems deploy voice recognition to help the user avoid distraction while driving or walking, and allows the user to concentrate and focus on their direction and traffic conditions. Speech recognition in the car is one of the applications that will be of high interest in the future (Pouteau, & Arévalo, 1998). As researched by Wang Mi Guo Bingxuan (2002), Via Voice is a Chinese mobile navigation system which adopts speech recognition and a dictation system to help the user while driving on the road. This system can transfer user speech into text without using their hands. It has the features of accent adaption to recognize different users' speech. The user can input phrases to Via Voice where vocabularies are used to recognize the spoken information. It is slightly different from our research in that the system can only recognize the user's speech and display results on the screen, whereas our research has the system and the user communicating with each other. There are similarities to the API.AI tool we used in our conversational navigation systems which can perceive most accents during the experiment, even though some of participants were not native English speakers. The system could determine their queries by using additional short sentences such as "what else?", "what is next?" and "then what?" making the system easy and flexible to use. The system felt more user friendly to the student.

Another system, mentioned earlier is a university map and parking system for "National Tsing-Hua University (NTHU) in Taiwan" (Chou et al., 2006, p.631). It is a two-in-one system which primarily uses voice instructions for guidance. It also can check the users position and give notification feedback when the user is off track. This system can be used only in National

Tsing-Hua University with a provided portable device, making it less convenient for usability. The downsides of these systems are:

- They are only implemented for a dedicated device. The user is required to request a device from the University which is inconvenient.
- They use voice recognition technology to help guide students and provide directions and recommendations for car parking. This system could have been improved if the user was able to respond back to the system. This requires the implementation of two-way interaction instead of one-way commands. This would provide options for people to ask for confirmation, to give feedback and to ask further questions.

Therefore, a better way to help the user to use navigate directions would be using speech recognition technology where the system can interact with the user as well as implement applications for a smartphone. It is convenient and easy for the user rather than using actual devices that can be complicated to use.

2.7 The conversational/dialog interfaces

Conversational interfaces are a technology system that allows the user to communicate with the device using their natural language, in order to make them feel like they are communicating with a human (Ortiz, 2014). The conversational interface can help the user gain information, access to the internet and other purposes, without typing. The use of a conversational interface is easier than other computer technologies and the user could spend less time learning how to use it. For example, all participants were asked to read the instructions and play around our systems for about 2-3 minutes before the experiment. During the test, many participants, especially in the age group of 40-50 were able to reach to their destination faster than any other age group. In general, at this age, learning a new thing related to new technologies can take more time compared to the younger age group. Communication between humans and smart devices has become recently possible (McTear et al., 2016). For example, Apple's SIRI, Google Now, and Samsung's S Voice helps the user perform tasks on their smartphones and obtain information by using voice search, find points of interest, find contacts and setting alarms. The conversational interface is also being used for education purpose, in order to develop student's learning ability. Pérez-Marín & Pascual-Nieto, (2013) investigated the prototype called "Pedagogic Conversational Agents (PCA)" (p.955) to develop student's learning skills. The computer system has features that motivate them to study because they felt like it is their friend and has a human feel. The challenging thing for this research was about the design of the system especially when being used by children which required more detail than with adult use, the context needs to be stricter and more suitable for children.

Being computer interactive is interesting and fun because users do not have to type into the system, only input their words and the system understands what they are saying. However, for designers and developers, creating this technology can be a challenge for them. Some researchers found that, developing a conversational interface requires dealing with several problems. Zue & Glass, (2000) revealed challenges of developing a conversational interface. It must deal with the transition from speech dictation and spoken document retrieval. Although the system is often trained in different ways this is unavoidable. Thus, the system must handle many situations especially background noise (e.g. wind or cars). Another problem is the diverse speaker population (Dekens et al., 2010) People have their own accent and way of speaking. A major problem for the conversational interface is recognising storing new words. Because it is impossible for the system to anticipate all the words that all users are likely to use, and the data base is constantly changing over time (Zue & Glass, 2000) the solution for this problem is to use keyword or phrase spotting in the system output to gain the meaning of sentences for a conversational interface to understand in context. This is a similar issue to our experiment as understanding all words from users can be difficult for the system. Our system has been designed to be as flexible as possible in this regard. We observed conversations between two people and added that data into the natural language platform (API.AI). The system can recognize the user’s utterances by keywords however but sometimes has difficulty with user speech. Another issue during the experiment was with background noise (wind) which was the biggest problem to affect the system’s performance.

The conversational interface should be designed based on real-world conversations, and the needs of users and their behaviours. As well as our research, we explored all kinds of conversations of users finding directions in our previous work (Longprasert, 2015). Because we would like the user to feel comfortable and enjoy using the system our design has been made as effortless as possible and suitable for all types of users.

2.8 The nature of conversation and conversational analysis

Conversation is interactive and occurs between two or more people to exchange ideas, share information, and express feelings. According to Hutchby and Wooffitt (2008) Conversational analysis (CA) is the process which describes and understands a human's natural communication in situations of everyday life.

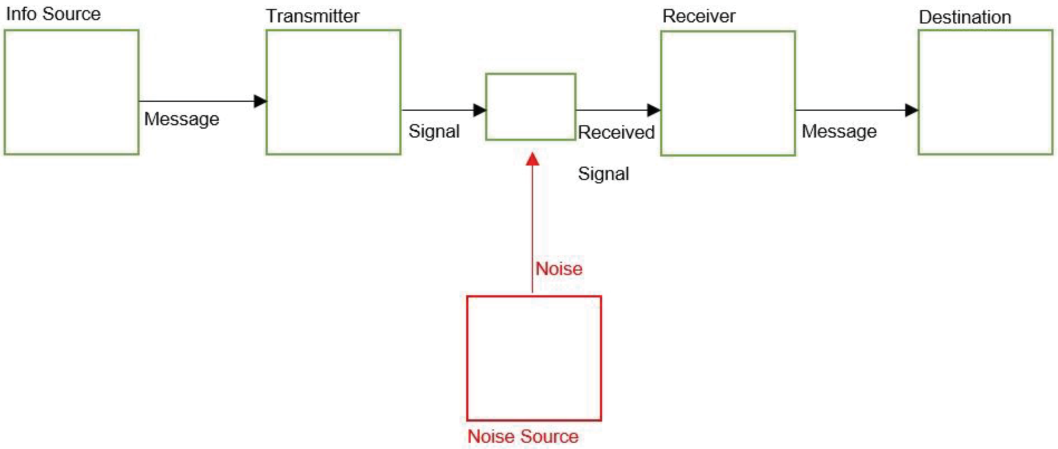


Figure 4 The conversational analysis process

(Shannon's Model)
Reprinted from (Dubberly and Pangaro, 2009, p.1)

Figure 4 shows the conversational analysis process based on Shannon's model which is from an information source (Person A) who creates messages, and has been sent through a transmitter as a means to capture a message from Person A and changes to an electronic signal, passing through a noise box which is an obstacle to the carried signal before the message is delivered to a receiver (Person B) and arrives to its destination (Dubberly and Pangaro, 2009).

There are three types of conversations which we used to analyse our research. According to Carletta et al., (1996) analysed the move coding Scheme to fit the kind of interactions in the map dialogues. There are three main move categories. First is the "**Initiating move"** (p.22), which occurs at the beginning of the game. We also identified this as the first conversation where the user starts to speak to the system. Second is the **"Response move"** (p.19) which is used after the system started and use to fulfill following conversations by the user. Lastly is the **"Ready move"** (p.22) which is used after the end of the conversation such as when the user has arrived at the place or the destination.

In our previous research, we identified all collected conversations into eight types of basis utterances (Longprasert, 2015) based on common conversations between two people when they talk about the directions which we were observed and collected from our previous participants. (1) “**Greeting”** (p.21) which is the first greeting in a conversation (e.g. “Hello”, “Hi”, and “How are you?”). (2) “**Request”** (p.22) is the conversation that Person A asks for help from another person such as “I need your help now”. (3) “**The Opening conversation”** (p.22) is the beginning of a conversation (e.g. “What can I help you with?”, “Do you know…?”). (4) “**The Identification conversation”** (p.22) is to identify people and ensure that they understand (e.g. “Are you at the bus stop?”, “Can you hear me?”). (5) “**The Confirmation conversation”** (p.23) is used when Person A and Person B confirm their answer to each other (e.g. “Yes, I am here”). (6) “**The Clarification conversation** is to clarify answers or explanations within a conversation” (p.23) (e.g. “Yes, then keep walking for five steps”). (7) “**The Feedback conversation”** (p.24) is reactions of the user after seeing things. (e.g. “Ok, I found ATM at this building”). And (8) “**The Closing conversation”** (p.25) used when the conversation is finishing. (e.g. “Thank you”, “See you later”).

2.9 Summary

It is obvious that most current navigation systems used for either driving or walking are mainly command-based systems which cannot be interactive. It might be difficult for the user which results in them getting lost or misunderstanding directions. For example, when people give directions they often do this through a conversational exchange in which the person being directed can seek confirmation, clarification and ask questions. However, most navigation systems do not support this functionality. Even though some studies adopted speech commands in their navigations, they were only one-way. For instance, Bingxuan, (2002) developed a navigation system for cars which can transfer user dictation into text instead of typing but still cannot be used interactively. As well as speech recognition in Apple’s SIRI where the user can ask anything, questions from the user are linked to the application within devices. It would be more beneficial if navigation systems could be interactive. This research presents the designing of a prototype two-way conversation for navigation where the user is able to communicate with the system via four different interfaces and then compare them by using the usability method along with interviews to obtain results.

Chapter 3

Methodology

This section explains the research methodology used. The research was conducted by combining qualitative and quantitative approaches. Both were used because different types of data were gathered and were analyzed using different methods to obtain the most useful results and fulfilling research questions that could be important for future development. For example, in the actual experiment we presented images of directions and text data which were not suitable for quantitative as well as interview data analysis. In contrast, we used quantitative approaches to indicate statistical results from closed-ended questions in the questionnaire which were clearer than using qualitative. As reported in Carvalho & White, (1997) using both quantitative and qualitative methods, we could get better results for products than using only one of them. We considered here are limitations in using a singular method. The methodology contained three parts:

3.1 Usability Study Methodology

Usability is one of the evaluating methods which helps developers improve their products, websites and systems. This method can get feedback of the use of products from the user in order to design the products to meet user's needs and become more effective (Dumas & Redish, 1999). The experiment of this research was conducted using qualitative and quantitative data to fill the research questions.

3.1.1 Pilot Study

To make sure all kinds of navigation interfaces worked well before the experiment and to evaluate the data collection instruments, the researcher ran an initial test to evaluate feasibility, duration and adverse events for the usability study and to estimate the sample size and confirm the design of the final experiment. The purpose was to avoid time and to identify any issues in the final experiment (Goodman et al., 2003). The six volunteers have different experience level of using navigation technology in the pilot study and they were recruited through personal contacts. They were four Asians and two Westerners. All volunteers could speak English fluently and some of them often used navigation systems.

3.1.2 Participants

100 people were selected to take part in the experiment. We distributed advertisements to the mailing list, notice boards and social networks (e.g. Facebook) within Massey University. All participants contacted the researcher to take part in the survey via email and text. After that, the researcher arranged a time to meet with them. They were divided into four groups with random experience of the use of navigation systems (Table 1). All were students or staff at Massey University and some were visitors to the country and spoke both English and a second language. For the interview, there were 70 participants from Groups B, C, and D.

Groups C and D had only 20 participants each because early expectations were that we would find only 90 participants and therefore set only three groups (Groups A, B, and C). After the experiment started, the researcher came up with a new idea of making the conversational interface in a different style This idea came from suggestions of earlier participants in Group B, which noted real images of landmarks might make location navigation easier. Eventually, Group D which was the conversational interface with image, was established and was limited to 100 participants in total due to budget. Finally, we reduced the participants from Groups C and D to 20 each making 100 in total.

Table 1 Four groups of participants

Group	The navigation interface	Participants
A	The map with command interface	30
B	The conversational-only	30
C	The conversational with map	20
D	The conversational with image	20

3.1.3 Experimental Design

The design of the experiment tested the effectiveness, efficiency and learnability of four kinds of navigation interfaces, including map with command interface, conversational only, conversational with map, and conversational with image. Group A includes audible commands but the user cannot communicate with the system while Group B, C, D can. The questionnaires were constructed as five Likert rating scales. “Participants were asked to rate their views with the statements, from “Strongly Disagree” to “Strongly Agree” (Lund, 2001, p.2) for all kinds of navigation interface. The four main sections of the questionnaire included:

- completing the given tasks to test four kinds of navigation interface;

- filling in the USE questionnaire form by Lund, (2001);

- presenting and comparing three sections: "Usefulness, Satisfaction, and Ease of Use" (p.1), between the existing navigation and the conversational interface;

- answering some questions about improvements of all kinds of navigation interface;

- interviewing participants' attitudes to obtain in-depth detail with Groups B, C, and D.

All participants were asked to:

1. Complete the navigation interface testing:

    I. Learn how to use the interface by followed the given instructions.

    II. Follow the system's instructions to the destination (Figure 5).

2. Complete some open-ended questions about their experience using the navigation interface.

3. Fill in the USE Questionnaire.

4. Discuss in the semi-structured interview with another three questions about future improvements and the in-depth experiences of using the navigation interface (except Group A). Participants' movements, voices, and times were recorded from the beginning to the end of the survey.

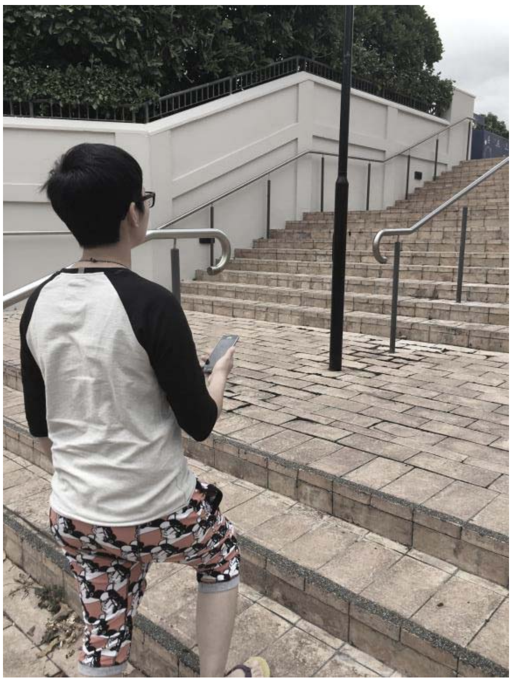


Figure 5 The participant is testing the interface

The survey took place on Massey University (Albany campus), Auckland, New Zealand and started from The Chicken Wing to The Student Accommodation. For the whole experiment, all participants did not know the destination because it had been set by only the researcher, the user can only ask for and follow the direction from the system without any idea of the end location. In order to find out interactions between the system and the user. Participants were required to complete the consent form before undertaking the test to ensure willingness to take part. The consent form contained three sections: General information (age, gender, nationality and position), mobile devices (the type of participant's phone and the operation), and the use of navigation systems (experience and participants' ideas about new features of navigation systems in the future). There were 30 participants each in Groups A and B and 20 participants each in Groups C and D.

The advertisement (Appendix J) of the experiment was sent out around the campus via email, social network (Facebook), and the news walls in Massey University, to get a variety of people in the experiment.

3.1.4 The questionnaire

The questionnaire (Appendix: E&F) for this research consists of four sections: the first section focused on the user's attitude to gain their feelings after using the system. The second and third sections focused on effectiveness and efficiency, learnability, usability and satisfaction. 26 questions in part three were based on a USE questionnaire by Lund, (2001) which has Usability appearing to consist of three dimensions: Satisfaction, Ease of Use and Usefulness. Each factor in turn drives user satisfaction and frequency of use. The last section, about the future improvement of the interface, used open-ended questions to gain the user's views and reasons for future development in the same way as the Interview.

The purpose of using the questionnaire was to evaluate users' attitudes for navigation interfaces that they have used, to show what is usable and what is not. The questionnaires were constructed as closed questions where a specific response set was provided (i.e. a five-point Likert rating scale), and users were asked to rate agreement with the statements, ranging from strongly disagree to strongly agree. There were also open questions where subjects were able to respond in any way they saw fit (e.g., what do you like most about this application?).

The first set of questions for Group A was slightly different from the other three groups because the capability of the system was different. Group A used the OsmAnd application which does not provide the conversational function.

The reason for using the USE question was that this instrument is suitable for this evaluation because the user evaluates the new product primarily using three dimensions:

Usefulness, Satisfaction, and Ease of Use. However, other dimensions were found, but those three served to most effectively discriminate between interfaces.

3.1.5 Interview

An interview was applied for participants in Groups B, C, and D - those that used the conversational interface - as this was the major evaluation of this research and enabled us to gain in-depth information by delving deeper into people's responses. Participants were able to give more feedback which was very useful when we were analyzing the results (DiCicco-Bloom & Crabtree, 2006). Two questions formed the interview section. They mainly focused on the future development of the conversational navigation which could be useful for future studies. The last question, which was not on the list, was about the feeling of using the conversational interface. After participants completed the survey, they were given a voucher card as a thank you. Group A did not include in the interview because the researcher would like to gain more feedbacks from participants who were in conversational interface groups but Group A used only map- interface from an instant application with no conversational feature. They would not have ideas what conversational navigation system should be.

In addition, the map with the command interface in this research used an existing navigation system named OsmAnd, an Android application. Four different kinds of navigation interfaces were required for comparison purposes (the only-map interface, the only-conversational interface, the map with conversational interface, and the conversational with image). The reason for testing four kinds of navigation interfaces is based on the current situation of using navigation systems where the user can only listen to the system giving directions. Some people may get lost and confused by the system's instructions but they cannot talk or ask for more information from the system. This research designed three conversational interfaces to compare with an existing navigation system to see how easy and useful it would be if the user was able to talk while they are finding locations.

3.2 User Evaluation Testing Protocol

This research used usability as a part of the method which refers to the improvement of ease of use during the design process (Marisa, 2010) including four components:

Efficiency and effectiveness: We monitored how users could complete their goals successfully.

Ease of use: The usefulness of the system.

Learnability: We compared both modalities to seek how easy it was for users to complete tasks the first time they used the system. We monitored and recorded them as they started learning the system and we also asked them about using the system in the interview section as the last question.

Satisfaction: To find out users' feelings about the system.

The reason for choosing usability is because usability is a method that could describe ease of use and gather real feedback about a product from potential users. It also could help the product to more effectively satisfy the needs and desires of users.

Below is a list of criteria based on some participation filters fixed for the selection of the suitable participants to the experiment.

Table 2 The selection criteria for user study

Filter	Criteria	Explanation
Student/staff/visitor	Participants could be studying, working in or visiting Massey University.	Participants could be a student, staff, or visitor as our research aimed to identify whether the conversational interface was useful for navigation by comparing four different kinds of navigation systems with a pre-defined route in Massey University campus.

Languages spoken	Any level of language skill but should be able to understand English. Native speakers are preferable but not compulsory.	The system has the ability to understand only English, so it would be an advantage for evaluating the system if participants are native speakers or have a good grasp of English.
Knowledge-Skills required	The perfect sentence is not needed. Experience using navigation systems is preferable but not compulsory. The reason for this criterion was the conversational interface could recognize short sentences, which could be easier for all kinds of users such as non-English speakers, and the elderly when they do not need to communicate in long sentences.	This research developed a conversational navigation system that is more flexible. Perfect sentences are not needed, so participants may say only a few words. This research categorized participants into four groups. Ex. Group A tested a visual modality named OsmAnd. Group B used a non-visual modality which is the conversational interface. Group C tested the conversational interface with the map. Group D used the conversational interface with image.

Additional document contains:
1. A list of questions (Appendix B)
2. Instructions of the OsmAnd application (Appendix C)

Table 3 The experiment's step

Section	**Approach**	**Duration**	**The use of equipment**	**Analysis Method**
The introduction	- Researcher introduced the study and the objective behind evaluating this system. - Researcher grouped participants to different tasks. - Researcher gave the instruction on how to use each navigation system. - Participants signed the consent form. - Pre-testing applications	5 mins		
Participant experimentation	- Participants are given a simple task. The first group used the map with command interface with an existing navigation system named OsmAnd launched on an Android. The second group used the conversational-only interface. The third group used the conversational interface with the map. The fourth group used the conversational interface with image. Along a pre-defined	10-15 mins	iPhone 6s, Samsung Galaxy J2	

	route in Massey University campus, the researcher recorded participants with audio and tracked the users' movements by GPS tracker.			
A questionnaire (Appendix C)	The participant who finished a task will be given a different questionnaire, in terms of the question in the questionnaire will be three and four sections.	5 mins	A Use questionnaire	
Interview (Appendix D)	The last step of the experiment was discussed to get more in-depth information. The reason for doing this is users will say much more verbally than they will in writing, and we can ask them to clarify, telling more, explaining things in more detail.	5 mins		
	Total	30 mins		

3.3 Data collection Techniques

3.3.1 The observation (all recorded data)

Observation takes two forms in this research. Firstly, the researcher is seated near the participants and observes them as they learn how to use the navigation system. The researcher is trained to focus on specific behaviors and times as they do the experiment (Kelly, 2009), taking notes that describe the participants' particular movements or spoken words. The observation was real-time and play-back time which means the researcher cannot interrupt or help the participant during the experiment, but can ask follow-up questions about particular events later during the post-search interviews (Kelly, 2009). The paths, voices, and duration of all participants were recorded as well as their interactions while using the system by different tools. For the participants' interactions, the researcher used a screen capture software named AZ Screen Recorder to record the participants while testing the system by using Samsung J7 for recoding data. The participants' paths used the OsMoDroid as the GPS tracker to monitor the participants while they were walking (Figure 6) by using iPhone 6s for recoding data. For the duration and voice recording, the researcher used a stopwatch for timing (Figure 7) and The Voice Memos for the voice recording (Figure 8) These two apps are from IOS operation (iPhone 6s).

Figure 6 OsMoDroid

Figure 7 Stopwatch

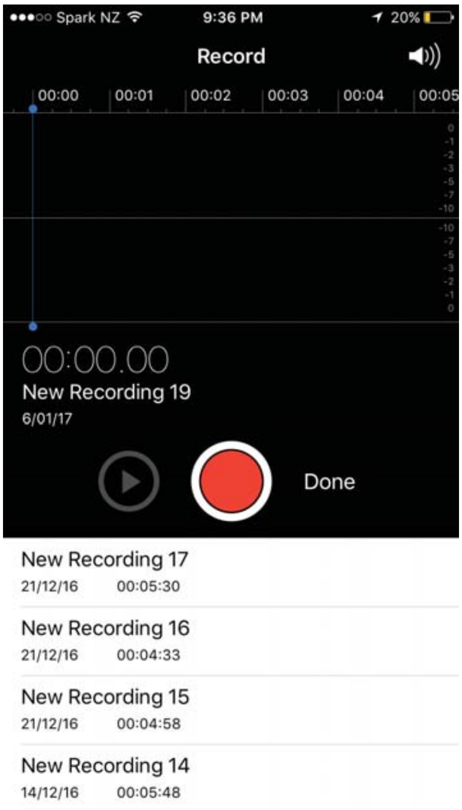


Figure 8 Voice Memos

For observation, it is crucial that the researcher has some practice before commencing observation and has some idea of what specific data needs to be recorded and how it will be used (Kelly, 2009). Training of observers is important. In this research, the researcher practiced several times before the real observers to ensure that the researcher recorded the same types of things and levels of detail without error.

Chapter 4

Prototype design

This section presents the design of the conversational interface prototype that was compared with an example of a conversational navigation system (OsmAnd). This question/answer response data for the prototype is hard-coded, and it is not intended to present that the kind of conversational interface that would be used in reality, since a dynamic way of responding to instructions is needed. The prototype was used to evaluate the potential benefits of conversational navigation interfaces and the parameters that might be considered in the design of such an interface. We describe below the software components and functionalities that were used.

4.1 The Overview of the Design

We used the Android studio to create the prototype and linked with other components such as API.AI by using ACCESS TOKEN KEYS for different locations obtained via GPS on the mobile device from the current user location. When the user wants to speak to the system, they are required to press the microphone button which triggers the system to find the closest location based on longitude and latitude that we added to each location. Then it determines the closest location and looks up the set of instructions for that location which are hard-coded in API.AI. Once the user reaches a closest location, the set of instructions of that location responds back to the user (Figure 10).

