Recommendations on a Test Infrastructure for Evaluation of Touchscreen Assistive Technology for Visually Impaired Users

Berglind Smaradottir, Jarle Håland, Santiago Martinez, Åsmund Rodvig Somdal, Rune Fensli

*Department of Information and Communication Technology, University of Agder, Norway
†Department of Health and Nursing Science, University of Agder, Norway
‡Department of Psychosocial Health, University of Agder, Norway

Abstract

Mobile technologies’ touchscreen allows the use of choreography of gestures to interact with the user interface. Relevant aspects in mobile technology design become crucial when targeting users with disabilities. For instance, when assistive technology is designed to support speech interaction between visually impaired users and a system, accessibility and ease-of-use of such technology should be included in the usability and technical evaluation of their effectiveness. This paper presents the analysis of the technical and physical infrastructure of a controlled laboratory environment for user evaluations made in the research project “Visually impaired users touching the screen - A user evaluation of assistive technology” where VoiceOver, a screen reader in Apple Inc. products was tested. The paper reports on challenges related to the use of the test infrastructure, such as how to obtain valuable data when interactive high-speed gestures are performed and how to optimise the recording and synchronisation between audio and video data. The lessons learned by the research group showed that there are effective alternatives for each challenge, and these should be customised for each particular test, type of participants and device.

Keywords:

Conclusion, visually impaired users, speech-assisted navigation, touch gestures, accessibility, assistive technology, laboratory infrastructure, usability, health informatics

Introduction

Mobile technology is used today in people’s life [1][2][3] for information and communication purposes. Mobile technologies usually incorporate touchscreen for the interaction between the user and device’s interface. Touchscreen technologies [4][5] allow users to interact with a system through touch gestures made with their fingers. However, this type of interaction becomes a challenge for visually impaired users who cannot see the screen with sufficient detail to distinguish interface dimensions, elements inside the interface and buttons without tactile feedback [6]. Globally, the number of people with visual impairment is estimated to be 285 million. The main impairment causes are uncorrected refractive errors, such as myopia, hyperopia or astigmatism, and cataracts. 39 million people are estimated to be blind because of cataracts [7][8]. The International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) provides a categorisation for visual impairments: normal vision, moderate visual impairment, severe visual impairment and blindness [9]. WHO estimates that about 65% of visually impaired people are older than 50 years and 90% live in developing countries [7].

In order to improve the accessibility and the interaction with the user interface, several solutions of assistive technology are available in the market for visually impaired users [10][11]. In this context, the research project “Visually impaired users touching the screen - A user evaluation of assistive technology”, aimed to evaluate the interaction of visually impaired users using VoiceOver, a built-in screen reader in Apple Inc. products (provided by default since April 2005, Mac OS X 10.4) that allows users to interact with the user interface (UI) through gesture-based (since June 2009, iPhone 3GS OS 3.0) speech-assisted navigation. One of the major aspects of the evaluation of touchscreen assistive technology is how accessible the UI is for users with and without visual impairments. For an optimal gathering of test data, a physical and technical infrastructure is essential to support a multiple visual and audio perspective for data collection of such interaction. The collected data will form the basis of a retrospective analysis where touch interaction details observed in the recordings can be coupled with comments and observations obtained during the test. It is relevant to note that because users are visually impaired, the touch gestures will be only seen by the researchers, and therefore a slow pace observation of them is necessary after the test to build up a meaningful analysis of the interaction. Another key requirement of a mobile device with assistive technology is the usability of the system. Considering the sensory limitations of the target user group, the assistive technology should be intuitive, with an optimal presentation of the information facilitating a general understanding of the functionality and distribution of the UI.

This paper presents the challenges related to the testing of touchscreen assistive technology from the perspective of how the technical aspects of a laboratory infrastructure can be used in an Information and Communication Technologies (ICT) and Health Informatics research environment. It reports on the lessons learned by the research group exploring how to effectively carry out accessibility and usability evaluations of the mobile applications and technologies used in the research project.

The research questions (RQs) of this study were:
**RQ1:** What technical infrastructure is suitable for evaluation of touchscreen assistive technology with disabled users?

**RQ2:** What are the learned lessons transferable for testing other mobile technologies?

Following this introduction, an overview of related research is presented. Analysis of the use of the technical and physical test infrastructure for user evaluations of touchscreen assistive technology and reflections on lessons learned during the project are presented in the next sections. Later, the discussion section highlights the benefits of having an optimal infrastructure for the type of the evaluation carried out. Finally, the conclusions regarding the characteristics of a technical infrastructure for accessibility and usability evaluations of touchscreen assistive technology are drawn.