Figure 9 The conversational interface

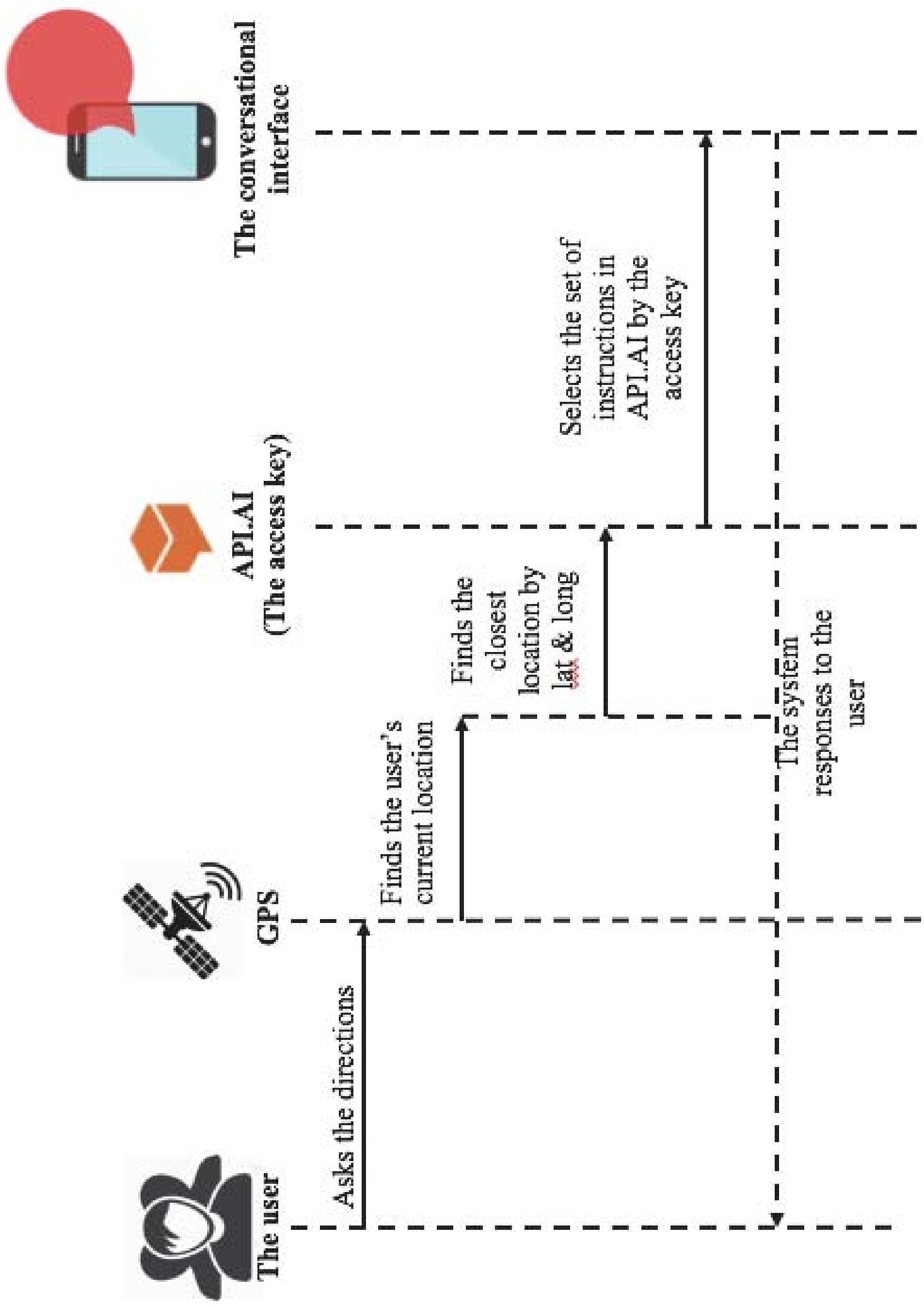


Figure 10 Overview of the Design

4.2 Software components

4.2.1 Android studio

The Android studio version 2.2 was used to design the conversational interfaces (the conversational only, the conversational with map, and the conversational with image). Android studio provides a helicopter view of the different tools that come with the Integrated Development Environment (IDE) (Van Drongelen, 2015; Zapata, 2015) created by Google. This software can support a variety of devices such as phones, phablets, tablets, wearable devices, Google Glass and TV apps. It contains all the Android SDK tools needed to design, test, and debug.

The reasons for selecting this software is that Android studio is a highly recommended IDE that has capability to develop so many kinds of Android applications and it is available free.

4.2.2 API.AI

All conversational data was created by using API.AI (Figure 11). It is a natural language platform which allows the user to create conversational user interfaces into mobile applications, web applications and devices (API.AI, 2016). API.AI contains five elements:

1) **Agents**: once we train and test an agent, it can be integrated with the app or device (API.AI, 2016). Figure 12 shows all the agents that the researcher created for this research.

Figure 11 The API.AI process

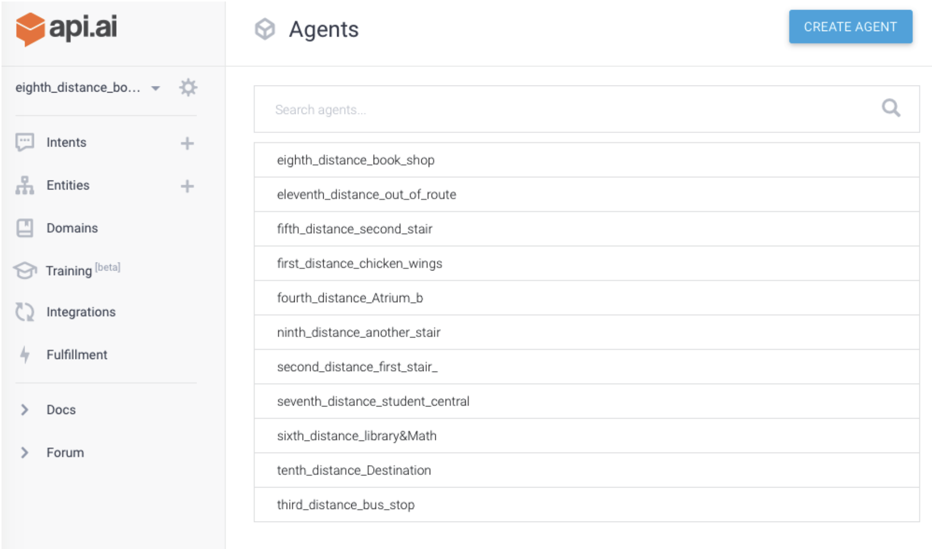


Figure 12 API.AI (Agents)

This research has 11 agents (Figure 12.) containing different intents and entities for conversations between the user and the system in each location from the starting point to the destination (Figure 13), and providing responses that are appropriate at each location.

Figure 13 The route from the starting point to destination

Each agent has only one access token key (Figure 14), which is used in Android Studio to identify the agent.

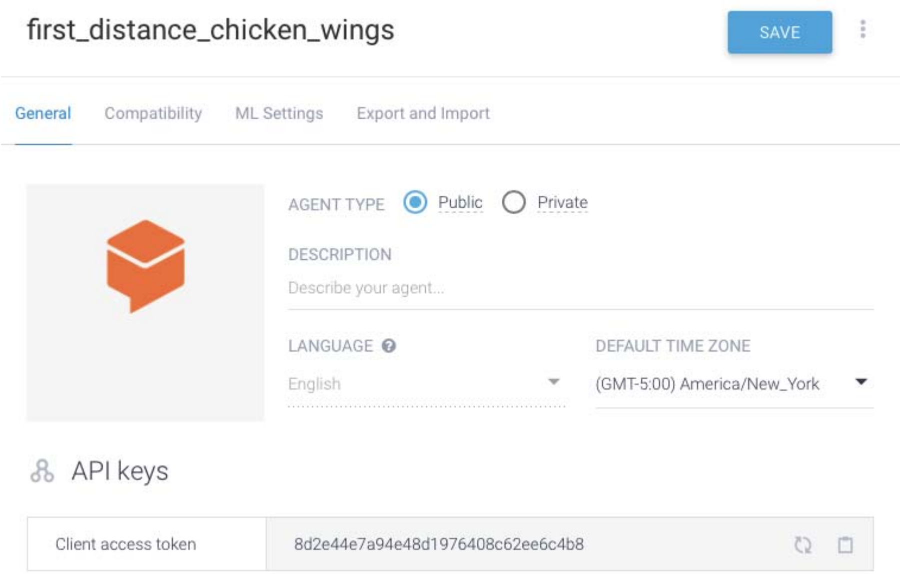


Figure 14 Api.ai (API keys)

2) **Intents** shows sets of specific conversations, the purpose is to transform natural user language into actionable data (API.AI, 2016). In this research, it contains questions and responses that are relevant from one location to another, (Figure 15) for example, from Chicken wings to the first stairs.

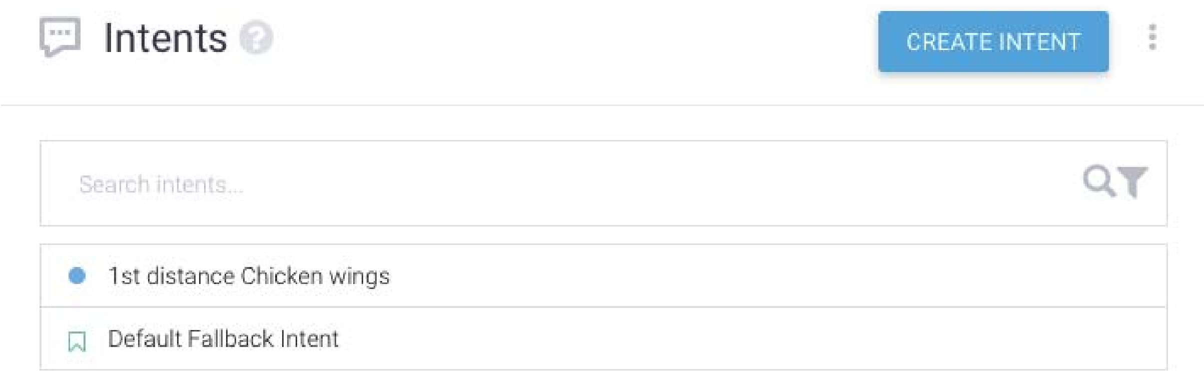


Figure 15 Api.ai (Intents)

3) **Entities** records sets of parameter values or common keywords from speech input. We do not need to create entities for every set that used in the agent, it can be recognized by agents (API.AI, 2016).

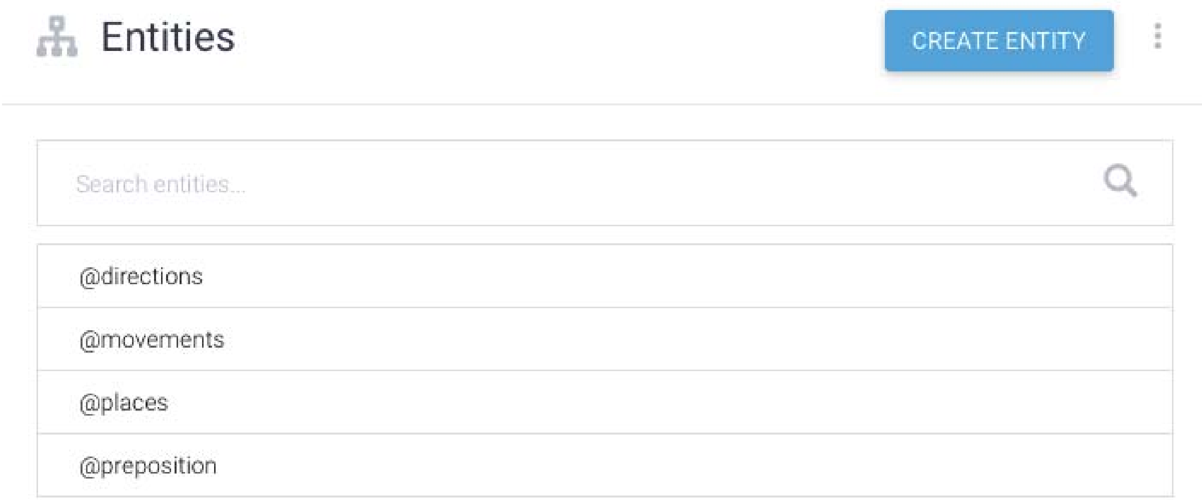


Figure 16 Api.ai (Entities)

Each entity contains all potential direction words, movement words, places and prepositions (Figure 13) that the user could use to interact with the system. We do not need to put specific words in again, only @ (the name of the entity) into sentences. We are also able to import entity files from CSV or JSON formats.

Example:

"I am @preposition to the Atrium building". Once the user says anything related to preposition, the system will be able to recognise the words (Figures 17 and 18) similar to other entities.

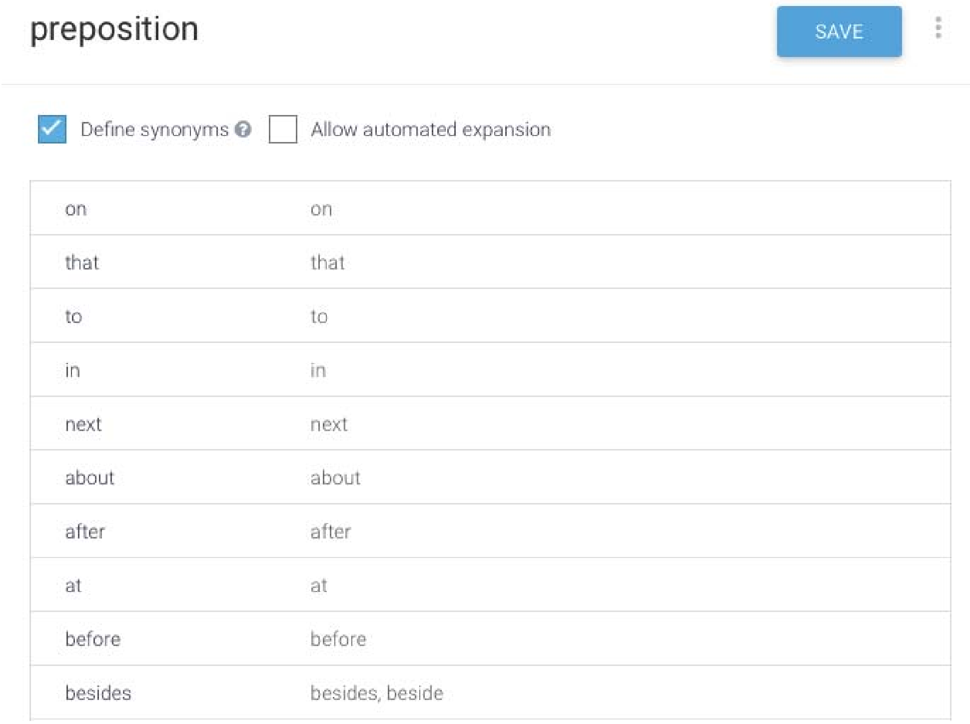


Figure 17 API.AI (An example of entity)

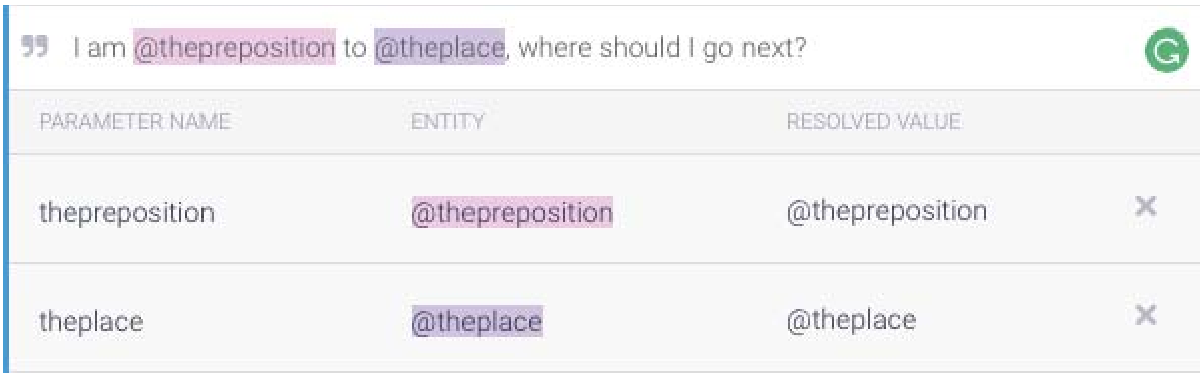


Figure 18 API.AI (The entity's process)

4) **Contexts** are expressions of the user (Figure 19). "This is helpful for differentiating phrases which may be vague or have different meanings depending on the user's preferences or geographic location, the current page in an app or the topic of conversation" (API.AI, 2016, p. https://api.ai/docs/contexts, line:1-3). For example, if the user reaches the library and wants to continue, they might say something like "I am next to the library, where do I go next?" In our design, we included several related answers for the user in order to make the communication between the system and the user more effective.

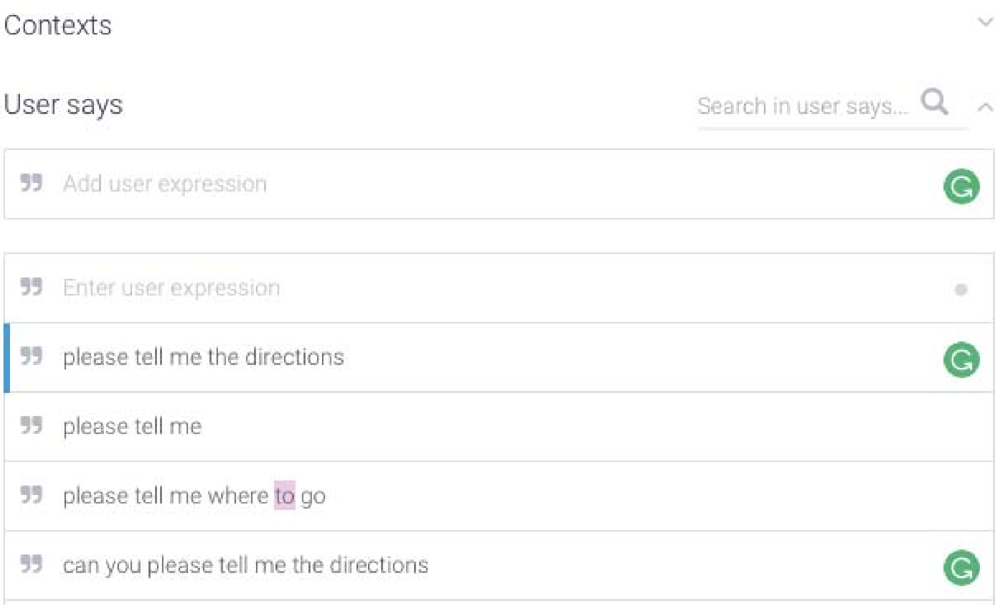


Figure 19 Api.ai (Contexts)

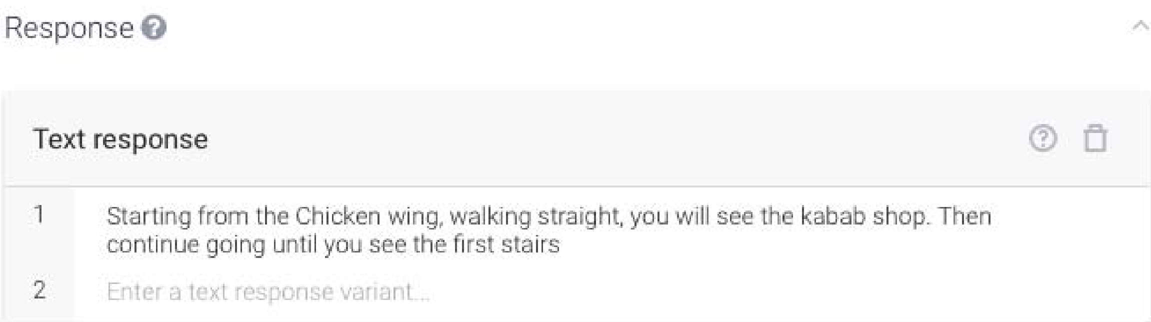


Figure 20 API.AI (Responses from the system)

Once the system receives the user requests, it will process the input and respond back to the user with the response that we have entered (Figure 20). All responses are shown in the move coding scheme section below.

4.3 Functionality

The conversational interface consists of four functions which help the interface meet research purposes.

4.3.1 The conversational function

All potential conversations between the user and the system is based on recorded data from our previous study (Longprasert, 2015). As explained above, we used the API.AI to create all the conversations. We created 11 agents, one from each location from the starting point to the destination and each agent had a unique access token key (Figure 21). The Android app was used to find the closest location by determining the current latitude and longitude of the user with GPS (Figure 22), and then submitting the user's speech with the matching access token key. For example, assuming the user arrives at point x on the map, the Android application gets the current user position using GPS, and determines the closest location. The access token key for that closest location is then used to submit the user's speech and API.AI responds with the answer that is appropriate for that location.

```
String ACCESS_TOKEN, place;
switch (closestWaypoint) {
    case 0:
        ACCESS_TOKEN = "8d2e44e7a94e48d1976408c62ee6c4b8";  //chicken wings
        place = "Chicken wing";
    //   ImageView img= (ImageView) findViewById(R.id.imageView);
        img.setImageResource(R.drawable.chickenwing);
        break;
    case 1:
        ACCESS_TOKEN = "c7d6e55704ef43ddbafad2d313af15ac"; // 1st stair
        place = "the first stair";
   //    ImageView img1= (ImageView) findViewById(R.id.imageView);
        img.setImageResource(R.drawable.firststairs);
        break;
    case 2:
        ACCESS_TOKEN = "f92809a3c3fe40bba90d35c393dfb5e6";  //bus stop
        place = "Bus stop";
     //  ImageView img2= (ImageView) findViewById(R.id.imageView);
       img.setImageResource(R.drawable.busstop);
        break;
    case 3:
        ACCESS_TOKEN = "785e24b3ee9940edb85886394ff504b2"; //Airtum building
        place = "Atruim building";
    //   ImageView img3= (ImageView) findViewById(R.id.imageView);
        img.setImageResource(R.drawable.atriumbuilding);
        break;
    case 4:
        ACCESS_TOKEN = "9669b7e13f48496bb26aef819ad37fdc"; //2ed stair
        place = "2nd stair";
   //  ImageView img4= (ImageView) findViewById(R.id.imageView);
      img.setImageResource(R.drawable.secondstirs);
        break;
    case 5:
        ACCESS_TOKEN = "c0649f74e7ad447191af3d6db151837a"; //library&Math building
        place = "Library&Math building ";
    //    ImageView img5= (ImageView) findViewById(R.id.imageView);
        img.setImageResource(R.drawable.mathandli);
        break;
    case 6:
        ACCESS_TOKEN = "8c0e92f3e4f947aab4015dd8088039cd"; //Student central
        place = "Student central";
     //   ImageView img6= (ImageView) findViewById(R.id.imageView);
```

Figure 21 Coding with ACCESS_TOKEN keys

```
private void initWaypoints() {
    this.waypoints = new double[15][2];
    waypoints[0][0] = -36.732378;  // chicken wings
    waypoints[0][1] = 174.701146;

    waypoints[1][0] = -36.732857; //1st stair
    waypoints[1][1] = 174.701198;

    waypoints[2][0] = -36.732926; // Bus stop
    waypoints[2][1] = 174.701175;

    waypoints[3][0] = -36.733036; // Atrium building
    waypoints[3][1] = 174.701188;

    waypoints[4][0] = -36.733326; // 2nd stair
    waypoints[4][1] = 174.701182;

    waypoints[5][0] = -36.733601; // Math building and library
    waypoints[5][1] = 174.701013;

    waypoints[6][0] = -36.733817; // student central
    waypoints[6][1] = 174.701425;

    waypoints[7][0] = -36.733991; // orbit&book shop
    waypoints[7][1] = 174.701724;

    waypoints[8][0] = -36.734040; // 3rd stair
    waypoints[8][1] = 174.701793;

    waypoints[9][0] = -36.734228; // Destination
    waypoints[9][1] = 174.702133;

    waypoints[10][0] = -36.734439; // out of way// (student accommodation village)
    waypoints[10][1] = 174.702523;
}
```

Figure 22 Coding with each waypoint

4.3.1.1 The move coding scheme

There are three types of conversations (1) "Initiating moves" (Carletta et al., 1996, p.22) are actions that happen at the beginning of the conversation. (2) "Response moves" (Carletta et al., 1996, p.19) always happen after the conversation has started e.g. the system's responses to the user's actions, movements or questions. (3) "Ready moves" (Carletta et al., 1996, p.22) often occur when the user almost arrives or has arrived at the destination and are used at the end of a conversation. And (4) "Repeat moves" which happens when the user makes mistake. Examples of the system's utterances were presented. All the system's utterances were based on real recorded conversations which the researcher collected the previous year (Longprasert, 2015).

A: Initiating moves

This often happens at the beginning of a conversation, where the user starts to speak to the system, including the instruct move, the explain move and the check move. In these and later examples, S is the system that gives directions, and U is the user, the one who is being given directions (Carletta et al., 1997).

(I.) "The Instruct Move" (p.15)

"An instruct move commands the system to carry out an action. Where actions are observable, the expected response could be performance of the actions" (Carletta et al., 1997, p.15).

Example 1

U: *Can you give me the directions, please?*

Tell me the directions, please?

Where should I go?

S: *Starting from the Chicken Wing, walking straight, you will see the kabab shop. Then continue going until you see the first stairs.*

(II.) "The Explain Move" (p.16)

An explain move, in the map task, the user asked simple questions to the system and it explained directions to the user. The user uses short sentences to interact with the system (Carletta et al., 1997).

Example 2

U*: I'm here, then what?*

S*: Great! Once, you have passed the first stairs, you will see the university bus stop on your left side. Then you will pass the Atrium building on the right side. Keep going straight to the stairs in front of you.*

Example 3

S: *You will see on the university bus stop on your left side. Then you will pass the Atrium building on the right side. Keep going straight to the stairs in front of you.*

Example 4

S: *Yes, that's right, after that, walk toward the stairs in front of you.*

Example 5

S: *after that, walk toward the stairs in front of you.*

Example 6

S: *Continue going in the same direction, until you see the Library entrance-way.*

Example 7

S: *Yes, in front of you should be Math building and the library should be on your right side. Then you turn left and pass student central.*

Example 8

S: *Sweet! you're getting there, keep walking and you will pass the Orbit Store and the Book Shop. And on your left side is the Sir Neil waters Lecture Theatre.*

Example 9

S: *Keep walking about 50 meters, you will see another stair.*

Example 10

S: *the book shop is on your right side next to Orbit shop and it is opposite the Sir Neil waters Lecture Theatres.*

Example 11

S: *Keep walking straight up to the stairs, you will see your destination, the student accommodation village.*

Example 12

S: you have arrived. Your destination is the student accommodation village.

B: Response moves

The response move is used within conversations after an initiation and includes the acknowledge move and the clarify move Carletta et al., 1997).

(I.) "The Acknowledge Move" (p.22)

An acknowledge move indicates the system has heard the user's voice and that the user has understood and accepted (Carletta et al., 1997).

Example 13

U: *I am lost*

S: *Oh dear! please go back to the previous location, once you see the location on your screen, you can ask me again.*

U: *Ok, I'm at (the place), what's next?*

(II.) The Clarify Move

A clarify move is a statement that applies to some questions where the system gives more explanations to the user in order to give the user a clearer idea of the location (Carletta et al., 1997). The conversation below shows that the user asked about the location and the system clarified the information, then explained more detail to help the user understand or see that location more easily.

Example 14

U: *Where is the book shop?*

S: *The book shop is on your right side next to Orbit shop and is opposite to the Sir Neil Waters Lecture Theatres.*

C: Ready Moves

A ready move is a move that happens when the user has almost arrived and after arriving at the location or destination. The system uses utterances such as “OK”, “that’s right”, “cool”, “great”. It is also followed by EXPLAIN (Carletta et al., 1997).

Example 15

U: *I can see it now*

S: *Sweet! You're getting there, keep walking and you will pass the Orbit Store and the Book Shop. And on your left side is the Sir Neil Waters Lecture Theatres.*

Example 16

S: *Yes, that's right, after that walk toward to the stairs in front of you.*

Example 17

S: *Awesome, you have arrived. Your destination is the student accommodation village.*

D: Repeat moves

A repeat move is a move that happens when the user walks through the wrong track or some unstable conditions (e.g. wind and noise) which beyond the system’s ability to recognize. The user has to repeat their questions or move back to the track that be able to communicate with the system until it able to repeat the directions information to the user.

Example 18

U: *I don’t know where to go.*

S: *Oh dear, you are lost, please go back to the previous location* **or**

S: *Sorry, I don't understand what you have said, repeat again.*

4.3.2 The map floating function

This function was used only with participants in Group C, to see how efficient the system was in helping the user achieve their goals of using conversation together with the map. In this design, the conversational interface is the main page and the map (OsmAnd) is a floating function on the map. We used an instant Android application named The C Floating to create the floating icon for the map (Figure 23). This allows the user to swiftly go to the map to check their position.

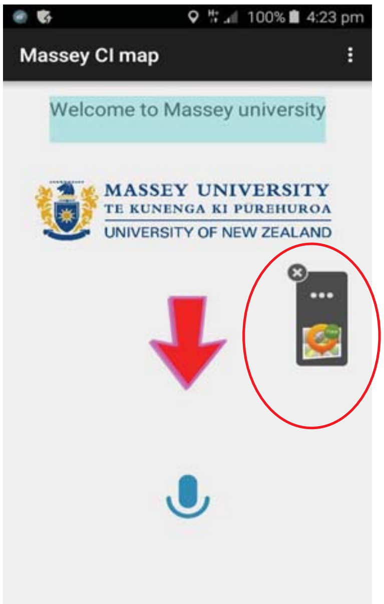


Figure 23 The conversational with map interface

4.3.3 The image function

This function, used with Group D, allowed participants to talk to the system in the same way as for Group B, but to also see a photo of the locations along the route. We uploaded real photos of each waypoint to **Res/drawble** of the conversational interface and used the "setImageResource" method to display the photo. *Res/drawble is a drawable resource for graphic that can be drawn to the screen. Once, the user walks to the closest location of that location, the photo of that location will be shown (Figure 24).

Example code:

```
case 8:
    ACCESS_TOKEN = "4854d592aceb44f495efc27abe94dbbb"; //3rd stairs
    place = "the 3rd stairs";
  ImageView img8= (ImageView) findViewById(R.id.imageView)
    img.setImageResource(R.drawable.thirdstairs);
    break;
```

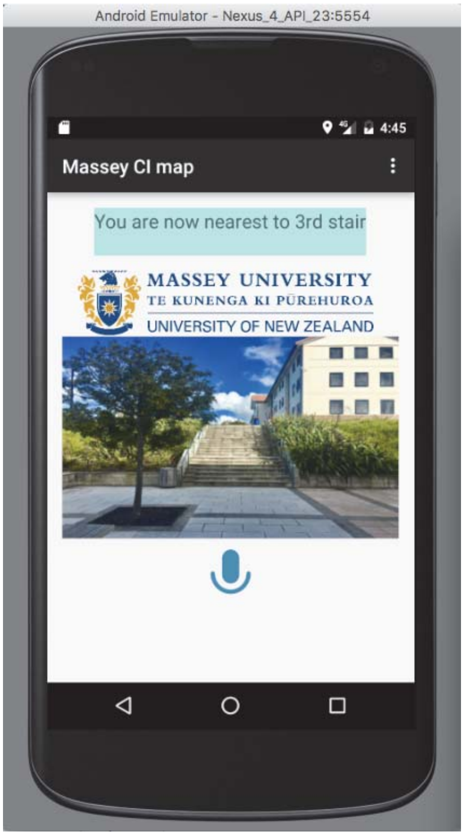


Figure 24 The conversational with image interface

Chapter 5

Results

This section presents and compares the results from four data collection exercises using task performance for each of the four different kinds of interface following the same route (from the Chicken Wing to the student accommodation village on Massey University Campus, Albany, New Zealand). There were 100 participants in this research and all participants were divided into four groups (map-only interface, conversational-only interface, conversational with map interface and conversational with images interface). In the first analysis, we looked at the accuracy of the experiment to see how different navigation interfaces compared with the route tracking in different interfaces. We then looked at the speed that participants took in the task for each location and different groups, and compared different subgroups of respondents. In the second analysis, we analyzed the conversation, identified the questions that most participants liked to ask the system and categorized them into three subgroups. Next, we analyzed the Usability questionnaire and looked at the interviews last to seek more information about future development.

5.1 Participant characteristic analysis

This section analyzes the answers to questions about the characteristics of the participants, including the age and experience using navigation systems. The age of participants was distributed from 18 to over 70 with different experience of using navigation systems. Participants in this research are mainly aged 18 to 28 years while only one of them was between 62 to 72 years in Group A (the map with command interface). Participants in all groups tended to use navigation systems almost every day, especially 16 participants in Group A. There were about 2 to 3 people who said, “Because of my job I need to travel every day, so the navigation system is a part of my daily life” and none of participants had no experience using navigation systems.

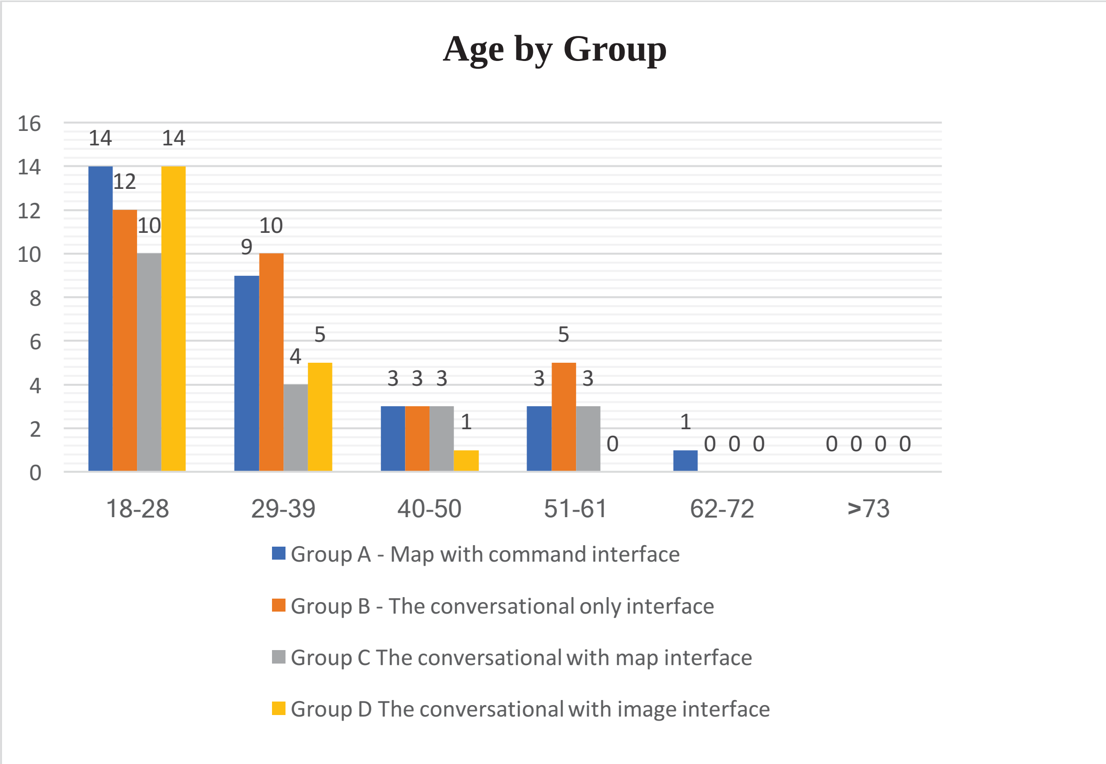


Figure 25 The graph of age of four groups

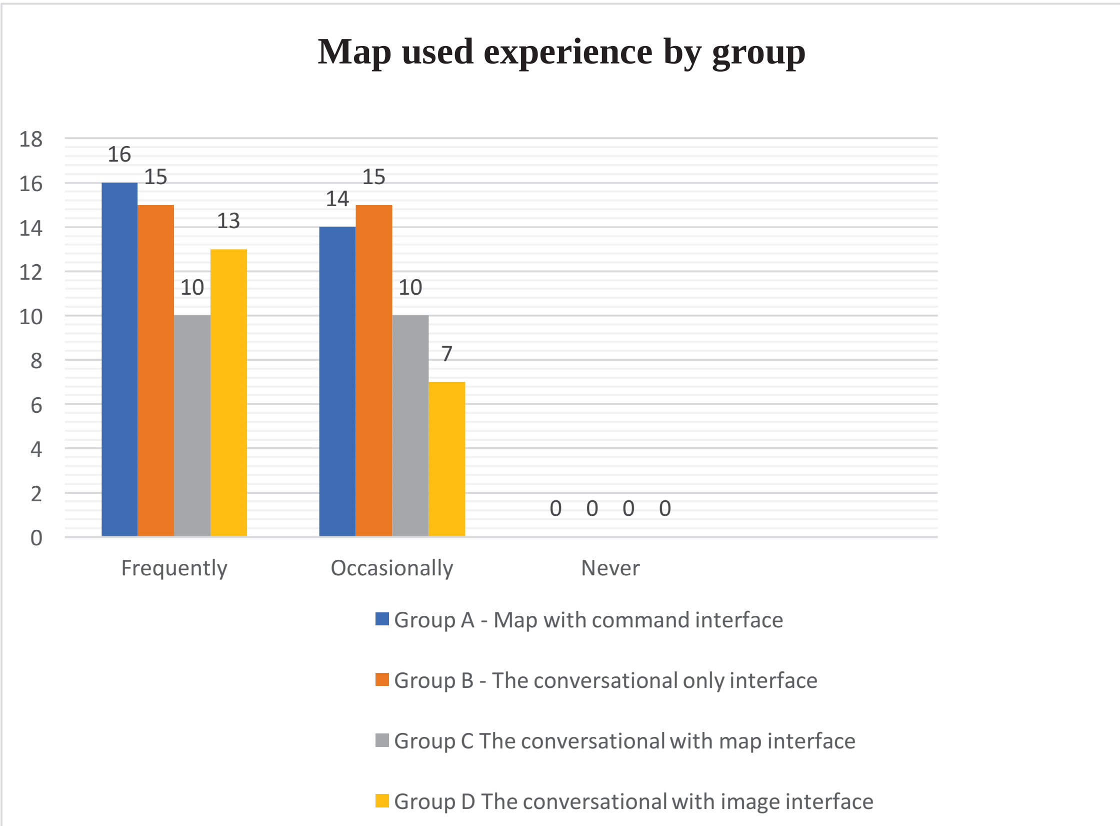


Figure 26 The graph of experiences of four groups

5.2 Route analysis

This section analyzes the route from the Chicken Wing to the student accommodation village on Massey University Campus, Albany, New Zealand for four different groups. The researcher used the SMODROID and the AZ Screen Recorder application to monitor the users' paths while they were walking to the destination. The purpose of this was to analyze the accuracy for the four kinds of interface, including the map-only interface, the conversational-only interface, the conversational with map interface and the conversational with images interface.

5.2.1 The Overview of the route tracking

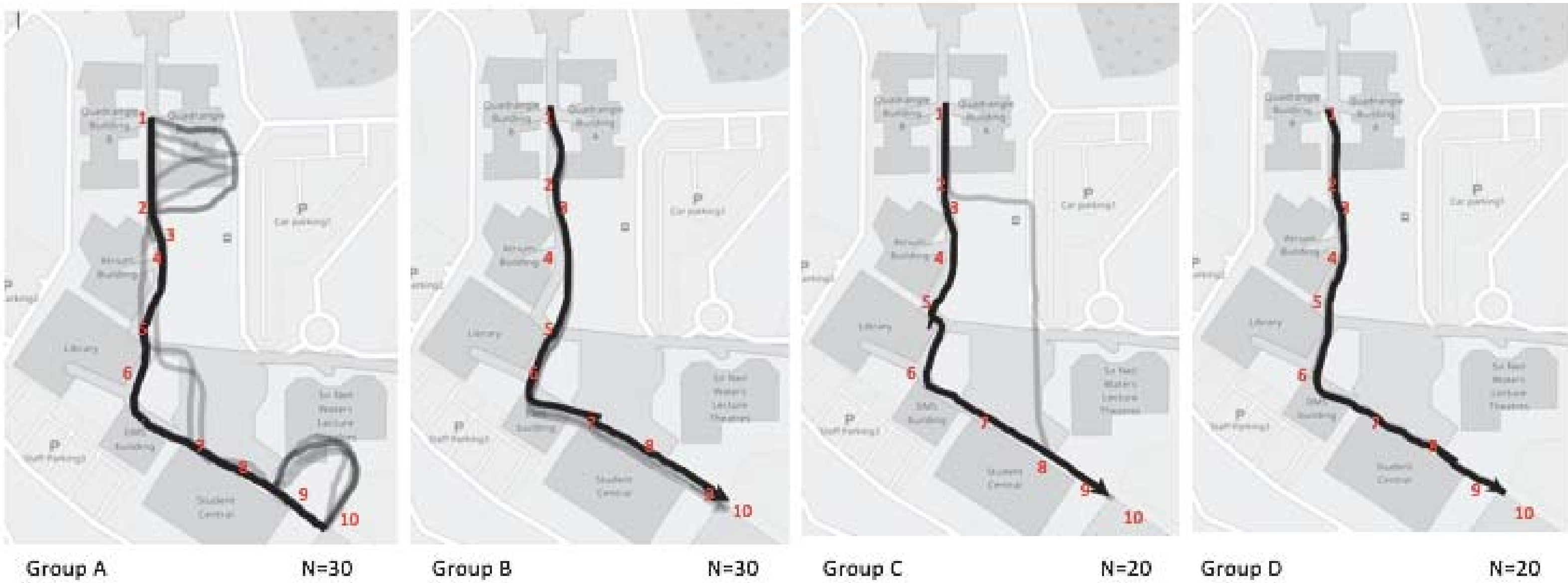
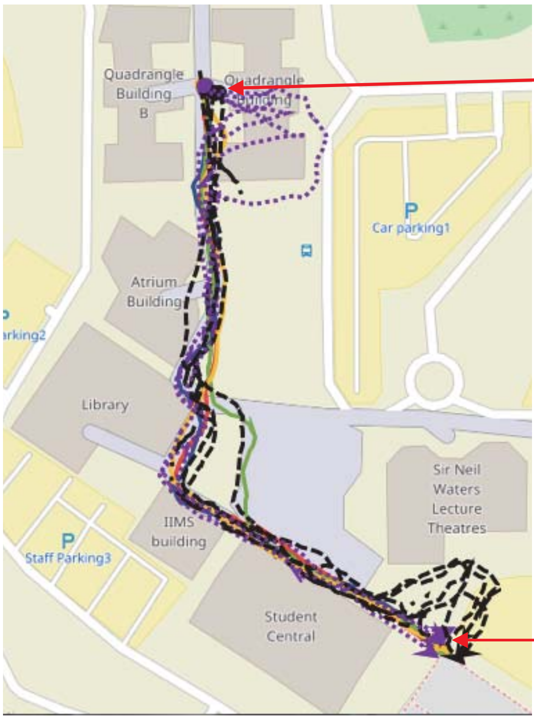


Figure 27 The overall of the route tracking

Figure 27 shows the accuracy to which members of each group were able to follow the directed path. Each participant's path is represented by a 70% greyscale line. Dark lines represent the common route that most participants took and gray lines show the route that fewer people followed to this area. It can be seen that most of participants (dark lines) walked past key points but some of them moved away from key points (gray lines) and got lost. This is especially evident of Group A who used the OsmAnd map application and Group C who used the conversational with map interface. Only one participant in Group C got lost half way from the starting point because of misunderstanding instructions provided by the system. After that, she switched the conversational page to the map page and followed the arrow back to the track. There was no problem with Groups B and D (more detail can be seen below). Dark lines represent the common route that most participants walked and gray lines show the route that fewer people followed to this area.

5.2.2 The map with command interface (Group A)



The starting point
(The Chicken Wing)

The destination
(The student accommodation village)

Figure 28 The route results of group A

Thirty participants tested the map and command interface. About four participants (in the purple lines) got lost between the starting point and the Quadrangle Building A due to confusion with instructions given by the system “50 meters turn slightly left and then keep right” while no directions were shown to turn either left or right at that location on the screen (Figure 29). Another location was around point 8 to 9 which is the book shop (B) and The Sir Neil Waters Lecture Theatres (A) to the 3rd stairs (Figure 28). About six participants (the black lines) got lost by the system’s instructions “turn left and then zero meter turn right” and then the system kept saying to the user to turn right which was different from the direction that presented on the screen (Figure 31). The one green line was the person who walked in a slightly different way, but still in the right track.

From observation, ten participants became lost because they relied on the system’s commands and did not pay attention to the screen which seemed to be showing more accurate information than the commands. When the system gave instructions, those participants just followed what the system said without double checking the directions on the screen. Once they got lost, they started looking at the map on the screen. Twenty participants reached the destination without any problem. They relied on both the map and the voice commands. Many participants said, “the system’s instructions were not corrected at some points, so I felt it was better to look at the screen as well as listen to the voice command so as not to be led the wrong way through unclear instructions”.

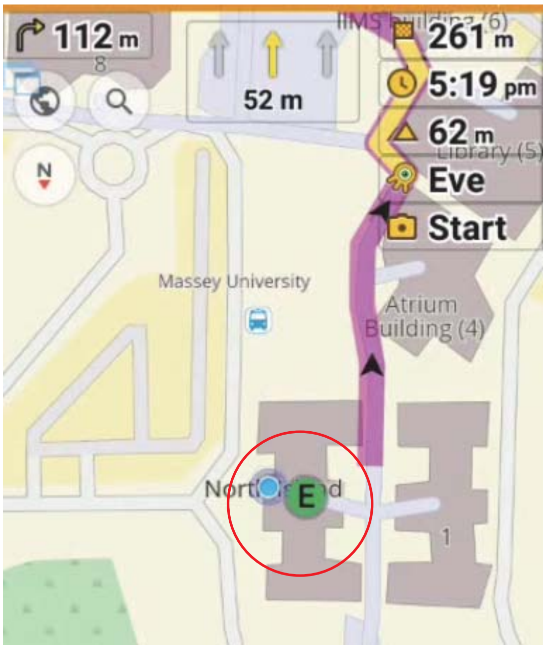


Figure 29 The route that many participants in Group A got lost (First point)

Figure 30 The route between the book shop and The Sir Neil Waters Lecture Theatres

Figure 31 The route that many participants in Group A got lost (Second point)

5.2.3 The conversational interface (Group B)

Figure 32 The route results of Group B

This group tested the conversational interface with which they could talk to the system, but had no visual aids.

Thirty participants reached the destination smoothly, no one getting confused or lost by the system's instructions. The system gave clear and accurate instructions to the users and was easy to understand "I love how it is very detailed with exact places and landmarks. It tells me how far I need to walk. I love how I can ask at any time where to go. I feel confident in asking it because the voice is calm and gentle." (participant)

However, most participants commented about the windy conditions which was a major problem for the conversational group. At some areas in the campus, such as between the Mathematical Sciences building and the Library building (Figure 32&33), most participants suffered from wind noise on windy days, so the system could not recognize their voices and gave information very slowly. They wasted the most time around this area.

Figure 33 The area that most participants were suffered from wind

5.2.4 The conversational with map interface (Group C)

Figure 34 The route results of Group C

Twenty participants tested the conversational with map interface, which has the conversational interface as the main page and the map as a floating function on the top of the page (which users can choose to open and then switch between the two screens). 18 participants smoothly followed the directions given by the system without any mistakes or confusion. They used the conversational interface along the way to the destination only one participant changed to the map mode because she suffered with wind noise around the Mathematical Sciences building and the Library building area. She spent a few minutes trying to talk to the system but it was very windy and finally she decided to switch to the map mode. However, she was still satisfied using the conversational interface as the main page. “The wind made the system unable to recognize my voice, but it was still good that I can talk to the system, not just only listen to it and it was great at the beginning, the system responded to me so well.” (participant). Another participant in this group got lost almost half way from the bus stop to student central (Figure 34), she was a visitor and was unfamiliar with the campus. She was using the conversational interface at the beginning till she reached student central, then she changed to the map mode, which finally helped her to get to the destination. She gave a comment that “I personally prefer to see the map more than listening, does not matter how good and clear instructions are, I still get lost and lose confidence if I cannot see anything on the screen” As in Yankelovich et al., (1995) reported that the user is more likely to be comfortable of using a visual interface than other kinds because seeing things on the screen could make them more confident to complete the task. However, with a conversational interface, the user can only communicate with the system by their voice without seeing any pictures or maps. It could cause the user to become confused and panic while using the system.