**Related Research/Background**

Assistive technology [12][13][14] includes devices or technological solutions that assist people with disabilities. Assistive technology is used as an alternative way of performing actions or interactions with technology. The accessibility [15][16] of a technology refers to how accessible a technology is regardless of user’s ability. Leporini et al. [17] investigated the interaction between Apple touchscreen devices with pre-installed VoiceOver screen reader through a usability inspection of the UI and an online survey with feedback from 55 blind users. They found that VoiceOver made the devices more accessible, but operations such as writing long text took too long or were uncomfortable for users. McGookin et al. [6] presented a study with 12 visually impaired participants operating two different touchscreen-based MP3 players. They found that participants could generally use the devices but they encountered problems in doing short time operations. They evaluated the touchscreen accessibility and provided guidelines for touchscreen technology design for visually impaired users. Phillips and Zao [18] did a study on user acceptance of assistive technology. They found that almost 30% of assistive devices were rejected by the users. Factors such as device performance, procurement and user need played an important role because they were related to the acceptance of technology. They concluded that involving users and focusing on their long-term needs would enhance user satisfaction. Demers et al. [19][20] described the development of a clinical instrument for evaluation of user satisfaction with assistive technology devices. They described several variables used to help user assess and rate the degree of satisfaction with assistive technology in a structured way. Svanæs et al. [21] presented a study on mobile ICT in clinical settings. They showed that the design of the graphical user interface (GUI) affects usability, ergonomic and social aspects. They concluded that usability tests of mobile ICT should be performed in a simulation environment with a high level of realism. Further, they stated that usability testing of mobile ICT for healthcare requires new ways of designing, recording and analysing the data collected.

**Test Infrastructure**

In order to test the infrastructure for evaluation of touchscreen accessibility, 6 visually impaired users participated in a study where they individually performed representative tasks related to gesture’s performance and task solving using the screen reader VoiceOver.

**The Research Group**

The evaluation research team consisted of three members with multidisciplinary background: one member with experience from teaching and supporting visually impaired students with assistive technology; the other two members with professional experience in health, ICT and human-computer interaction (HCI). All had professional experience in working with visually impaired people. One team member was the moderator in all the tests. In addition, an external senior researcher advised regarding planning and execution of the research study. A technician provided technical expertise and was available in case of need for assistance during the tests.

**Test Environment Infrastructure**

The evaluation of mobile assistive technology was held in the Usability Laboratory at the Centre for eHealth and Healthcare Technology of the University of Agder, Norway. The Usability Laboratory had two rooms: the Test room and the Observation room, connected through one-way mirror (visualisation from the Observation room towards the Test room). The complete infrastructure is described in details in [22].

The technical infrastructure for the usability evaluation is illustrated in Figure 1. The moderator and participant were in the Test room, while the other two members of the research team were in the Observation room. The moderator sat down on a table in the middle of the Test room with the participant besides. The elements used in the room were a smartphone, a task list, a table microphone and a tablet for additional sound-recording. The participant had the smartphone in their hands. The room had 2 IP cameras, 1 fixed and 1 portable with an external microphone. The Observation room had a desktop PC connected to three monitors. The observers followed the evaluation, remotely controlled the zooming of the fixed camera and made recordings and annotations of the test sessions.

The Observation room and the Test room were connected with a dedicated segment of the LAN infrastructure of the Centre for eHealth and Healthcare Technology, making use of VLAN technology. This connection was used for the IP-based streaming of video and audio signals from the Test room to the Observation room, using Wirecast [23] as capture and encoding software.

**Materials**

The material used during the study is presented below grouped by rooms for reproducibility and information purposes.

**Test room:**
- Apple Inc. iPhone 4 MD128B/A iOS 7.1.2 with VoiceOver activated.
- Fixed Camera: SONY BRCZ330 HD 1/3 1CMOS P/T/Z 18x Optical Zoom (72x with Digital Zoom) Colour Video Camera.
- Portable Camera: SONY HXR-NX30 Series.
- Apple Inc. iPad MD543KN/A iOS 8.1 for additional sound-recording.
- Sennheiser e912 Condenser Boundary Microphone.
- Landline phone communication
Observation room:
- Stationary PC: HP Z220 CMT Workstation, Intel Core i7-3770. CPU@3.4 GHZ, 24GB RAM, Windows 7 Professional SP1 64 bit.
  Monitor: 3x HP Compaq LA2405x
- Streaming: 2x Teradek RX Cube-455 TCP/IP 1080p H.264.
- Software Wirecast 4.3.1.
- Landline phone communication.

Data Collection
The test sessions were audio-visually recorded in the F4v video file format, exported to the Windows Media Video (WMV) format and then imported from QSR NVIVO 10 [24]. The recordings from two independent audio-visual sources were merged into one video file using the software Wirecast v.4.3.1, with multiple video perspectives and a single audio channel. In addition, annotations were made by the evaluation team during the test. After the evaluation, all recordings were transcribed verbatim and divided into categories for a qualitative content analysis [25]. The data collection of the study was approved by the Norwegian Social Science Data Services (NSD) [26] with the project number 40636.

Lessons Learned from the Test
This section presents the challenges and lessons learned about the technical infrastructure in the laboratory through the evaluation of touchscreen mobile assistive technology.

Optimisation of the Test Environment
Before the start of each test session, participants were asked to sit in a natural and relaxed position with the mobile phone in their hands. The cameras were then adjusted for optimal recording of the screen and hand gestures. The remote controlled camera zoomed on the mobile interface, visualised in full screen on one of the PC monitors in the Observation room. The portable camera was placed near the participant’s side. In general, both cameras were slightly angled from above to record the interaction and provide the best possible shot of mobile user interface and participant’s hands.