5.2.5 The conversational with images interface (Group D)

Figure 35 The route results of group D

The conversational with real images interface (Figure 35) was used by 20 participants in the last group. All participants arrived at the destination. However, many participants from this group also suffered with the windy conditions in the same area as Groups 2 and 3 (around the Mathematical Sciences and Library buildings). Some of them took a while to pass that area but eventually, they arrived at the destination and nobody in this group got lost.

5.2.6 The average distance for each location in different groups

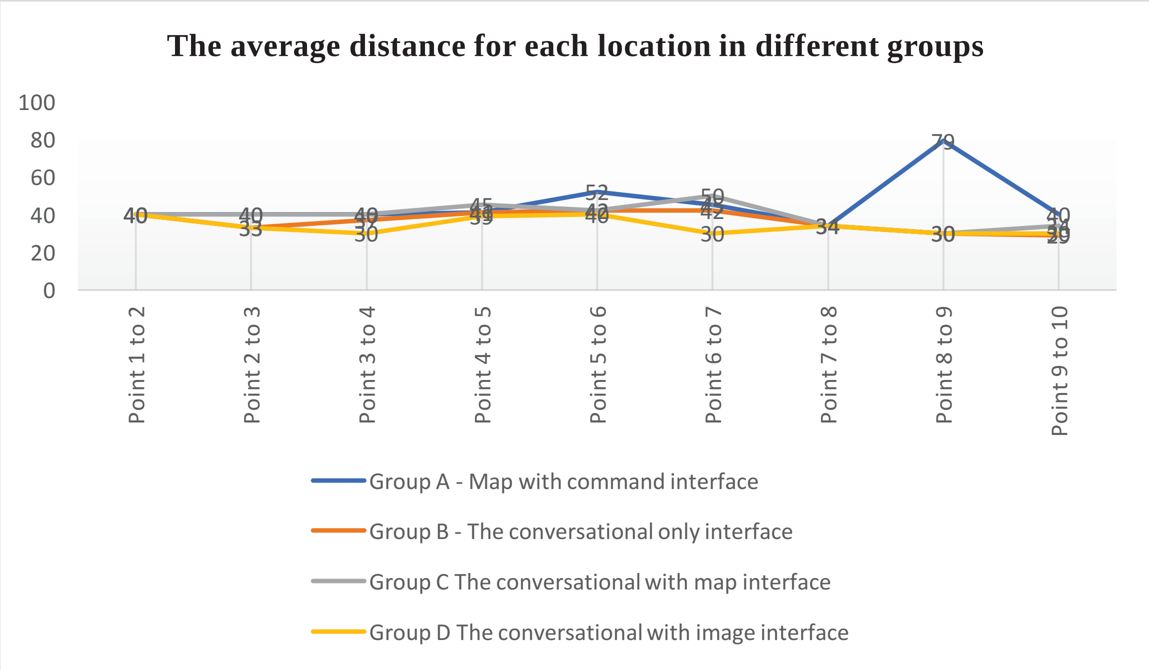


Figure 36 The average distance for each location in different groups

Table 4 The average distance between each location in different groups

Participants	Group A	Group B	Group C	Group D
Mean total	411	320	355	302

Unit = metre

Table 4 and figure 36 indicates the average distance between each point for the four groups. On average, people who used the conversational interface walked a shorter route than people who used the OsmAnd. As the results show, Group A had the longest distance (411 metres) as several participants in this group got lost and out of two points (point 1 to 2, and point 8 to 9.) For each point in group A, participants walked the average distance of 79 metres, which was the longest distance compared to other points. It was at point 8-9 that several participants got lost. Groups B and D from point 6 to 7 was the longest distance that participants walked through (42 and 50 metres). Group D differed from Groups B and C by walking a shorter distance.

5.2.7 The differences of the average for distance in different groups

Table 5 The different of the average for distance in different groups

	Group A	Group B	Group C	Group D
Group A	0	**91**	56	**109**
Group B		0	-35	18
Group C			0	33
Group D				0

Unit = metre

Table 5 indicates the difference in average distance between each pair of groups. Across for the whole route. The figures that are highlighted are statistically significant at 0.05 confidence level.

It can be seen that Group A walked a longer distance than Group B (91 metres) and Group D (109 metres) because several participants in Group A got lost while in Groups B and D, participants found all waypoints to the destination without any mistakes. Participants in Group C walked a longer distance than Group D (33 metres). This was because a participant in Group C was totally lost half way from the starting point. She pointed out that she personally prefers to see the map more than listening, it does not matter how good and clear instructions are, she still gets lost and loses confidence if she cannot see anything on the screen. During the experiment, she spoke to the system and changed to the map mode but the arrow on the screen was a bit slow, that meant she walked out of the track for a few seconds. After that, she was fine to use the map page which successfully directed her to the destination.

5.2.8. Summary

The route analysis compared the accuracy of the four different interfaces. The results above showed that ten participants in group A got lost with the main problem of the group being the voice commands given by the OsmAnd system which were inaccurate and some commands were incorrect or confusing. For instance, (1.) “turn slightly right/left” while the user was walking straight as they were following the map on the screen. (2.) “50 meters turn slightly left and then keep right” most of participants got lost and confused by this instruction. Groups B, C, and D had the conversational interface with which they could interact. All arrived at the destination without any mistake or misunderstanding of what the system said. Group C had only one participant who got lost by using the conversational with map interface. She reported that it was because she does not use navigation technology so often and she prefers seeing the map on the screen. However, the major barrier to using the conversational interface were the weather conditions, especially the wind noise. The system could not recognize the users’ voices and they needed to keep speaking to the system for some time before being understood. We also looked at the accuracy in distance travelled between each pair of waypoints for the four interfaces. Results showed that points 6 to 7 which is from The Library & Math building to the student central was the longest distance that participants in Group B and C walked, and points 8 to 9 which is from the Book Shop to the 3rd Stairs was the longest distance (79 metres) for participants in Group A.

Overall, for the routing the conversational interface has instructions to guide people to reach to directions more accurately than the normal navigation system (OsmAnd). They walked to exact key points better than people who used the OsmAnd.

5.3 Time-taken analysis

5.3.1 The average of time-taken analysis for each location

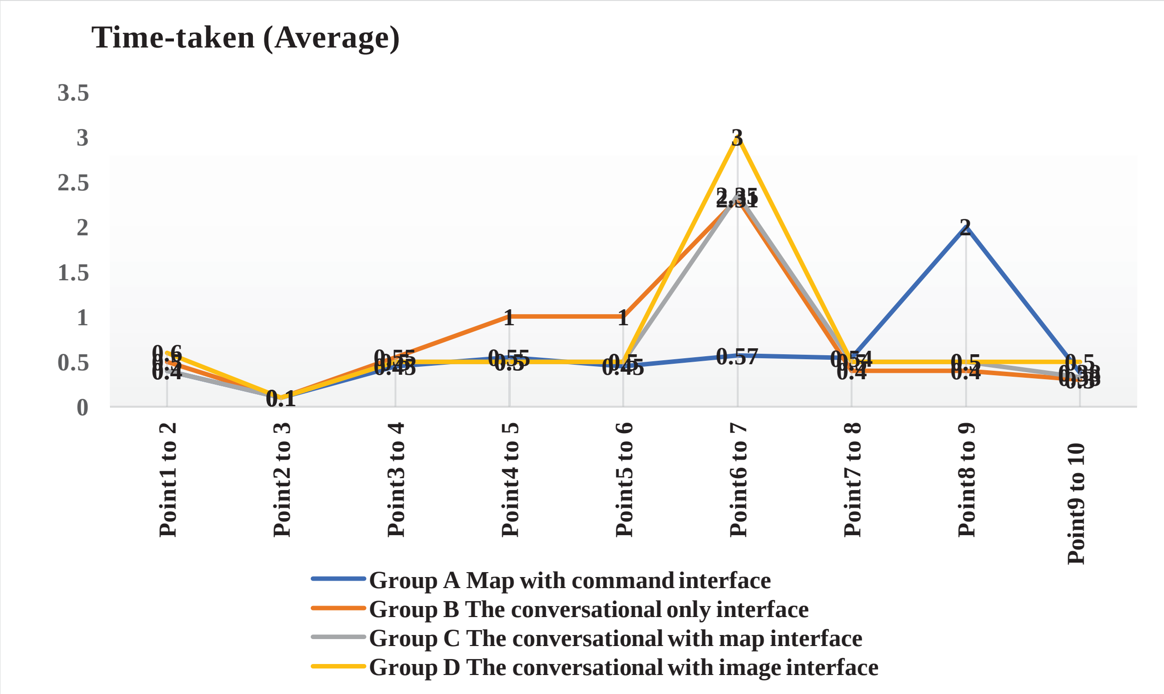


Figure 37 The graph of the time-taken analysis for each location

The data used in this section presents the time taken for different kinds of navigation interfaces. Figure 37, compares the average time taken in each location from the starting point (Chicken Wing) to the destination (the student accommodation village) of the four kinds of navigation interfaces which the users were testing. On average, Group A, which used the map with command interface, spent the shortest time, 5.44 mins, to reach the destination, while the other three groups spent 6.56 mins (Group B), 5.68 mins (Group C) and 6.70 mins (Group D). At point 2 to point 3 which is from the second stairs to the bus stop, participants spent only 10 seconds for all kinds of navigation interface. However, at point 8 to point 9, which is between the bookshop to the third stairs area (Figure 34), the users of Group A were getting lost by the system's instructions and turned a different direction. They spent about 2 mins finding the way back to the third stairs. Groups B, C, and D, using the conversational interface, didn't suffer this problem and spent about 40 to 50 secs. to pass this point. The users in Groups B, C, and D spent more time between point 6 and point 7: from the Mathematics & Library building to the student central (Figure 32) because wind noise interrupted the performance of the system. At this point, many users attempted to communicate with the system to get more information and they took about 2.30 to 3 mins to cross this point while the users in group A spent less than 1 min with this position and without any obstacles.

Table 6 The mean time-taken of four groups

Participants	Group A	Group B	Group C	Group D
Point 1 to 2	40 secs.	50 secs	40 secs.	1 min
Point 2 to 3	10 secs.	20 secs	10 secs.	10 secs.
Point 3 to 4	45 secs.	55 secs	50 secs.	50 secs.
Point 4 to 5	55 secs.	1 min	50 secs.	50 secs.
Point 5 to 6	45 secs.	1 min	50 secs.	50 secs.
Point 6 to 7	57 secs.	2.31 mins	2.35 mins	3 mins
Point 7 to 8	54 secs.	40 secs.	50 secs.	50 secs.
Point 8 to 9	2 mins	40 secs.	50 secs.	50 secs.
Point 9 to 10	38 secs.	30 secs.	33 secs.	50 secs.
Mean total	5.44 mins	6.56 mins	5.68 mins	6.7 mins
Median Total	5.08 mins	6.10 mins	5.53 mins	7 mins

5.3.2 The differences of the average time-taken of each group

The table below indicates the difference in average time taken between pairs of groups. The highlighted numbers are statistically significant at the 0.05 confidence level.

Table 7 The differences of the average time-taken of different groups

	Group A	**Group B**	**Group C**	**Group D**
Group A	0	-1.12	-0.24	1.26
Group B		0	0.88	-0.14
Group C			0	-1.02
Group D				0

Unit=min

The data of table 7 indicates the differences in the time taken are significant at a level of 0.05 in several cases such as Groups A and B, Groups A and D and Groups C and D and shows similar patterns as the analysis for distance travelled. The table shows that Group A spent less time (-1.12 mins) to reach the destination than Group B as well as Group D. The reason was most participants were in conversational interface groups, and they stopped at each location to speak to the system. Sometimes the system could not recognize their voices, and they had to keep speaking, so it took a while for them to reach the next location. Another problem of the conversational interface groups was the wind noise, which seriously distracted them at some points during the experiment. As report in Yankelovich et al., (1995) the challenges for the speech interface are the users' utterance or background noise that can cause the system to be unable to recognize the speech. The user had to repeat or rephrase their speech until the system was able to understand.

5.3.3 The difference in average time taken gender

A: Gender

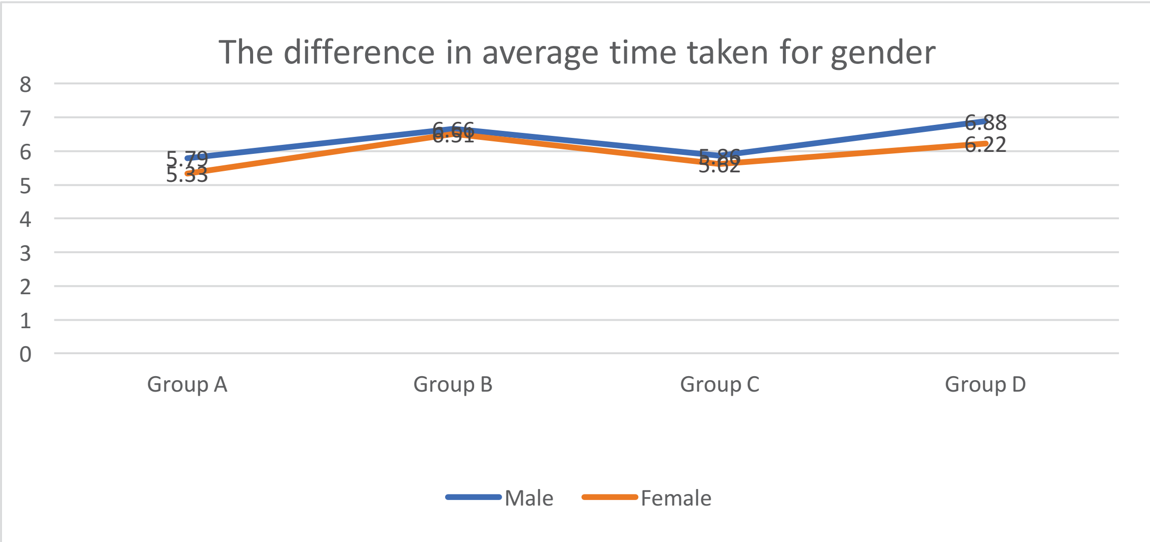


Figure 38 The different in average time-taken for gender

B: Age

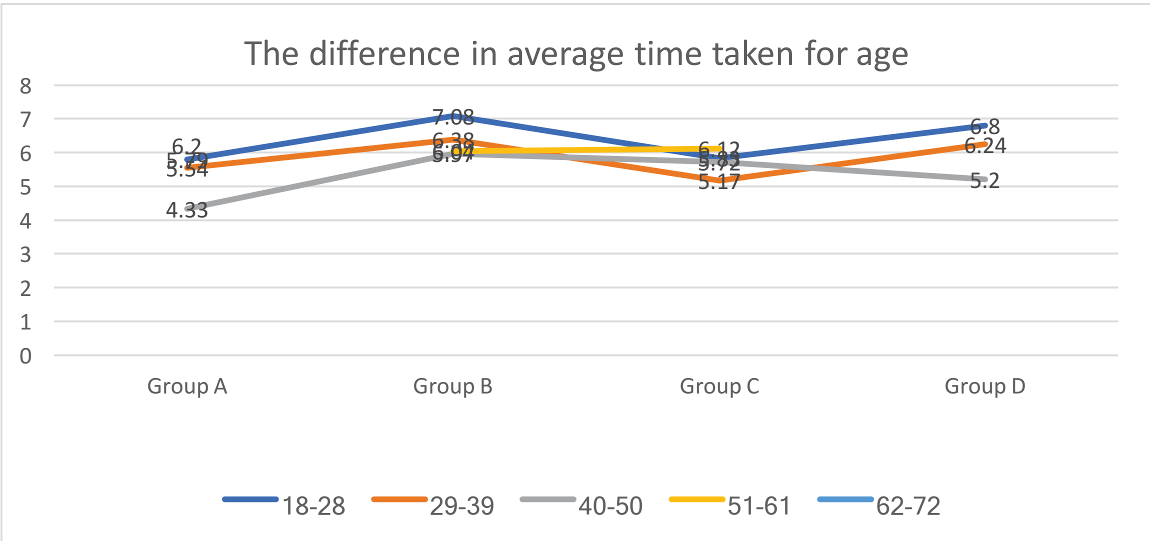


Figure 39 The difference in average time-taken for age

C: Experience with navigation

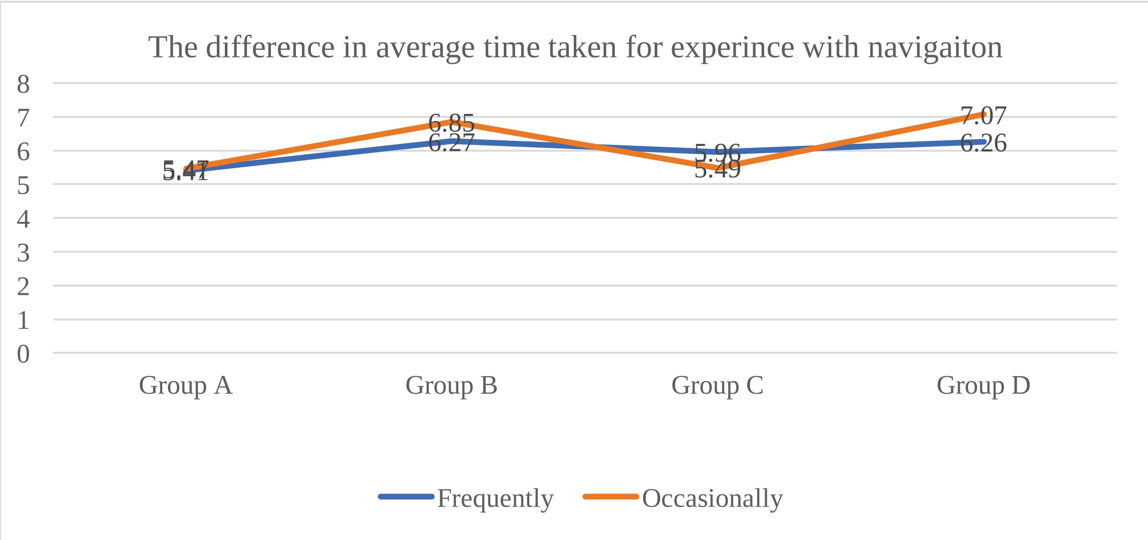


Figure 40 The difference in average time-taken for experience with navigation

Figures 38, 39 and 40 demonstrate the difference in average times taken for Gender, Age, and Experience using navigation. None of these differences were statistically significant at level 0.05. As for gender, females walked faster than males in all groups of navigation interface. Surprisingly, the ages of 40 to 50 in Groups A, B and D spent less time walking to the destination than other age groups, except participants in Group C, the fastest time being 4.33 mins in Group A while the slowest time was participants who were aged 18 to 28. The last criterion was about experience using navigation. Those people who used navigation frequently spent time less than people who used it just occasionally, except people in Group C who had used navigation on some occasions and spent a longer time than people who often used navigation, about 5.96 mins.

5.3.4 Time taken analysis for different groups of navigation interface

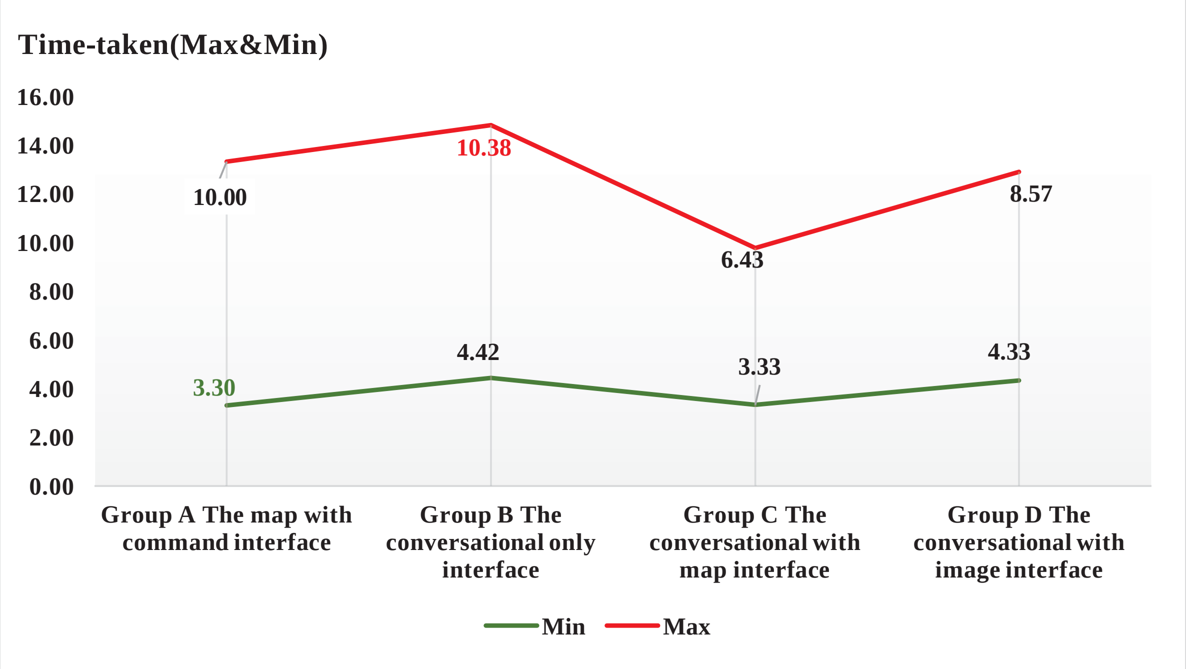


Figure 41 The graph of time-taken (Max & Min) for different groups of navigation interface

Apparently, from the maximum and minimum time taken analysis (Figure 41), group B which is "the conversational only interface" spent a longer time than "the map with command interface". The maximum for Group C is much lower than all the others, even though the average is higher than for Group A, and there was one participant who went quite a long way out. She decided to change the mode to the map and walked as fast as she could to the destination feeling uncomfortable with the instructions.

People that walked the longest distance were the fastest (Group A) because, they need only to look at the screen, which enabled them to easily see directions. Even though the commands of the system were not accurate enough, they could see the object on the screen and that made them feel confident to walk through to the destination. Similarly, Groups B, C and D, used the shortest distance but slowest time compared to Group A, because while they walked to the next location, they stopped to speak to the system, and were distracted by the windy conditions, a major issue for the "conversational interface" groups.

5.3.5 Summary

The location where the three groups spent much time was between the Mathematics and Library buildings, about 2.31 to 3 mins. The major problem for these three groups was the windy conditions at that point which made the system unable to easily recognize participants' voices. And another obvious reason was that when participants talked to the system they tended to stop walking and asked directions rather than continuing to walk. This is another reason the "conversational interface" groups took longer than the "map-only interface" group.

Another method was analyzing the maximum and minimum time taken the whole way for four kinds of navigation interface. The "conversational-only interface" group spent the longest time reaching the destination and, surprisingly, the "map-only interface" group spent a longer time than the "conversational with map" and the "conversational with images" groups. Whereas, the minimum time-taken for the "map-only interface" group was only 3.30 mins while other three groups were 4.42 mins (Group B), 3.33 mins (Group C) and 4.33 mins (Group D). From both the average and minimum time taken analyses, participants in the "map with commands" interface group (Group A) spent less time arriving at the destination than the other three groups.

5.4 The conversational analysis

This section presents the conversational data between participants and the system for three kinds of navigation interface: the conversational-only, the conversational with map, and the conversational with images interfaces. The map with command interface did not allow users to talk to the system and is not discussed in this section. Lists of questions were provided for participants in these three groups at the beginning of the experiment (Appendix B) to provide some ideas of sentences that participants could use. They were also told that they could use other sentences.

The speech part of the system within conversations focused on a three moves coding scheme by Carletta et al., (1997) for the map task dialogues. (1) "**Initiating moves**" (p.22) which always happens in the first conversation. (2) "**Response moves"** (p.19) are used after the conversation has started. (3) "**Ready moves"** (p.22) occur at the end of a conversation, and prepares new conversations to be started over again, with common words such as "OK" and "Alright".

The first part analyzes the three moves coding scheme for the system's speech followed by the rate of the user's utterances. From observation, all participants of the three groups followed given sentences about half of the time and a few of them came up with their own sentences.

5.4.1. Results of the user's utterances for three groups of conversational interfaces in initiating moves

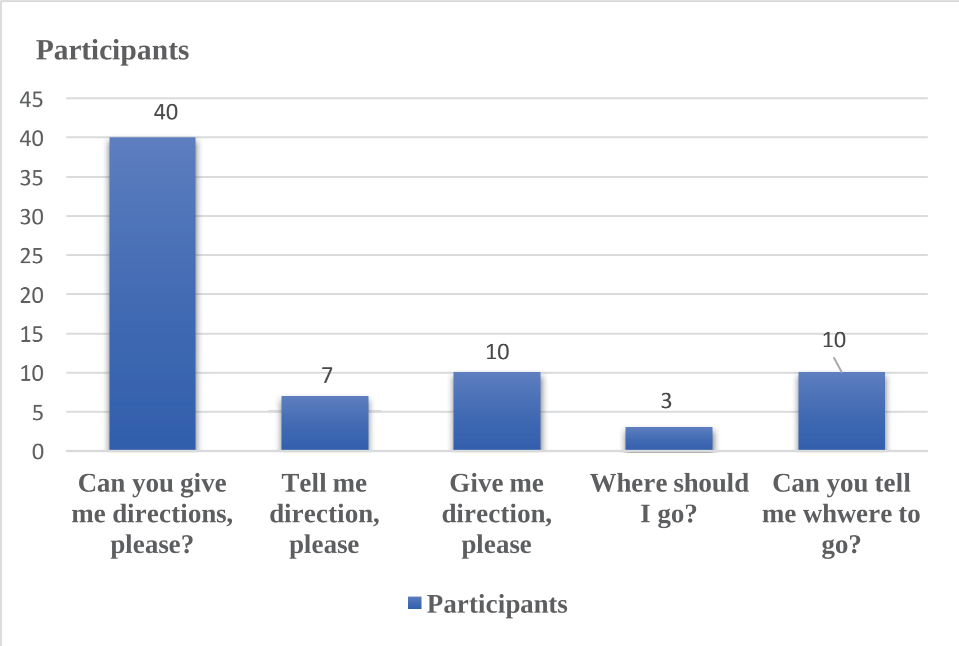


Figure 42 The graph of the user's utterances for three groups of conversational interface (The starting sentences)

Figure 42 presents initiating moves that happen at the beginning of a conversation. In the experiment section, the researcher has already set the destination, so the user asks for directions from the system and follows its instructions. As the result below shows, 40 participants were using "Can you give me directions, please" as the starting conversation. We observed that most participants had trouble knowing what to say at the beginning. The researcher decided to give the list to participants and told them to try not to refer to the list. Some participants provided the feedback that "I can talk to the system in very short sentences, it is very easy to communicate. If I need to say long sentences, I'm afraid the system will not understand me as I am not a native English speaker." They only followed the example sentences list, even though the researcher had pointed out that they were only examples. At the beginning of conversations, no participants tried any other sentences.

5.4.2 Results of the user's utterances for three groups of conversational interfaces in response moves

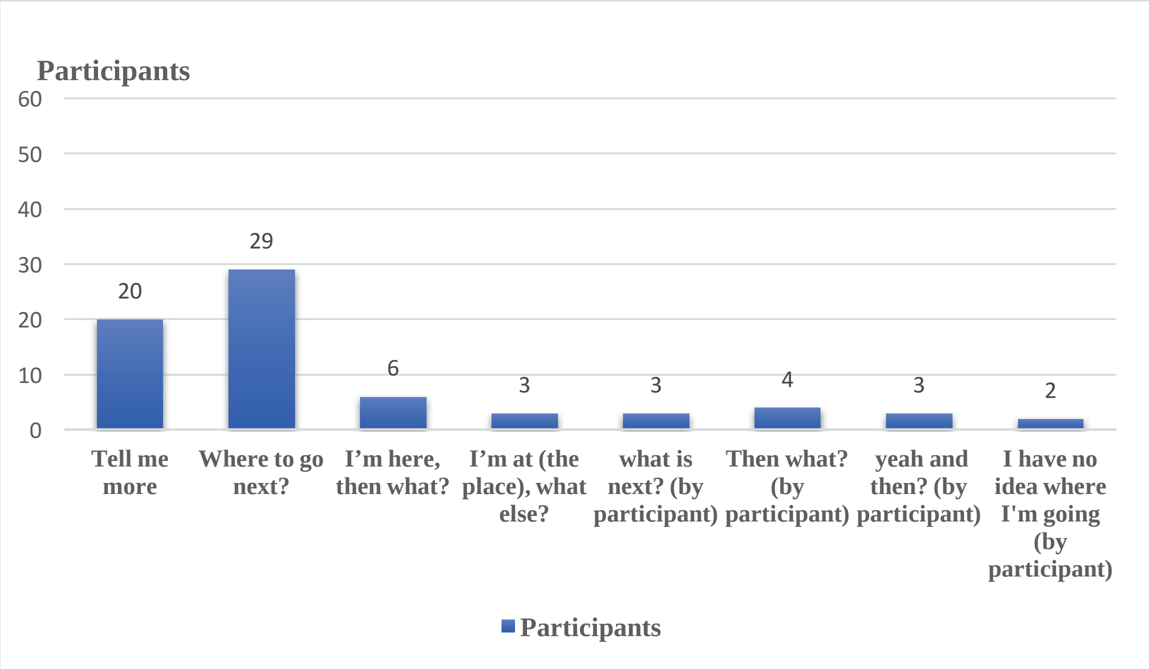


Figure 43 The graph of the user's utterances for three groups of conversational interface (the middle conversation sentences)

Figure 43 shows the response move which is used within conversations after an initiation (the second utterance) and includes the acknowledge move and the clarify move. These are sentences for confirming their position when the user has arrived at the location, and they interact with the system through the provision of a confirm sentence. There were about 29 out of 70 participants used the "where to go next?" sentence to respond to the system and 20 participants responded by saying "tell me more" to get more information from the system. After passing the Atrium building area, some participants felt confident enough to respond to the system without referring to the list, and they came up with some sentences. For example, three participants asked, "what is next?", while four of them asked "then what?" and "yeah and then?" and another two said "I have no idea where I'm going". This participant got lost in Group C (conversational with map).

5.4.3 Results of the user's utterances for three groups of conversational interfaces in ready moves

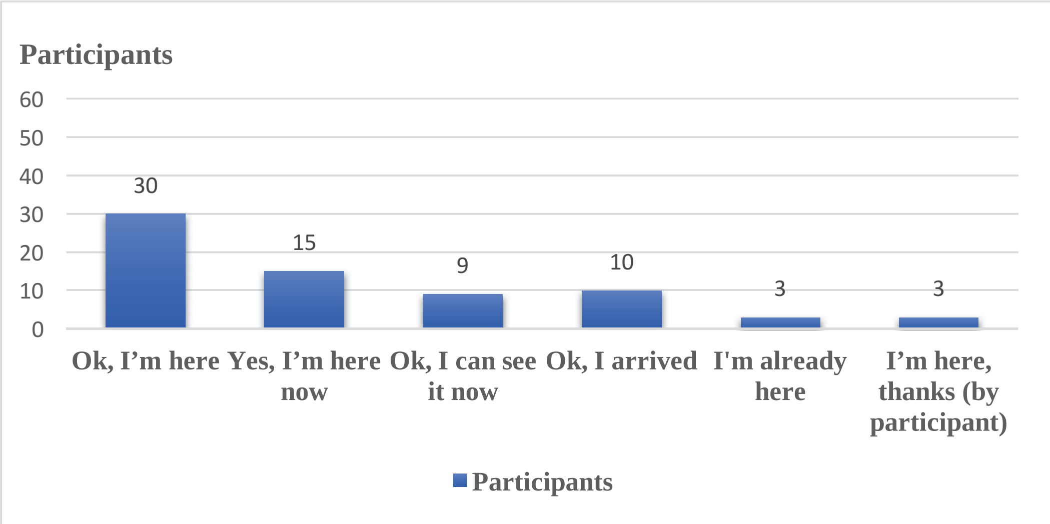


Figure 44 The graph of the user's utterances for three groups of conversational interface

A ready move (Figure 44) is a move that happens when the user has almost arrived or has arrived at the location or destination. The system uses utterances such as "OK", "that's right", "cool", "great". The sentences below were used when the user arrived at the destination (the student accommodation village). These are similar to confirming position sentences but these sentences are often used at the end of a conversation or when arriving at a destination. Once the user arrived at "the third stairs":

Example:

S: *Keep going straight to that stair in front of you. You will see the student accommodation village.*

U: *Ok, I can see it now.*
S: *Awesome! That is your destination. The student accommodation village.*

U: *Ok, I'm here.*

In the observation, the sentence that most participants said was "Ok, I'm here" (30 participants). Even though this sentence is on the list, they were saying it naturally without thinking and most of them did not look at the list. The expressions "Yes, I'm here now" and "Ok, I can see it now" were used with similar frequency (about 10 for each). Only three participants tried another sentence, namely "I'm here, thanks"

5.4.4. Summary

The rates of common sentences that the user spoke to the system. During the experiment, most participants felt panicked and they had nothing to say to the system for the first two minutes. The research showed the list of utterances but when doing the real test, they could not think or start asking the system. Several participants spoke to the system without relying on the list and the system could recognize their words (e.g. “Then what?”, “I have no idea where I’m going”). The utterance, “Ok, I’m here” was the most commonly used by participants at almost the end of a conversation (participants almost arrived or arrived at the destination). **(The Ready move)** at this state they were confident to talk to the system and did not look at the list.

Overall, the reason to analyze the conversational data was because we wanted to know which utterances would be most commonly used by the user, in order to guide future design and development of conversational navigation systems. We categorized the system’s instructions into groups because, as reported by Carletta el at, (1997), several researchers are beginning to try to code the dialogue corpora at a higher-level structure in the hope of providing more reliable findings. It would be useful for future development to know the basis of conversation for the map task.

5.5 Usability questionnaire analysis

The questionnaire consisted of four sections:

1. The user’s attitude about the system.
2. Effectiveness and efficiency of the system.
3. Learnability, usability and satisfaction of the user after using the system.
4. Future improvements. The full questionnaire in the Appendix D & E.

5.5.1 The user’s attitude questions analysis

P1.1. Have you ever used this kind of application before? (Yes/No formatting)

Most participants used the visual map interface (Group A). The conversational interface is a new approach, and many participants had never used it. Some of them had used a system that provided a conversational interface but not in a navigation system (e.g. SIRI function on iPhone, and Google Voice on Samsung).

P1.2. Do you think it is important to have this kind of function included in this application? (Yes/no and free text formatting)

Group A, 100% of participants think that the visual map interface was important. 10 participants wrote that the visual map was easy and not complicated to use for navigation. "It also shows how far to the destination, which is helpful when in unfamiliar places." Nine of them pointed out that using the visual map saved time for the user to find the place because they could know which direction they needed to go without guessing. Four participants liked using the visual map interface to guide them to the place.

Groups B, C, and D, all participants wrote that the conversational interfaces were important for navigation. Most of them gave reasons like "useful for low-tech people, who need specific information (e.g. blind and elderly)." They could communicate with the system without typing. Six participants wrote that the conversational interface was helpful for people new to the area (e.g. new students in the University).

P1.3. What do you like most about this application? (Free text formatting)

Group A, 15 participants liked visual instructions on the map, which has the names of buildings on the screen, as it was easier for the user to recognize the places. 10 participants said that the visual map interface is not complicated and was easy to use. A few of them liked the interface design of the system.

Group B, C, and D, most participants in these three groups liked the way the systems can be interactive and that participants can talk to the system to seek more information. They like the system's instructions, being not complicated, and providing specific information "I love how it is very detailed with exact places/landmarks. It tells me how far I need to walk. I love how I can ask it any time to tell me where to go again. I feel confident in asking it because the voice is calm and gentle". (participant) They pointed out that it is more accurate than the visual map. In Group D, about 10 participants liked seeing images on the screen and having them change while they were walking.

P1.4. What do you find most difficult about this application? (Free text formatting)

Group A: the first thing that more than half of participants mentioned was the system gave the wrong instructions, which made many participants get lost and frustrated. Four of them mentioned that the arrow display on the screen did not update and the map could not be rotated while the user was walking - they had to turn the position of the device by themselves. The last comment was about zoom-in and zoom-out which the system did not run automatically.

Group B, C, and D: the most difficult thing identified was the weather, which we have outlined previously. In addition, the sentences that the system could recognise were limited because people had different ways of speaking. There were about three participants who pointed out the microphone button on the screen, saying "it would have been better, if we did not need to press the button every time we talk to the system" (participant).

5.5.2 Effectiveness and Efficiency questions analysis (Likert formatting)

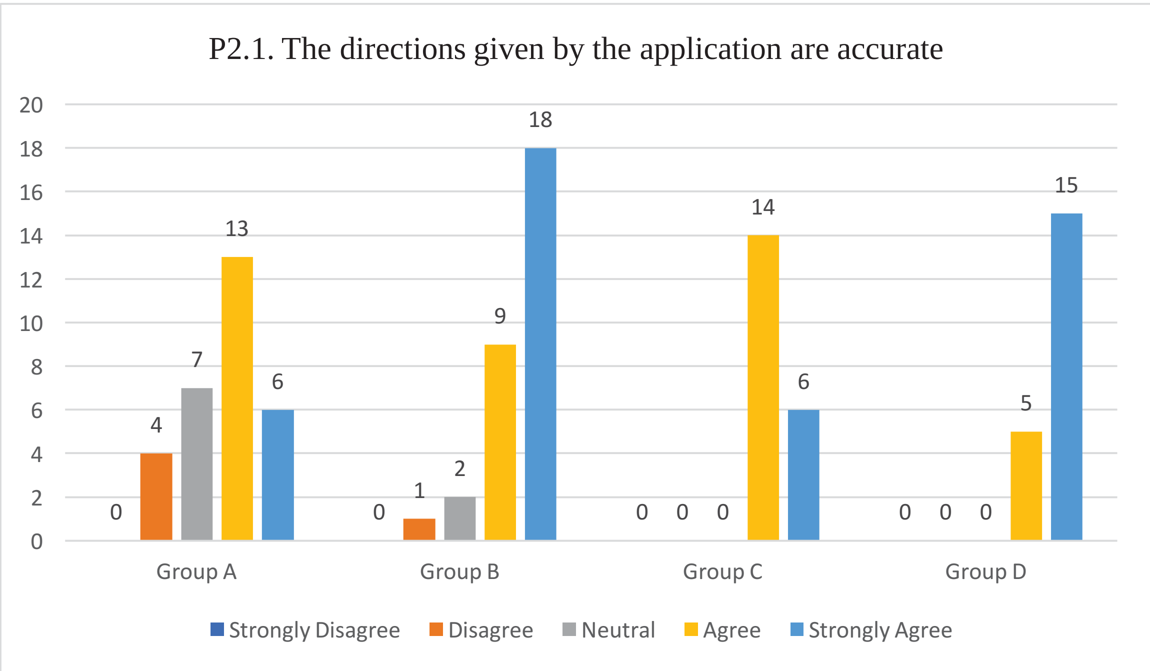


Figure 45 The results of P2.1

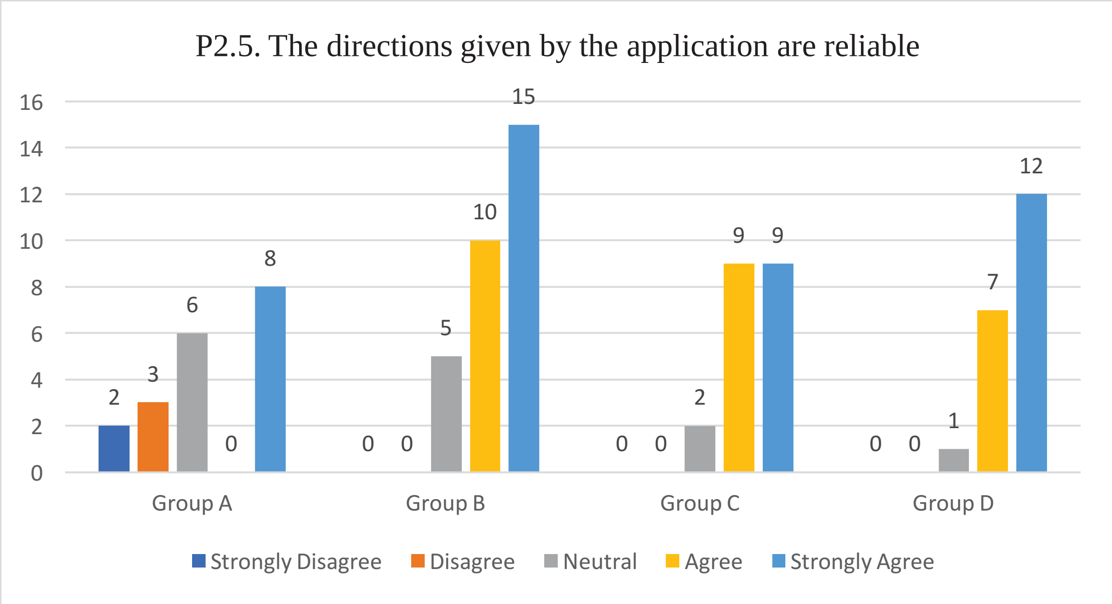


Figure 46 The results 0f P2.5

The results in Figures 45 and 46 show the users' perception of the accuracy and reliability of the application respectively. A higher proportion of participants who "Strongly agree" and "Agree" that the system that includes the conversational interface (Groups B (41.5%), C (43%), and D (49.5%)) is accurate and reliable than for the visual interface (OsmAnd). Only a few participants put "Disagree" and "Neutral" for Groups B, C, and D about reliability but still less than Group A which had quite high negative feedback from participants.

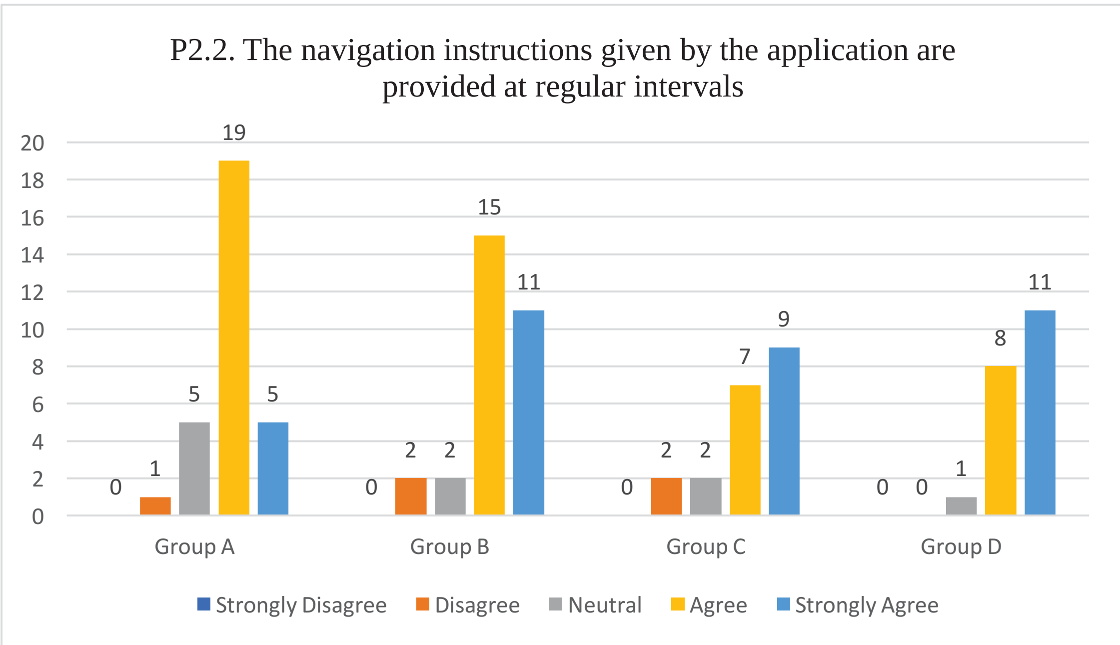


Figure 47 The results of P2.2

In Figure 47, 11, 9, and 11 participants in Groups B, C and D respectively, “Strongly Agree” that the conversational interface provided appropriate times to give instructions. The system did not give instructions too frequently or irregularly. As well, in Group A, 19 participants “Agree” and 5 “Strongly Agree” that intervals of instructions provided by the OsmAnd are sufficient.

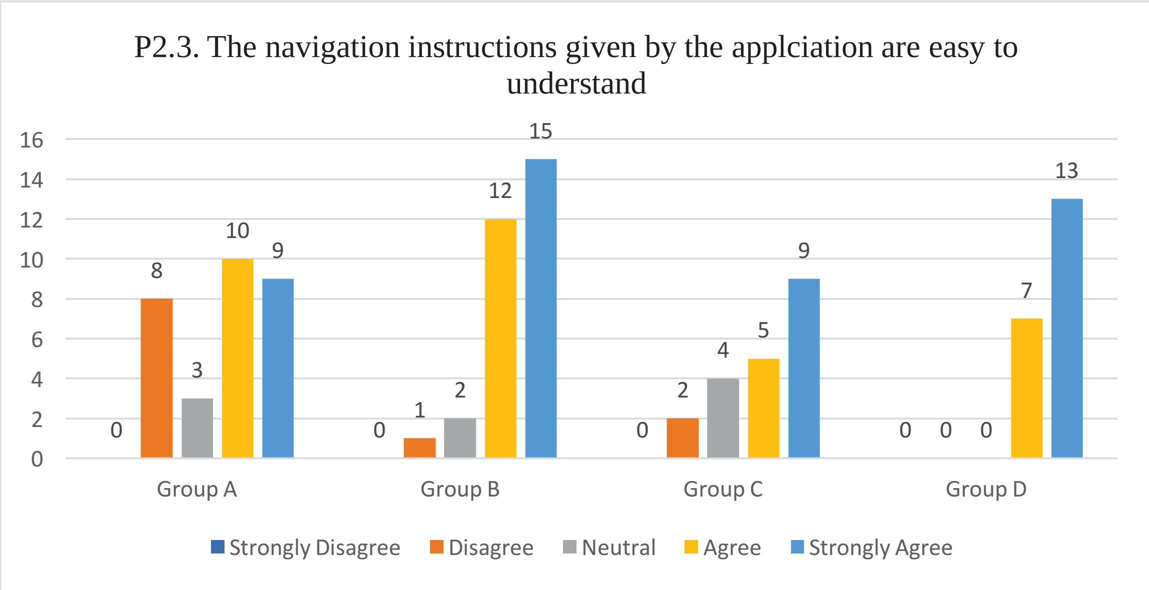


Figure 48 The results of P2.3

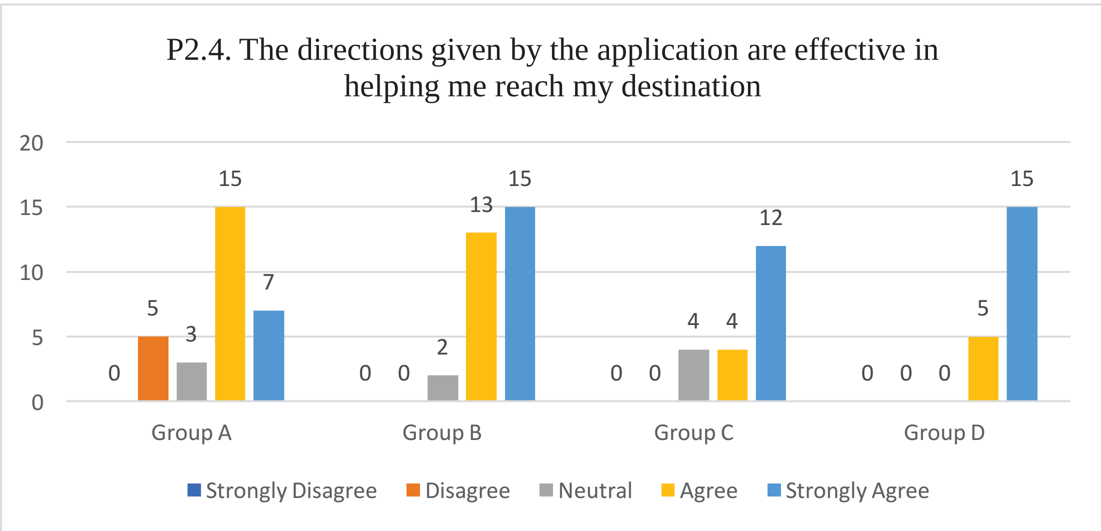


Figure 49 The results of P2.4

Figure 48 shows how clear and easy the instructions given by the system were to understand. Results show that eight participants in Group A ticked “Disagree” for this question.

This was because the OsmAnd application attributed to several participants getting lost and some verbal commands were incorrect. While most of participants who used the conversational interface (Groups B & D) “Strongly Agree” that the system has given simple and easy instructions, only a few in Group C gave negative answers for this question. Also, a higher number of participants in Groups B, and D “Strongly Agree” that the conversational interface was effective than the OsmAnd (Group A) in helping them to reach their destination (Figure 49).

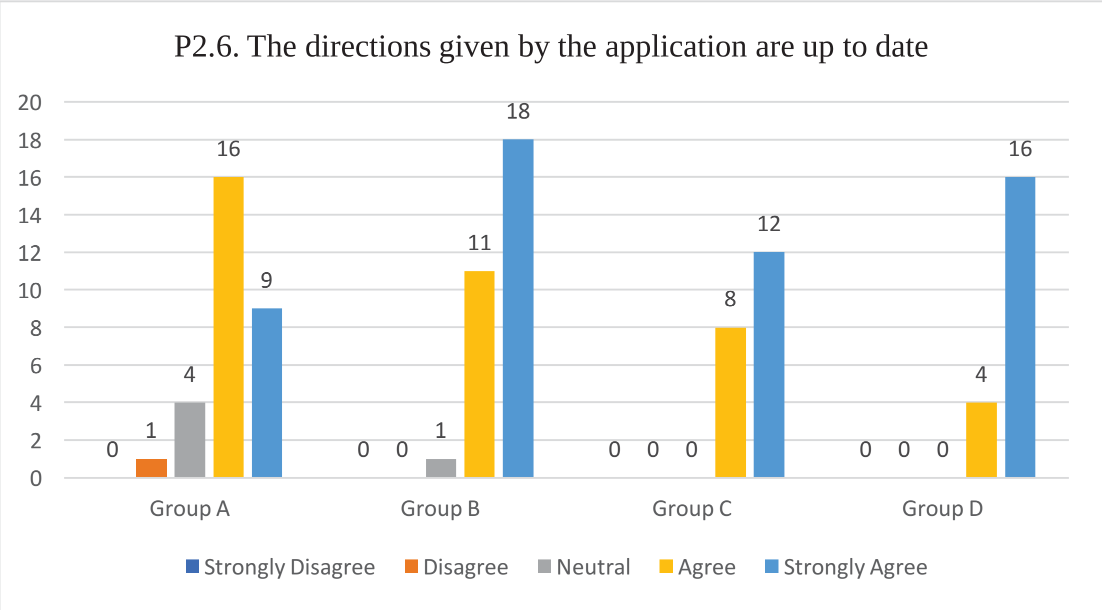


Figure 50 The results of P2.6

Figure 50 shows information about perceptions of direction currency in the OsmAnd and three conversational interfaces. The results show that most of participants in Groups B, C and D “Strongly agree” that the information within the conversational systems was up-to-date as they could see if pictures or other information from the system were correct and reliable. While Group A, who used the OsmAnd, ticked “Agree” with this statement, some of them were just “Neutral” as they gave comments that the system was not always correct in some locations both on the map and instructions.

The following five questions are especially for Groups B, C, and D focused on the Effectiveness and Efficiency of the conversational interface.

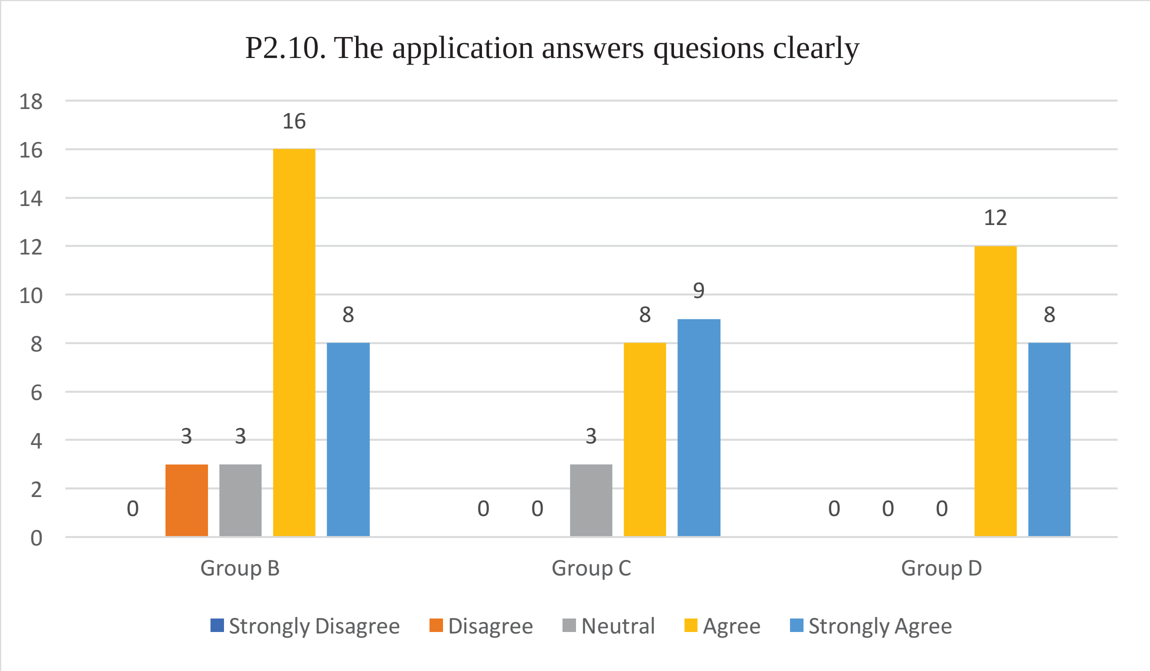


Figure 51 The results of P2.10

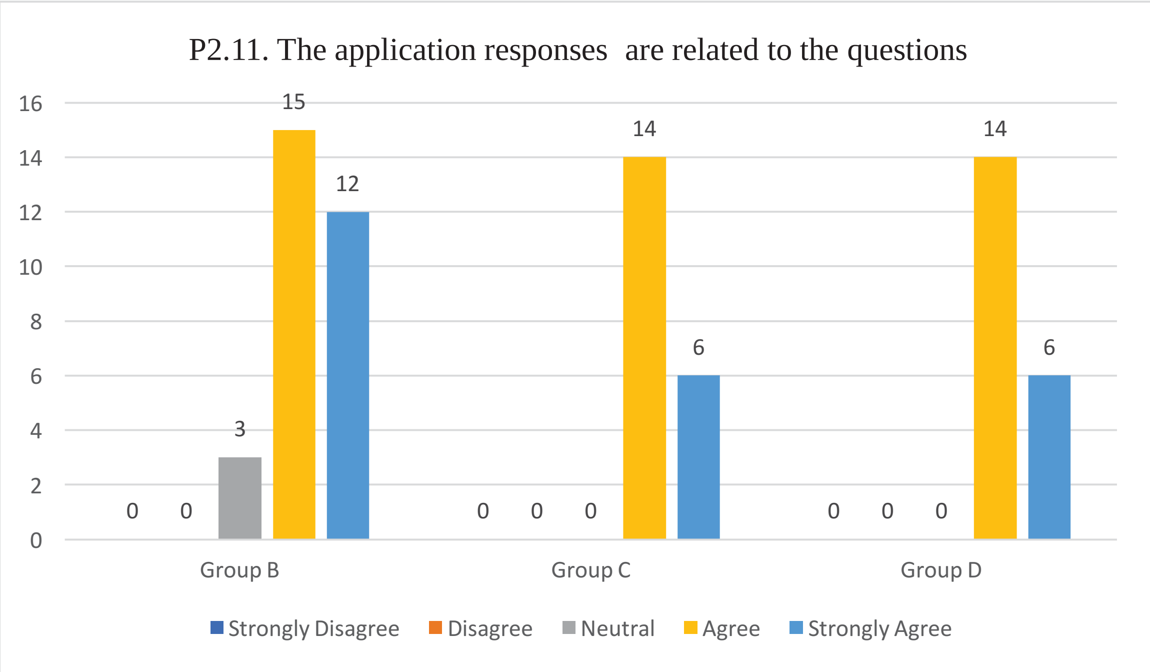


Figure 52 The results of P2.11

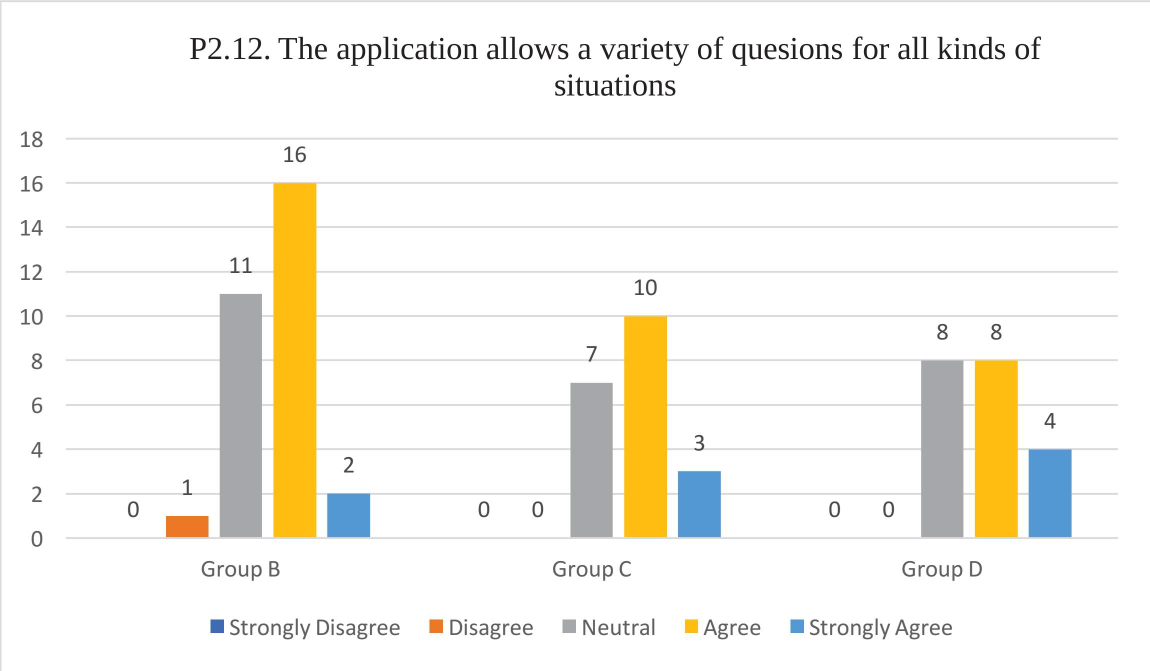


Figure 53 The results of P2.12

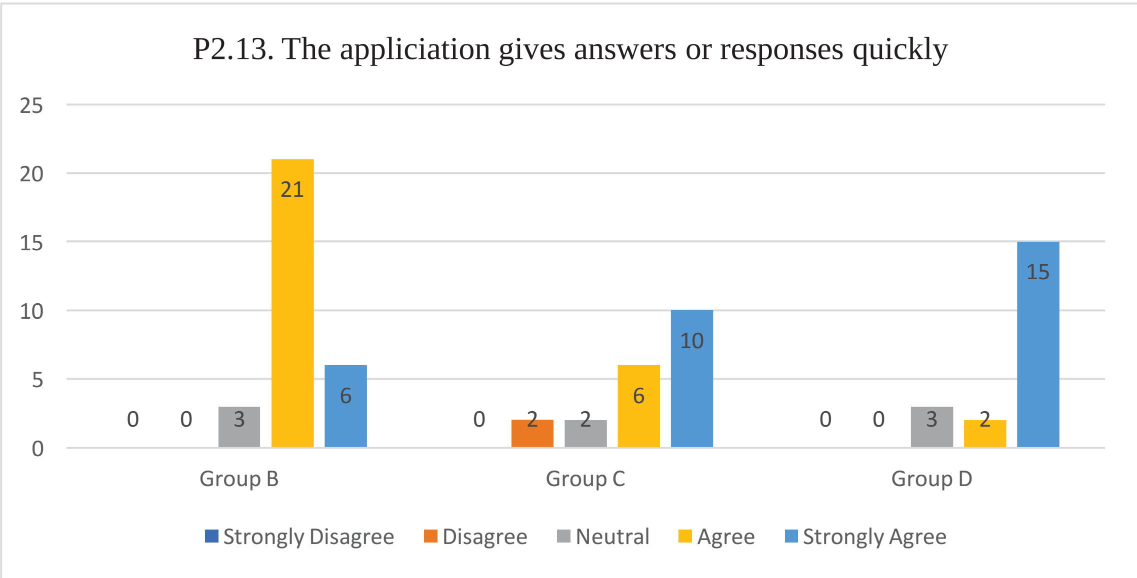


Figure 54 The results of P2.13

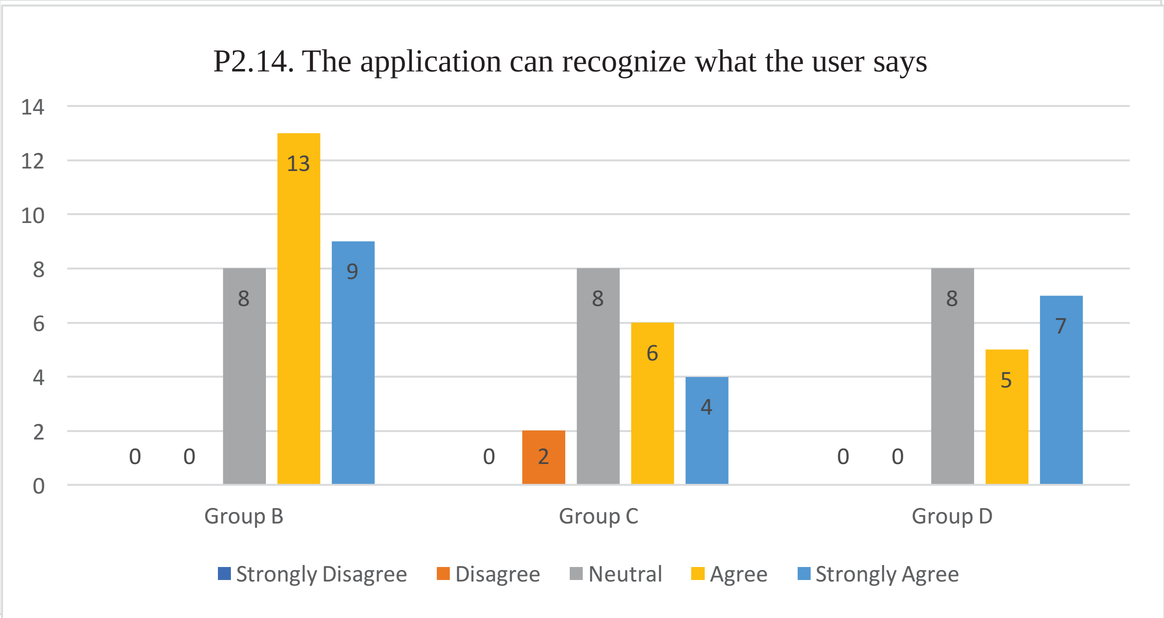


Figure 55 The results of P2.14

5.5.3 The average of USE questions analysis

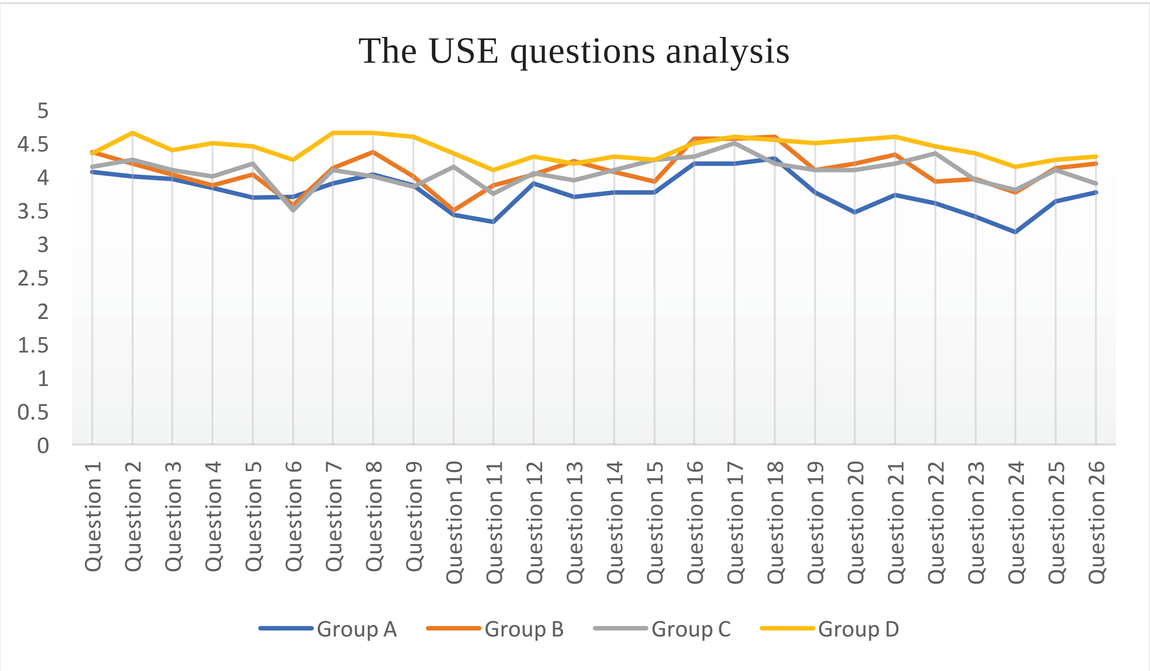


Figure 56 The average of USE questions analysis

Figure 56, shows the comparison of USE questions between Groups A, B, C and D. The results present that the average rate of Likert scale of Group A is lower than the other three groups which means that conversational interfaces, from participants' points of view preferred the OsmAnd interface. For example, on the Learnability question regarding the ease of the system (Question 7), Group D shows the highest scale "Strongly agree", while Groups B and C ticked "Agree" and Group A has a lower scale than the other groups. In Question 24, which asks about the necessity of having the system, Group A has lowest scale ("Neutral") compared to Groups B, C and D, which all ticked "Agree". As noticed for Question 6, Groups B and C have the lower rate compared to Group A, regarding the functionalities and capabilities of the system. This is because conversational interfaces in Groups B and C were not very attractive since participants could see only the microphone button and texts on the screen while Group D had the highest rate, which system displays images for all locations. In contrast, where questions 16 to 18 were about ease of using the system, Groups B and C jumped significantly higher than Group A. (The average of USE questions calculation can be seen in Appendix: I)

5.5.4 The average and max and min of the total across all questions for each map

1.) The map with command interface (Group A)

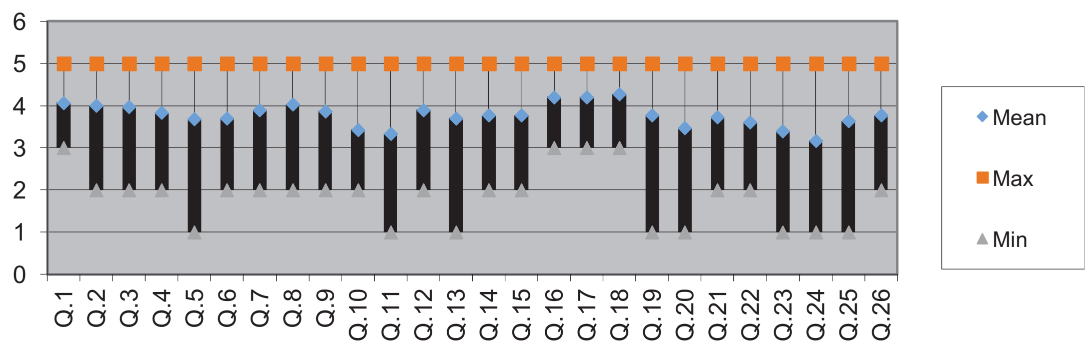


Figure 57 Mean and max and min of the total across all questions for Group A

2.) The conversational- only interface (Group B)

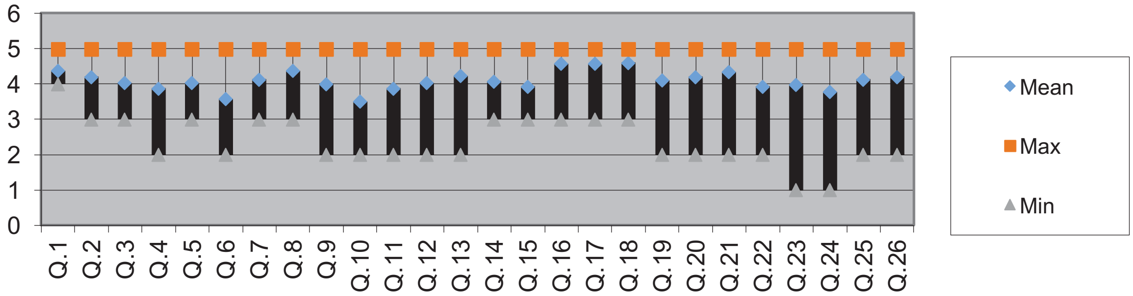


Figure 58 Mean and max and min of the total across all questions for Group B

3.) The conversational with map interface (Group C)

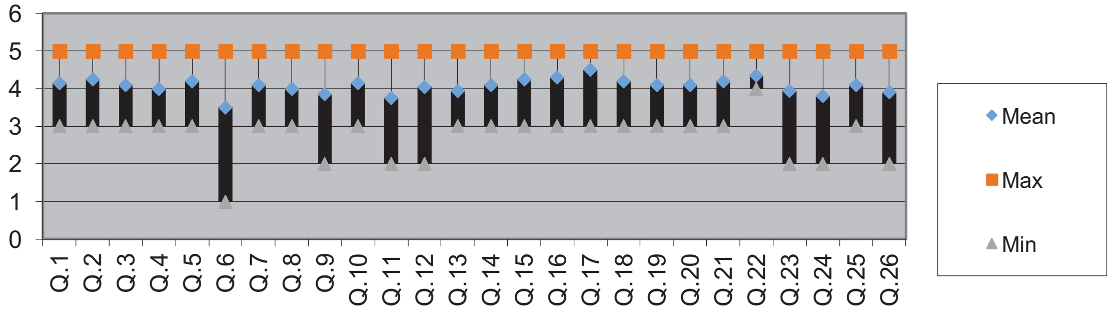


Figure 59 Mean and max and min of the total across all questions for Group C

4.) The conversational with image interface (Group D)

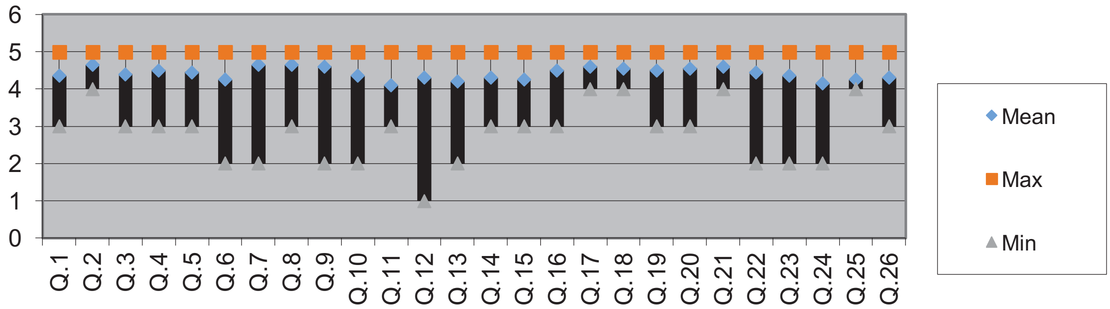


Figure 60 Mean and max and min of the total across all questions for Group D

These four graphs present Mean, Max and Min of the total across all questions for each map. It can be seen that three conversational interface groups (Group B, C and D) have higher averages than Group A especially in question 24 regarding the importance of having this kind of navigation system. Group A has an average of only three while other three groups of the conversational interface are higher at a four average.

5.5.4.1 The average response figures across all the questions for four interfaces by Microsoft Excel

Table 8 The average response figures across all the questions for four interfaces

Group	Average
Group A	3.82
Group B	4.13
Group C	4.14
Group D	***4.28**

(question 1-9: part2) and (question 1-26: part3)

5.5.4.2 The average response figures across all the questions for three conversational interfaces by Microsoft Excel

Table 9 The average response figures across all the questions for Three conversational interfaces

Group	Average
Group B	4.11
Group C	4.14
Group D	***4.28**

(question 1-14: part2) and (question 1-26: part3)

The two tables above present the average response figures across all the questions. None of these differences were statistically significant at level 0.05. Group D is highest number for both calculations, which means that the conversational with image is the most usable for the user regarding to all responses.

5.5.5 Future improvement questions analysis

In this section, questions for Group A and Groups B, C and D were slightly different because of the system's capabilities.

P4.1 *Do you have any suggestions about how to improve this application?*
(Free text formatting)

In group A, a large number of participants pointed out the improvement of voice instructions in the OsmAnd application. Obviously, some instructions were not reliable and need to be resolved as well as the interface design of the system, eight participants thought that it could have provided more information about the places instead of just their names. Some participants also mentioned the orientation of the map not rotating while walking which required them to turn it by themselves, which can be an annoying point for users of the OsmAnd application.

Groups B, C and D, gave more improvement suggestions. They wanted to see a variety of sentences to interact with the system. For example, some participants were asking the system "Am I in the right place?" (the check system) but the system could not recognize this sentence. Wind protection was the major problem during the experiment. Improving language options was also common feedback, "if the conversational interface could support more than English, it would be really ideal for all kinds of people." (participant). The last popular feedback from participants was about the check system, which was the system response showing notification to remind the user that they had almost arrived or they were off-track. They said it could have been better if the system could have alerted the user or they could ask the system to check their location. Another comment that came up in all groups was a speed reminder to record and remind the user what the speed should be (for driving).

P4.2 *Are there other kinds of functions you would like to be able to see in the application? (Group A)* (Free text formatting)

This question applied only to Group A. The result shows that Point of Interest (POI) was the first thing that 10 participants mentioned. This application had already been built in, but prior to the experiment this was removed as it was beyond the research purpose. The second popular feedback from this question was about the speed reminder for driving and the last common comment was about the visual graphics (e.g. 3D images and real images of the place).

The reason for applying this question for only Group A was the map with the command interface has very basic functions which all navigation systems should have and we wanted to know what new functions they would like to see in the future.

P4.2 *What is good about this application?* (Free text formatting)

About 20 participants in Group A said that the OsmAnd application was easy to use, especially as they could see the names of the buildings. They felt more secure than when only listening to voice commands. The setting location icon is not too difficult to find, even for low-tech people and the elderly.

Most participants in groups B, C and D were impressed with the interactive nature of the system. 50 out of 70) participants thought that the conversational interface was very easy to use and gave simple instructions. It suited all kinds of people (low-tech, high-tech, and disabled). They mentioned that the system made them feel as if they were communicating with a real person, rather than a robot. They felt comfortable talking to the system. Some said that it was good not to look at the screen all the time and not needing to type as with the traditional map application.

P4.3 *What is missing from this application? (Group B, C, and D)* (Free text formatting)

Visual appeal was a commonly mentioned; however, rather than wanting to see a map, they wanted to see 3D images or real images on the screen and 20 participants mentioned this (this was the reason for the creation of Group D during the research execution). A check system was another common suggestion, where the system could send notifications to the user to let them know about distance or the arrival duration or a sound to remind them they were back to the right place when they went off-track.

5.6 Interview analysis

This section presents analysis of the semi-structured interviews which took place after the task and the questionnaire were completed. This interview was conducted with 70 participants in Groups B, C and D to get more in-depth feedback on the conversational interface. The reason for this is that users will say much more verbally than they will in writing, and the researcher can ask them to clarify, tell more and explain things in more detail because questionnaires are quite limited in the information they provide.

1st question: Were there types of questions that you would have liked to ask, but you could not?

50 people said that all questions were good and simple enough for navigation while 13 people pointed out the absence of (1) the ability to ask for confirmation, including ascertaining whether the user is on the right track, allowing the user to confirm information that the system has given, but the user is not entirely sure about. Example include “Am I going to the right way?” and “Can you say that again?” This case is where the user could not understand the first time. (2) The ability to specify the names of locations to the system directly instead of typing. For example, “I want to go to …….” or “How can I get to…..”. 7 participants also identified the need for the point of interest question, where the user could ask about restaurants, petrol stations, hotels, or sightseeing locations nearby as well as a weather report allowing the user to ask about the daily weather “Any good restaurants around me?” or “where can I see a good view?”, “what is the weather today or tomorrow?”

Table 10 Types of questions created by participants

Types of questions	*Definition*	*Example*
The check question	The question that the user could ask while using the system whether to see they are on the right way and make sure that the user understands the system.	I. Am I going to the right way? II. How far from here to ….? III. Can you say that again? IV. Do you mean ….?

The set-up location question	The user could tell or ask the system about where to go and the name of their locations.	I. I want to go to …? II. How to go to ….? III. Do you know how to go to ….?
The POI question	The user could ask about points of interest to visit.	I. Where I can see a good view? II. Which place is nice for dinner? II. Where can I get petrol around here?
The weather question	The user could ask about the daily weather.	I. How is the weather today? II. Do I need to bring umbrella today? III. Will it rain tomorrow?

2nd question: Are there any other kinds of interactions you would like to see for future development?

Most participants mentioned the check function where the system could check the user's position by responding with an alert to keep the user on the right track. It was considered that this would be more secure for the user while using the system. "It would be better if the system had a function that will BEEP, shake or say something whenever I am straying away from the path." (participant) Another participant said, "This system would be perfect if it could alert when I have almost arrived at a main landmark or the destination." (participant). 6 participants mentioned that they would like to see the system report on the traffic on the road. For example, accidents, traffic congestion and road hazards.

3rd question: How do you feel about the conversational interface?

This question was not on the list but the researcher asked this question to all participants of Groups B, C and D after they arrived at the destination to try to gain their feelings for using this kind of navigation. Most of them had never seen this kind of navigation and they were very impressed. They liked the way the system responded, it was simple and easy to understand. Many of them suggested developing the system as an indoor navigation, a university navigation or a navigation for use in a shopping mall or small areas. About five participants said this system would be ideal for blind people as they can talk through their phones. Many researchers are now focusing on navigation systems for blind people (Loomis, Klatzky, Golledge, Cicinelli, Pellegrino, & Fry, 1993).

5.6.1 Summary

This section analyzed three questions from interviews of participants in Groups B, C and D to gain more information about the user's feeling when they used this kind of navigation system. It consisted of three questions:

1. types of questions that the user would like to ask. Most participants were satisfied with the types of questions that were in the system. Some of them mentioned the **check question** which is the question that users ask to check their positions or paths; the **set-up location question**, where the user can tell or ask the system about the location instead of typing on the screen; the **point of interest question**, where the system can understand and give information about places or landmarks when the user asks and the **weather question**, where the user can ask about the weather from the system.

2. These were about the interactions that the user would like to see in the future. Most participants also mentioned a check function where the system could alert the user whenever they are lost or getting close to a place, a turning point or the destination, as well as the traffic function to report to the user about accidents, road hazards and traffic jams. The last question was not on the list,

3. This question was about the impression of the conversational interface. Most participants were interested in this kind of navigation. They said it was very handy and more convenient than typing on the screen. All participants in groups B, C and D were happy and satisfied with the system's instructions. They also pointed out that "It is cool to communicate with the system rather than only listen to it, it is a very great idea and it could be ideal for the university, the shopping mall and in small areas." (participant)

Chapter 6

Discussion and recommendations

The purpose of this research was to compare the performance between the OsmAnd application and three types of conversational interface, to see whether the conversational interface was able to be easily used and useful for navigating. Prior studies have developed a navigation system by adopting speech recognition to improve its capability, however, these systems have only one-way interaction. That is the navigation systems have the capability to recognize the user's spoken words and transfer the user's dictation into text, then it displays directions, but the system is not able to talk back (Chou et al, 2006; Bingxuan, 2002). Thus, this research designed prototypes of conversational interface for navigation to compare with the map with command (OsmAnd) interface to find whether the conversational interface is effective, efficient and more usable than the conventional map interface.

The results revealed that, for the accuracy of the route, many participants in Group A who used the OsmAnd application got lost at two locations due to the system's instructions. Some participants were looking at the screen instead of listening to commands and they reached the destination without problems, even though they commented about an arrow on the screen being a bit delayed. However, the overall result has shown that participants Group D has spent the shortest average distance compares to those three kinds of navigation system. It is because they can see the real images of the locations clearly plus the system explained instruction in a simple way and easy to understand. It could make the user know exactly the next location without any confusion or hesitation.

The three other conversational interface groups walked to the destination without any confusion from the system's instructions. The major problem for these groups was the noisy environment (e.g. the windy noise), which affected the microphone of the system as it could not recognize the user's voice. Unwanted background noise that interrupted the desired speech signal caused a mismatch between the speech signal and the training data of the acoustic models of the system. This degraded the performance of the system (Dekens et al., 2010). This problem was not present in other types of map interface.

6.1 The problem with the noisy environment

Speech recognition in a construction setting is difficult because of the loud noise produced by wind and heavy machines. As reported in Moreno, (1996) main sources of distortion are the type of noise that occurs from things, parts of machines, or even some movements (e.g. a computer fan, lorry engine, door slams, radio, and other speakers).

There are some solutions that previous researches have undertaken and can be used in further research. According to Huang et al., (2001) one of the best ways to solve this problem is to train the acoustic model with data gathered from the operating environment. Dekens et al., (2010) They apply a throat microphone to help improve the performance of speech recognition. This is as a sensor which has the human spoken signal as an important part and is used with a Bluetooth close-talk microphone which helps to avoid the air signal and non-spoken segments. Another basic solution from participants during our experiment is that they used a mobile headphone which we noted could help the system recognize their speech easier and can be used in light wind.

6.2 The problem with OsmAnd's commands

We selected the OsmAnd application because it is open source and all data contributes to the OpenStreetMap database which we can add extra information into the application (e.g. the name of buildings in Massey University). During the experiment, many people got lost while using this application as outlined in the results section. They had trouble with the command system giving incorrect information and the green arrow on screen moves quite slowly. We found critical reviews about this application on Amazon.com and Google Play. They said that "Couldn't find my location for 30 minutes. The voice navigation for walking is equally bad as Google Maps and is unable to function as turn for turn (The user from Google Play)" and another "I really want to like this program but I don't find it user friendly or terribly accurate (The user from Amazon)".

As reported in Karimi, (2013) implementing the navigation system for walking is more difficult than other kind of navigations because of the level of detail required. In addition, the user generally uses footpaths which are sometimes not noted on the Global Navigation Satellite System (GNSS) signals, e.g. inside buildings, small roads, or in tree shadowed environment. A further concern is that sometimes the arrow on the screen is a little bit delayed when walking.

For example, the user arrives to place A but on the screen still displays the previous location that the user walked past.

For the time taken, the average time taken illustrates in each point. There were about 10 points which participants walked past. In Group A, participants took the quickest time to reach the destination while the other three groups spent longer. Participants in Group A spent the most time in points 8 to 9 (the bookshop to the 3rd stairs) because of the system's instructions. They spent about 2 mins finding their way back to the third stairs. Groups B, C and D, using the conversational interface, didn't suffer this problem and spent less time than Group A to walk past this point. In contrast, participants in Groups B, C and D suffered from the windy conditions that interrupted the performance of the system. They attempted to communicate with the system to get more information. The average minimum time also shows participants in group A spent only 3.30 mins, while Groups B, C and D spent 4.42, 3.33, and 4.33 mins respectively, to reach the destination. We noticed that conversational interface groups spent longer time than the map with command group because of problems from (1) the noisy environment (wind) (2) the user stopped in every landmark to communicate with the system, and sometimes, couldn't understand the user's utterances. The fastest time taken for the map interface is particularly significant because those participants travelled further on average. This it means they are moving quite a lot faster.

6.3 Recommendations

Further research can focus on;

1. Add the ability for users to ask confirmation questions (e.g. "Am I going to the right way?" and "Can you say that again?").

2. Support question regarding point of interest where the user could ask about restaurants, petrol stations, hotels, or sightseeing locations nearby as well as weather report questions "Any good restaurants around me?" or "where I can see a good view?", "what is the weather today or tomorrow?"

3. Support 3D images of real places or landmarks on the screen, that will build confidence for the user when they travel to unfamiliar places.

4. Warn users of digression while they are using the system, in order to make sure that the user is on the right track or almost arriving at their destinations. It could be shown as a BEEP, vibration or a spoken word (e.g. "you are getting off-track, please go back to the previous location") to alert the user.

5. Less press button, users can talk to the system without stopping and press the button every time when they would like to communicate. This could be used to talk about other things not just directions.

6. For the future research, it could be better if the training section will be extended for a longer time in order to make participants feel more familiar and confidence to use the system before the experiment in order to gain more efficiency results.

6.4 Answering the research questions

In this section, we go back to the research questions mentioned in section 1.3.

1.) **Q: Which kind of navigation interfaces are the best for pedestrians?**

A: The results we have from the analysis showed that both kinds of navigation interfaces are good for pedestrians in different situations. There are two points to mention:

1.) From the route analysis, the conversational interfaces obviously had clearer instructions than the OsmAnd. As a result, participants who used the conversational interfaces (Groups B, C and D) reached all key points, while some participants in Group A got lost. Moreover, the information in the OsmAnd was not always correct so it made participants in Group A confused and led them in wrong directions.

2.) From the time taken analysis, using The OsmAnd took less time than using the conversational interfaces (Table1). Because of the users' need to press the button each time they wanted to get instructions from the system, they normally stopped walking and asked. As well, many participants in the conversational interfaces groups suffered from the wind, which interfered with voice recognition. This was the main obstacle in using the conversational interfaces.

2.) **Q: Does a conversational interface allow users to navigate more effectively and efficiently than a more conventional map interface?**

A: All participants in Groups B, C, and D answered that the conversational interface for navigation system was very useful and could be more effective than the map and command system as it could inform the user of specific landmarks. This is especially helpful for disabled and low-tech people as users do not need to look at the screen all the time.

3.) **Q: Does a conversational interface provide better usability than a more conventional map interface?**

A: The results from the questionnaire and interview questions indicated the conversational interface provided better usability than the conventional map across a range of parameters (Appendix I). The conversational interface is more flexible to use than the conventional map interface as the user does not have to look at the screen, thus avoiding possible accidents. Some participants said that "If instructions in the conversational navigation are clear enough, the map is no longer necessary for navigation."

4.) **Q: How should a conversational navigation interface be designed to optimise user experience?**

A: Refer to section 6.3 for recommendations regarding design and development of a conversational user interface for navigation.

Chapter 7

Conclusion

In conclusion, we have presented the results of an evaluation of four kinds of navigation interfaces tested with 100 participants. The design and comparison looked at the following kinds of navigation interface:

- The map with command interface (OSMAND) (Group A)
- The conversational-only interface (Group B)
- The conversational with map interface (Group C)
- The conversational with image interface (Group D)

The goal was to identify and compare which kind of navigation interface is the most useful for pedestrians. The research shows that all three of the conversational interfaces were more accurate than the OsmAnd application. No participants got lost while using the conversational interfaces. Only one in the conversational with map interface group got lost around half-way from the starting point. The conversational interfaces have much simpler instructions and are easy to understand by the user. In contrast, many participants got lost while they were using the OSMAND because the instructions were unclear and incorrect. However, when we looked at the time taken of all kinds of navigation interface on average, the map with command interface (OSMAND) was quicker to use compared to all three of the conversational interface.

Overall, the conversational interface that allows the user to communicate with the system is more accurate than the normal navigation interface (OsmAnd) for pedestrian, because it can give more information in specific locations to the user and is much easy and simple for them. Results showed that the conversational with image was the most usable for the user.

Chapter 8

Reference

Adler, J L. (2001) Investigating the learning effects of route guidance and traffic advisories on route choice behaviour *Transportation Research Part C 9*(1), 1-14

Akaho, K., Nakagawa, T., Yamaguchi, Y., Kawai, K., Kato, H., & Nishida, S. (2012). Route guidance by a car navigation system based on augmented reality. *Electrical Engineering In Japan*, *180*(2), 43-54.

Api.ai. (2016). *API & Docs.* Retrieved from https://docs.api.ai/docs/key-concepts

Arentze, T., Feng, T., Robroeks, J., van Brakel, M., & Huibers, R. (2012). Compliance with and influence of a new in-car navigation system for trucks: Results of a field test. *Transport Policy*, *23*42-49. doi:10.1016/j.tranpol.2012.06.011

Asano, M. (2008). Smart Technology for Aging, Disability, and Independence: The State of the Science. *Physiotherapy Canada*, *60*(2), 193.

Battleson, B., Booth, A., & Weintrop, J. (2001). Usability testing of an academic library web site: a case study. *The Journal of Academic Librarianship*, *27(3),* 188e198.

Begampure, S., Garad, D., Patwardhan, K., & Sarve, P. (2016). Navigation System for the Blind. *Imperial Journal of Interdisciplinary Research*, *2*(6).

Bingxuan, 等. (2002). Speech Recognition Technology Applied to Intelligent Mobile Navigation System. *地球空间信息科学学报：英文版 / Geo-Spatial Information Science*, *(4)*, 37.

Bowles, H., & Seedhouse, P. (2007). *Conversation analysis and language for specific purposes* (63). Peter Lang.

Brainline.org. (n.d.) *Speech Recognition for Learning.* Retrieved from http://www.brainline.org/content/2010/12/speech-recognition-for-learning_pageall.html

Brett F. (2008) *The Top Five Uses of Speech Recognition Technology*. Retrieved from https://www.callcentrehelper.com/the-top-five-uses-of-speech-recognition-technology-1536.htm

Britta, O, B. (2015). *5 best navigation apps: Helping you get around as quick as possible.* Retrieved from http://www.cmswire.com/cms/web-engagement/the-what-why-and-how-of-usability-testing-007152.php

Bruce, C. Edmundson, A., Coleman, M. (2003). Writing with voice: an investigation of the use of a voice recognition system as a writing aid for a man with aphasia. *International Journal of Language & Communication Disorders*, *38(2)*, 131-148.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Cardoso, J., Sheth, A., & Miller, J. (2003). Workflow quality of service. In *Enterprise Inter- and Intra-Organizational Integration* (pp. 303-311). Springer US.

Carletta, J., Isard, S., Doherty-Sneddon, G., Isard, A., Kowtko, J. C., & Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Computational linguistics*, *23*(1), 13-31.

Cavallo, F., Sabatini, A., & Genovese, V. (2005). A step toward GPS/INS personal navigation systems: Real-time assessment of gait by foot inertial sensing. *2005 IEEE/RSJ International Conference On Intelligent Robots And Systems, IROS*, (2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS), 109-113. doi:10.1109/IROS.2005.1544967

Carvalho, S., & White, H. (1997). *Combining the quantitative and qualitative approaches to poverty measurement and analysis: the practice and the potential* (Vol. 23). World Bank Publications.

Chou, T. S., Ku, I., Wu, C. H., Hsu, L. C., Lin, Y. S., Chen, Y. T., & King, C. T. (2006). CaNPAs: a campus navigation and parking assistant system. In *2006 IEEE International Conference on Systems, Man and Cybernetics*, (1), 631-638.

Colman, S. (2016). Voice recognition evidence. *Journal of Criminal Law*, (1). 5.

Darves, C., Oviatt, S., & Coulston, R. (2002). Adaptation of users' spoken dialogue patterns in a conversational interface. In *Proceedings of the 7th International Conference on Spoken Language Processing* (1), 561-564.

Dekens, T., Verhelst, W., Capman, F., & Beaugendre, F. (2010). Improved speech recognition in noisy environments by using a throat microphone for accurate voicing detection. In *Signal Processing Conference, 2010 18^th^*, 1978-1982. IEEE.

Dellea, P. (2011) Talking to yourself: Using voice recognition software. *Journal of Nurse Life Care Planning, 11(1),* 321-328. 1942-4469

Dembczyński, K., Gaweł, P., Jaszkiewicz, A., Kotłowski, W., Kubiak, M., Susmaga, R., & ... Zielniewicz, P. (2012). Community Traffic: a technology for the next generation car navigation. *Control & Cybernetics, 41*(4), 867-883.

D'Este, G. M., Zito, R., & Taylor, M. A. (1999). Using GPS to measure traffic system performance. *Computer-Aided Civil and Infrastructure Engineering*, *14*(4), 255-265.

Dharun, V. S., & Karnan, M. (2012). Voice and Speech Recognition for Tamil Words and Numerals. *International Journal of Modern Engineering Research (IJMER) Vol, 2,* 3406-3414.

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education*, *40*(4), 314-321.

Ding, J. W., Wang, C. F., Meng, F. H., & Wu, T. Y. (2010). Real-time vehicle route guidance using vehicle-to-vehicle communication. *IET communications*, *4*(7), 870-883.

Dodson, A. H., Moon, G. V., Moore, T., & Jones, D. (1999). Guiding blind pedestrians with a personal navigation system. *Journal of Navigation*, *52*(03), 330-341.

Dong, W. (2011). An overview of in-vehicle route guidance system. In *Australasian Transport Research Forum.*

Dubberly, H., & Pangaro, P. (2009). What is conversation? How can we design for effective conversation. *Interactions Magazine*, XVI, 4, 22-28.

Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*. Intellect books.

Du, J., & Aultman-Hall, L. (2007). Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues. *Transportation Research Part A: Policy and Practice*, *41*(3), 220-232.

Elizabeth, L. (2013) *New Navigation Technology Predicts Traffic Conditions*. Retrieved from http://www.voanews.com/a/traffic-technology-clearpath/1616682.html

Geutner, P., Denecke, M., Meier, U., Westphal, M., & Waibel, A. (1998). Conversational speech systems for on-board car navigation and assistance. In *ICSLP* ,98, 14447-1450).

Goodman, R., Meltzer, H., & Bailey, V. (2003). The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *International Review Of Psychiatry*, *15*(1/2), 173.

Helal, A., Moore, S. E., & Ramachandran, B. (2001). Drishti: An integrated navigation system for visually impaired and disabled. In *Wearable Computers, 2001. Proceedings. Fifth International Symposium on*, 149-156. IEEE.

Higgins, E.L., & Raskind, M.H. (2004). Speech recognition-based and automaticity programs to help students with severe reading and spelling problems. *Annals of Dyslexia, 54(2),* 365-392.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & cognition*, *13*(3), 208-217.

Huang, X., Acero, A., Hon, H. W., & Foreword By-Reddy, R. (2001). *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice hall PTR.

Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis*. Polity.

Ian, P. (2015). *Google Maps adds offline turn-by-turn directions, search.* Retrieved from http://www.pcworld.com/article/3003913/apps/google-maps-adds-offline-turn-by-turn-directions-search.html

Jacob, R., Shalaik, B., Winstanley, A. C., & Mooney, P. (2011). Haptic feedback for passengers using public transport. In *International Conference on Digital Information and Communication Technology and Its Applications*, 24-32. Springer Berlin Heidelberg.

Kaparias, I. (2008) Reliable dynamic in-vehicle navigation. *Ph.D. Dissertation* Imperial College London

Karimi, H. A., Jiang, M., & Zhu, R. (2013). Pedestrian navigation services: challenges and current trends. *Geomatica, 67*(4), 259-271.

Kawamura, K. (2000). Perceived value of time for truck operators. *Transportation Research Record: Journal of the Transportation Research Board*, (1725), 31-36.

Kawasaki, H., Murao M., Ikeuchi K., and Sakauchi M. (2001). “Enhanced Navigation System with Real Images and Real- Time Information.” Proceedings of the 8th world congress on Intelligent Transport Systems (ITSWC ’01), Sydney, Australia, September 30, 221–228.

Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval, 3*(1—2), 1-224.

Kitchin, R., & Blades, M. (2002). *The cognition of geographic space* (Vol. 4). Ib Tauris.

Lewis, J. R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction, 7(1),* 57-78.

Lin, J. S. C. and Hsieh, P. (2006). The Role of Technology Readi- ness in Customers’Perception and Adoption of Self-Service Technologies. *International Journal of Service Industry Management 17 (5)*, 497–517.

Lin, W., & Da Young, J. (2015). Concurrent use of an in-vehicle navigation system and smartphone navigation application. *Social Behavior & Personality: An International Journal*, *43*(10), 1629-1639.

Lohr, S, & Markoff, J. (2010). *Smarter than you think: Computers learn to listen, and some talk back.* New York Times. Retrieved from http://www.nytimes.com/2010/06/25/science/25voice.html?ref=science.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *Journal of Experimental Psychology: General*, *122*(1), 73.

Longprasert, N. (2015). *Data Collection and Analysis for a Conversational Interface for Pedestrian Navigation.* Unpublished manuscript, Department of Information Technology, Massey University, Auckland, New Zealand.

Lund, A. M. (2001). Measuring Usability with the USE Questionnaire12.

Margaret R. (2007) *Voice recognition (speech recognition).* Retrieved from http://searchcrm.techtarget.com/definition/voice-recognition

Marisa, P. (2010). *The What, Why and How of Usability Testing.* Retrieved from http://www.cmswire.com/cms/web-engagement/the-what-why-and-how-of-usability-testing-007152.php

May, A. J., Ross, T., Bayer, S. H., & Tarkiainen, M. J. (2003). Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing*, *7*(6), 331-338.

McTear, M., Callejas Carrion, Z., & Griol Barres, D. (2016). *The conversational interface. [electronic resource]: talking to smart devices*. Switzerland : Springer, [2016].

Medenica, Z., Kun, A. L., Paek, T., & Palinko, O. (2011). Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, 265-274.

Medenica, Z., Kun, A., Paek, T., & Palinko, O. (2010). Comparing Augmented Reality and Street View Navigation.

Michon, P. E., & Denis, M. (2001). When and why are visual landmarks used in giving directions?. In *Spatial information theory*, 292-305.

Minami, N., & Madnick, S. (2010). *Using Systems Analysis to Improve Traffic Safety* (Vol. 4). Working Paper CISL 2010.

Moreno, P. J. (1996). *Speech recognition in noisy environments* (Doctoral dissertation, Carnegie Mellon University Pittsburgh).

Ortiz, C. L. (2014). The Road to Natural Conversational Speech Interfaces. *IEEE Internet Computing*, *18*(2), 74-78. doi:10.1109/MIC.2014.36

Park, E., & Kim, K. (2014). Driver acceptance of car navigation systems: integration of locational accuracy, processing speed, and service and display quality with technology acceptance model. *Personal & Ubiquitous Computing*, *18*(3), 503-513. doi:10.1007/s00779-013-0670-2

Park, E., Kim, H., & Ohm, J. Y. (2015). Understanding driver adoption of car navigation systems using the extended technology acceptance model. *Behaviour & Information Technology*, *34*(7), 741-751.

Park, E., Kim, K. J., & Del Pobil, A. P. (2013). An examination of psychological factors affecting drivers' perceptions and attitudes toward car navigation systems. In *IT Convergence and Security 2012* (555-562). Springer Netherlands.

Park, E., Kim, K. J., Jin, D., & del Pobil, A. P. (2012). Towards a successful mobile map service: an empirical examination of technology acceptance model. In *International Conference on Networked Digital Technologies* (420-428). Springer Berlin Heidelberg.

Peck, T. C., Fuchs, H., & Whitton, M. C. (2012). The design and evaluation of a large-scale real-walking locomotion interface. *IEEE transactions on visualization and computer graphics*, *18*(7), 1053-1067.

Pérez-Marín, D., & Pascual-Nieto, I. (2013). An exploratory study on how children interact with pedagogic conversational agents. *Behaviour & Information Technology*, *32*(9), 955-964.

Pouteau, X., & Arévalo, L. (1998). Robust spoken dialogue systems for consumer products: a concrete application. In *ICSLP*.

Rehrl, K., Häusler, E., Steinmann, R., Leitinger, S., Bell, D., & Weber, M. (2012). Pedestrian navigation with augmented reality, voice and digital map: Results from a field study assessing performance and user experience. In *Advances in Location-Based Services*, 3-20. Springer Berlin Heidelberg.

Rose India (n.d.) *Advantages and Disadvantages Global Positioning System.* Retrieved from http://www.roseindia.net/services/trackingsystem/advantaesanddisadvantagesofgps.shtml

Silver-Pacuilla, H. (2006). Access and benefits: Assistive technology in adult literacy. *Journal of Adolescent and Adult Literacy 50(2)*,114-25.

Skog, I., & Handel, P. (2009). In-Car Positioning and Navigation Technologies-A

Survey. *Ieee Transactions On Intelligent Transportation Systems*, *10*(1), 4-21.

Smith, A. (2015). US smartphone use in 2015. *Pew Research Center*, *1*.

Strothotte, T., Petrie, H., Johnson, V., & Reichert, L. (1995). MoBIC: user needs and preliminary design for a mobility aid for blind and elderly. In *The European Context for Assistive Technology: Proceedings of the 2nd TIDE Congress, 26-28 April 1995, Paris* (Vol. 1, p. 348). IOS Press.

Tarkiainen, M., Kauvo, K., Kaikkonen, J., & Heine, H. (2001). Simple turn-by-turn route guidance for pedestrians. In *8th World Congress on Intelligent Transport Systems*.

Turk, M., & Robertson, G. (2000). Perceptual user interfaces (introduction). *Communications of the ACM*, *43*(3), 32-34.

Van Drongelen, M. (2015). *Android Studio Cookbook*. Packt Publishing Ltd.

Vehicle-to-vehicle communication *IET Communications 4*(7), 870-883

Vonk, T., Van Rooijen, T., Hogema, J., & Feenstra, P. (2007). Do navigation systems improve traffic safety. *Report TNO*.

Yankelovich, N., Levow, G. A., & Marx, M. (1995). Designing SpeechActs: Issues in speech user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 369-376.

Yannakoudakis, E. J., & Hutton, P. J. (1987) *Speech Synthesis and Recognition Systems, Ellis Horwood Limited*. Chichester, UK,

Yoshino, K., & Kawahara, T. (2015). Conversational system for information navigation based on POMDP with user focus tracking. *Computer Speech & Language*, (1), 275.

Zapata, B. C. (2015). *Android Studio Essentials*. Packt Publishing Ltd.

Zue, V., Seneff, S., Glass, J. R., Polifroni, J., Pao, C., Hazen, T. J., & Hetherington, L. (2000). JUPlTER: a telephone-based conversational interface for weather information. *IEEE Transactions on speech and audio processing*, *8*(1), 85-96.

Zue, V. W., & Glass, J. R. (2000). Conversational interfaces: Advances and challenges. *Proceedings of the IEEE*, *88*(8), 1166-1180.

Chapter 9

Appendices

Appendix A: Lists of example questions

The research gives a list of example question to the user to see before do the experiment.

1. Start asking directions

Tell me the directions, please
Where I can go?
Where should I go?
Give me the directions
Give me the directions, please
Can you tell me, please?
Can you tell me where to go?
Can you please give me the directions

2. Confirm position

Tell me more
Where to go next?
Ok, I'm here, where to go next? Ok, I'm here
I'm here, then what?
I'm at (the place), what's next?
I already here
I can see it now
Yes, I'm here now

**** These are only examples of sentences; you can try out other sentences as well**

Appendix B: User instructions for the OsmAnd application

Instructions for using the OsmAnd application

1st Open the OsmAnd application

2nd choose “Direction” which is the first icon. The arrow will be pointing at our destination,

3rd Change to pedestrian mode 

4th press “Go” and start following directions until you get to the place. 

Appendix C: The consent form

This form uses after the researcher introduced about the research and experiment.

Informed Consent

- The objective of the experiment, task and actives involved have been explained to me.
- I understand that I may withdraw from the experiment at any time during the experimental session.
- I understand that the data collected will be stored and used for research only
- I understand that I may request that any data I submit during the experiment may be destroyed at my request or removed by me.
- I understand and agree with the procedures which guarantee my anonymity with respect to data gathered by questionnaires
- I understand that my participation will be recorded through audio and screen capture.

I agree to take part in this research study (tick where relevant)

Yes	No

Signed:

....

Print Name:

....

Date:

....

Participant Details

Please complete the following details. Your responses will be kept confidential and will be used solely for research purposes.

Thank you for your co-operation.

Part1. General questions

1. What is your age?

❍ 18-28	❍ 62-72
❍ 29-39	❍ >73
❍ 40-50	
❍ 51-61	

2 What is your gender?

❍ Male

❍ Female

3. Are you a native English speaker?

❍ Yes

❍ No

4. What is your native language?

....

4. What position are you in Massey University?

❍ Student

❍ Staff

❍ Visitor

Part2. Mobile devices

1. Do you use a smartphone?

❍ Yes

❍ No

2. What type of mobile phone do you use?

❍ Android

❍ IOS

❍ Symbian OS

❍ Don't know

Part3. Navigation systems

1. Have you ever used an app in your phone to help you find directions using maps, or to find your current location? (If yes, go to 2, otherwise, return this document to the researcher)

❍ Yes

❍ No

❍ I don't know

2. How often do you use map navigation apps?

❍ Frequently

❍ Occasionally

❍ Never

3. Do you think map navigation applications are useful?

❍ Useful

❍ Not useful, why ……………………………………………………..…..

4. Are there any new features that you think would be useful in mapping applications that are not currently available?

…………………………………………………………………………………………………

…………………………………………………………………………………………………

Appendix D: The questionnaire Group A

Questionnaire

Part1: The following questions are about your opinions of using the application, please tick as appropriate and complete with as much information as you can provide

P1.1. Have you ever used this application before?

- ❍ Yes, I have been used this application
- ❍ No, I've never used this application

P1.2. Do you think it is important to have this kind of function included in this application?

- ❍ Yes, why? ……………………………………………………………………
- ❍ No, why?……………………………………………………………………

P1.3. What do you like most about this application?

……………………………………………………………………………………………………
……………………………………………………………………………………………………
…………………………………………………………

P1.4. What do you find most difficult about this application?

……………………………………………………………………………………………………
……………………………………………………………………………………………………
…………………………………………………………

Part 2 & 3: Please rate the following statements that represent how you feel about interactions with the application you used.

P2.1. The directions given by the application are accurate.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P2.2. The navigation instructions given by the application are provided regular intervals.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P2.3 The navigation instructions given by the application are easy to understand.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P2.4. The directions given by the application are effective in helping me reach my destination.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P2.5. The directions given by the application are reliable.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P2.6. The directions given by the application are up to date.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P2.7. The application provides the shortest route to the user.

Strongly disagree — Disagree — Neutral — Agree — Strongly agree

P2.8. The application shows current locations correctly.

Strongly disagree — Disagree — Neutral — Agree — Strongly agree

P2.9. This application is easier to use than other navigation applications.

Strongly disagree — Disagree — Neutral — Agree — Strongly agree

Part 3:

P3.1. The application helps me to achieve my goals.

Strongly disagree — Disagree — Neutral — Agree — Strongly agree

P3.2. The application is useful.

Strongly disagree — Disagree — Neutral — Agree — Strongly agree

P3.3. The application makes the things I want to accomplish easier to get done.

Strongly disagree — Disagree — Neutral — Agree — Strongly agree

P3.4. The application saves me time when I use it.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.5. The application meets my needs.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.6. The application has all the functions and capabilities I expect it to have.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.7. The application is easy to use.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.8. The application is simple to use.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.9. The application is user-friendly.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.10. The application is flexible.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.11. Using this application is effortless.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.12. I can use the application without written instructions.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.13. Both occasional and regular users would like it.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.14. I can recover from mistakes quickly and easily.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.15. I can use it successfully every time.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.16. I learned to use it quickly.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.17. I easily remember how to use it.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.18. It is easy to learn to use it.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.19. I am satisfied with it.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.20. I would recommend it to a friend.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.21. It is fun to use.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.22. It works the way I want it to work.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.23. It is fantastic.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.24. I feel I need to have it.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.25. It is pleasant to use.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.26. It feels comfortable using this application.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

Part 4: Please give us your feedback

P4.1. Do you have any suggestions about how to improve this application?

....

....

P4.2. Are there other kinds of functions you would like to be able to see in the application?

....

....

P4.3. What is good about this application?

....

....

Appendix E: the questionnaire Group B, C, and D

Questionnaire

Part1: The following questions are about your opinions of using the application, please tick as appropriate and complete with as much information as you can provide

P1.1. Have you ever used this kind of application before?

- ❍ Yes, I have been used this application
- ❍ No, I've never used this application

P1.2. Do you think it is important to have this kind of function included in this application?

- ❍ Yes, why? ……………………………………………………………………
- ❍ No, why?………………………………………………………………………

P1.3. What do you like most about this application?

……………………………………………………………………………………………………

……………………………………………………………………………………………………

………………………………………………………………

P1.4. What do you find most difficult about this application?

……………………………………………………………………………………………………

……………………………………………………………………………………………………

………………………………………………………………

Part 2 & 3: Please rate the following statements that represent how you feel about interaction with the application you used.

P2.1. The directions given by the application are accurate.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P2.2. The navigation instructions given by the application are provided regular intervals.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P2.3 The navigation instructions given by the application are easy to understand.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P2.4. The directions given by the application are effective in helping me reach my destination.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P2.5. The directions given by the application are reliable.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P2.6. The directions given by the application are up to date.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P2.7. The application provides the shortest route to the user.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.8. The application shows current locations correctly.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.9. This application is easier to use than other navigation applications.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.10. The application answers questions clearly.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.11. The application responses are related to the questions.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.12. The application allows a variety of questions for all kinds of situations.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.13. The application gives answers or responses quickly.

Strongly disagree Disagree Neutral Agree Strongly agree

P2.14. The system can recognize what the user says

Strongly disagree Disagree Neutral Agree Strongly agree

Part 3:

P3.1. The application helps me to achieve my goals

Strongly disagree Disagree Neutral Agree Strongly agree

P3.2. The application is useful.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.3. The application makes the things I want to accomplish easier to get done.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.4. The application saves me time when I use it.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.5. The application meets my needs.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.6. The application has all the functions and capabilities I expect it to have.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.7. The application is easy to use.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.8. The application is simple to use.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.9. The application is user friendly.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.10. The application is flexible.

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.11. Using this application is effortless.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P3.12. I can use the application without written instructions.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P3.13. Both occasional and regular users would like it.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P3.14. I can recover from mistakes quickly and easily.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P3.15. I can use it successfully every time.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P3.16. I learned to use it quickly.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
○	○	○	○	○

P3.17. I easily remember how to use it.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.18. It is easy to learn to use it.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.19. I am satisfied with it.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.20. I would recommend it to a friend.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.21. It is fun to use.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.22. It works the way I want it to work.

Strongly disagree Disagree Neutral Agree Strongly agree

P3.23. It is fantastic.

○—○—○—○—○

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.24. I feel I need to have it.

○—○—○—○—○

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.25. It is pleasant to use.

○—○—○—○—○

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

P3.26. It feels comfortable using this application.

○—○—○—○—○

Strongly disagree | Disagree | Neutral | Agree | Strongly agree

Part 4: Please give us your feedback

P4.1. Do you have any suggestions about how to improve this application?

………………………………………………………………………………………………

P4.2. What is good about this application?

………………………………………………………………………………………………

P4.3. What is missing from this application?

………………………………………………………………………………………………

Appendix F: Interview questions

For only participants in Group B, C, D. Interview ran after they finished the experiment and questionnaire.

The following questions were discussed during the interview.

1. Were there types of questions that you would like to (or) have been able to ask, but you couldn't?
2. Are there any other kinds of interactions you would like to see for future development?

Appendix G: The structure of the experiment

1. A researcher introduces themselves to participants

"Hello, my name is Eve. I am doing an experiment in conversational interfaces for navigation for pedestrians**."**

2. The explanation of the research objective

"The objective of this research is to find out what is the most useful conversational interface for navigation systems."

3. The explanation of tasks

"We will start from the Chicken Wings, then I will give you some time to play with the application then you can start to use the system to guide to the destination. What you will need to do is :

1st fill a consent form before doing the task.

2nd start doing the task that I have explained.

3rd complete a usability questionnaire."

4. Filling a usability questionnaire

"After you have finished your tasks, then complete the last part by filling a questionnaire to give us your feedback, and I will give you a gift voucher for thanks."

5. Confirmations from participants

"Are you happy to do it? :D

If **Yes**

"That is cool! Let's complete the consent form first, and after you're done, we can start doing the task. Before we start, I will give you a guideline on how to use the application and you can play with the application for about 5 minutes, then I will follow you and record your interactions with my phone until you reach the destination, but please behave as naturally as possible and I will not be able to help or answer any questions"

If **No**

"That is ok, thanks."

6. Explanations for each of group

Group A

"I provide brief instructions for the OsmAnd application for you to have a look at and play around with the application for one minute. When you are ready, tell me."

Group B

"This is the conversational interface. Once you press the 'Microphone' icon, you can start asking questions. This is an example of the questions that you could ask the system. You also can think of other sentences. Continue asking questions naturally until you get to the destination. At this stage, I have set the destination already. What you need to do is ask questions to the system and follow its instructions."

Group C

"You will be testing this navigation interface to guide you to somewhere within our campus. What you need to do is ask questions to the system and follow instructions to the destination. I will give you some examples of questions that you could ask or say to the system. If you get confused or get lost by the system's instructions, there is a small icon at the top. Once you press this, you will see the map icon, so you can press this and it will change to the map interface, so you will be able to check where you are."

Group D

"You are going to test this navigation interface to guide you to the place in this area and these are examples of sentences that you could say to the system. You have to ask and follow instructions from the system until you reach the destination. You will also see pictures of landmarks that the system mentions once you get close to them"

Appendix H: The brochure

Do you ever get lost?
Come and help us to test a new way to navigate!

Come and participate in our research!!!

get A Westfield $ 20 GIFT CARD for FREE!

No experience needed.

If you are

1. Massey students, staffs or visitors.
2. Willing to participate in tasks to test our apps (20-30 mins) at the big gold Chicken Wings (between Quadrangle Building A and Quadrangle Building B) at Massey University (Albany campus).

Please contact me if you are interested.! **021- 083- 74594, Longprasert.n@hotmail.com (Eve)**

"This project has been evaluated by peer review and judged to be low risk. Consequently it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named in this document are responsible for the ethical conduct of this research.
If you have any concerns about the conduct of this research that you want to raise with someone other than the researcher(s), please contact Dr Brian Finch, Director (Research Ethics), email humanethics@massey.ac.nz. "

Appendix I: The average of USE questions calculation by SPSS

1. Group A

Descriptive Statistics

	N	Mean	Std. Deviation		Variance
	Statistic	Statistic	Std. Error	Statistic	Statistic
P3.1	30	4.07	.117	.640	.409
P3.2	30	4.00	.152	.830	.690
P3.3	30	3.97	.162	.890	.792
P3.4	30	3.83	.152	.834	.695
P3.5	29	3.69	.173	.930	.865
P3.6	30	3.70	.167	.915	.838
P3.7	30	3.90	.154	.845	.714
P3.8	30	4.03	.148	.809	.654
P3.9	30	3.87	.190	1.042	1.085
P3.10	30	3.43	.171	.935	.875
P3.11	30	3.33	.205	1.124	1.264
P3.12	30	3.90	.147	.803	.645
P3.13	30	3.70	.174	.952	.907
P3.14	30	3.77	.157	.858	.737
P3.15	30	3.77	.177	.971	.944
P3.16	30	4.20	.111	.610	.372
P3.17	30	4.20	.111	.610	.372
P3.18	30	4.27	.106	.583	.340
P3.19	30	3.77	.196	1.073	1.151
P3.20	30	3.47	.202	1.106	1.223
P3.21	30	3.73	.166	.907	.823
P3.22	30	3.60	.183	1.003	1.007
P3.23	30	3.40	.201	1.102	1.214
P3.24	30	3.17	.204	1.117	1.247
P3.25	30	3.63	.182	.999	.999
P3.26	30	3.77	.149	.817	.668
Valid N (listwise)	29				

2. Group B

Descriptive Statistics	N	Mean	Std. Deviation		Variance
	Statistic	Statistic	Std. Error	Statistic	Statistic
P3.1	30	4.37	.089	.490	.240
P3.2	30	4.20	.130	.714	.510
P3.3	30	4.03	.122	.669	.447
P3.4	30	3.87	.150	.819	.671
P3.5	30	4.03	.102	.556	.309
P3.6	30	3.57	.157	.858	.737
P3.7	30	4.13	.115	.629	.395
P3.8	30	4.37	.102	.556	.309
P3.9	30	4.00	.152	.830	.690
P3.10	30	3.50	.133	.731	.534
P3.11	30	3.87	.142	.776	.602
P3.12	30	4.03	.148	.809	.654
P3.13	30	4.23	.133	.728	.530
P3.14	30	4.07	.106	.583	.340
P3.15	30	3.93	.106	.583	.340
P3.16	30	4.57	.104	.568	.323
P3.17	30	4.57	.114	.626	.392
P3.18	30	4.60	.103	.563	.317
P3.19	30	4.10	.139	.759	.576
P3.20	30	4.20	.147	.805	.648
P3.21	30	4.33	.138	.758	.575
P3.22	30	3.93	.135	.740	.547
P3.23	30	3.97	.169	.928	.861
P3.24	30	3.77	.177	.971	.944
P3.25	30	4.13	.150	.819	.671
P3.26	30	4.20	.162	.887	.786
Valid N (listwise)	30				

3. Group C

Descriptive Statistics					
	N	Mean	Std. Deviation		Variance
	Statistic	Statistic	Std. Error	Statistic	Statistic
P3.1	20	4.15	.167	.745	.555
P3.2	20	4.25	.123	.550	.303
P3.3	20	4.10	.191	.852	.726
P3.4	20	4.00	.205	.918	.842
P3.5	20	4.20	.138	.616	.379
P3.6	20	3.50	.336	1.504	2.263
P3.7	20	4.10	.216	.968	.937
P3.8	20	4.00	.178	.795	.632
P3.9	20	3.85	.254	1.137	1.292
P3.10	20	4.15	.167	.745	.555
P3.11	20	3.75	.260	1.164	1.355
P3.12	20	4.05	.256	1.146	1.313
P3.13	20	3.95	.198	.887	.787
P3.14	20	4.10	.191	.852	.726
P3.15	20	4.25	.143	.639	.408
P3.16	20	4.30	.179	.801	.642
P3.17	20	4.50	.154	.688	.474
P3.18	20	4.20	.172	.768	.589
P3.19	20	4.10	.161	.718	.516
P3.20	20	4.10	.161	.718	.516
P3.21	20	4.20	.172	.768	.589
P3.22	20	4.35	.109	.489	.239
P3.23	20	3.95	.223	.999	.997
P3.24	20	3.80	.225	1.005	1.011
P3.25	20	4.10	.161	.718	.516
P3.26	20	3.90	.261	1.165	1.358
Valid N (listwise)	20				

4. Group D

Descriptive Statistics					
	N	Mean		Std. Deviation	Variance
	Statistic	Statistic	Std. Error	Statistic	Statistic
P3.1	20	4.35	.131	.587	.345
P3.2	20	4.65	.109	.489	.239
P3.3	20	4.40	.152	.681	.463
P3.4	20	4.50	.154	.688	.474
P3.5	20	4.45	.135	.605	.366
P3.6	20	4.25	.123	.550	.303
P3.7	20	4.65	.131	.587	.345
P3.8	20	4.65	.131	.587	.345
P3.9	20	4.60	.112	.503	.253
P3.10	20	4.35	.167	.745	.555
P3.11	20	4.10	.143	.641	.411
P3.12	20	4.30	.128	.571	.326
P3.13	20	4.20	.156	.696	.484
P3.14	20	4.30	.128	.571	.326
P3.15	20	4.25	.143	.639	.408
P3.16	20	4.50	.136	.607	.368
P3.17	20	4.60	.112	.503	.253
P3.18	20	4.55	.114	.510	.261
P3.19	20	4.50	.115	.513	.263
P3.20	20	4.55	.114	.510	.261
P3.21	20	4.60	.112	.503	.253
P3.22	20	4.45	.135	.605	.366
P3.23	20	4.35	.167	.745	.555
P3.24	20	4.15	.150	.671	.450
P3.25	20	4.25	.099	.444	.197
P3.26	20	4.30	.128	.571	.326
Valid N (listwise)	20				

Appendix J: Categories of the user’s utterances

This is lists of the user’s utterances that they used to speak to the system during the experiment.

Initiating moves	Response moves	Ready moves
Can you give me the directions?	Tell me more	Ok, I’m here
Tell me the directions, please	Where should I go next?	Yes, I’m here now
Give me the directions, please	I’m here, then what?	Ok, I can see it now
Where should I go?	I’m at (the place), what else?	Ok, I arrived
Can you tell me where to go?	What is next?	I’m already here
	Then what?	I’m here, thanks
	Yeah and then?	
	I have no idea where I’m going	