Moderator’s View
The moderator was sitting beside the test participant to guide them through the tasks on the smartphone. Two factors negatively influenced the accurate observation of the interaction between participant and the device: the mobile device’s small-size screen and the high speed of gestures.

In order to improve the moderator’s view and allow the possibility of following the actions of the participant and screen response on-line, a screen capture tool (e.g., software Mirroring 360 [27], Apple Airplay [28]) could be used to show the screen interface on a larger external screen in the Test room. The screen interface could be simultaneously recorded by a screen recording program (e.g., software Snagit [29]). In order to observe and record the finger interaction and the system’s response a screen capture tool (e.g., UX Recorder [30]) would also allow detecting, in time, when the hand interaction touches the interface. To closely observe gesture choreography, one common alternative in mobile usability testing is to place a macro-focused camera on the mobile phone to record user’s hand gestures. Its signal could also be displayed on an external screen in the Test room if necessary. However, its suitability for testing visually impaired users has not been yet tested by the researchers.

Clarity of Screen Reader Sound for Moderator
In the Test room, the moderator had in some cases difficulty to adequately listen the feedback from the VoiceOver, even when the settings were at maximum volume for the screen reader.

In order to improve the sound quality, Bluetooth or dedicated software such as Mirroring 360 could be used for transmission of sound to an external loudspeaker in the room. The use of external loudspeaker could increase the perception of sound for the moderator. However, this would create a new different setting for a test participant that would not directly hear the sound as usual from the mobile device, but instead from an external loudspeaker.

Effective communication between research team members was essential to perform on time any readjustment of equipment or task necessary during the test. The landline phone communication was available between the two rooms and used when the test was being recorded and none of the researchers could leave the Observation room. In order to improve the communication, an ear plug to connect moderator...
to observers watching the recordings would allow instant 1-way communication to do the adjustments without interrupting the test session.

Quality Optimisation of the Recordings

A high level of quality of the recordings is generally recommended for an optimal retrospective analysis of data in usability studies. The audio-visual recordings in the usability evaluation had a F4v video file format and were converted to the WMV format to be imported into the qualitative analysis program QSR NVIVO 10, used for watching and transcription purposes. Several factors associated with the quality of video and sound were identified that influenced the analysis in detail of the actions performed by participants during task execution. They are next described in 4 subcategories: visual improvements, sound improvements, video and sound synchronisation and storage.

Visual Improvements

In the Test room, the light source was directed down to the floor. Some footage showed glares that impeded the correct view of the mobile interface during the analysis. An alternative would be to have a light source directed to the walls of the Test room instead of directly down to the floor. In addition, a dimmer device could be used to reduce the brightness of the light sources that produced the glare. The Test room had one remotely controlled camera and another that was controlled manually. An advantage would be to also have the second camera remotely controlled for adjusting the angle and the zooming in case of glare or unexpected movement by a participant. Participant’s gestures were usually performed at high speed. This impeded the ability to accurately distinguish the finger gesture several times when retrospectively analysing the video at normal speed. In those cases, instead of using QSR NVIVO 10 that only allowed reducing up to 50% of the speed, the software Cyberlink [31] was chosen to show the footage even at lower speed, down to 20%.

Sound Improvements

In the recordings, in spite of the fact of having one wired microphone placed on the table and another on the external camera, the quality of sound reception was not sufficient at times. When testing mobile assistive technology, it should be taken into consideration that the VoiceOver of the smartphone gives a speech feedback that may interfere with other sounds listened during the test, e.g., participant’s answers or comments. For instance, there were up to three other sound sources (i.e., moderator’s voice, participant’s voice, smartphone’s VoiceOver speech) recorded simultaneously in several occasions. Recording overlapped sound sources obstructed the accurate perception of the sound during the analysis phase. It would be then advised to try to implement the policy of speaking one at time during the test, even though the VoiceOver could interfere at any point. A wireless microphone worn by participant and moderator would increase sound reception quality in addition to a stable sound source place nearby. This would remove the constraint of placing the participant beside the table microphone and allow them to freely move around.

In the case of insufficient quality of sound recordings, an additional sound recording during the session is recommended as a backup. In the usability evaluation, a tablet device was used as backup for sound recording; very useful when sound recordings from the main sources were not optimal. To improve the sound during the analysis, the VLC media player [32] was used to adjust frequencies of sound.

Video and Sound Synchronisation

When analyzing the recordings, video and sound signals were not perfectly synchronized, with a delay of the video signal of approximately 0.5 s. regarding the audio. This was probably due to the network latency added to the video signal streaming. This issue that may seem generally unimportant, is however especially relevant when the study includes rapid movements and actions of high order of magnitude. A potential solution could be to record all sources separately with digital audio workstation software (e.g., ProTools by Avid [33]) and transfer them to an editing program (e.g., Cyberlink [30], Final Cut Pro X [34] or Adobe Premiere Pro CC [35]). In such programs, the synchronisation can be adjusted frame by frame. This software also allows discretionary switching between the different video and sound recordings and zooming. However, substantial technical knowledge is required for the correct use of these digital audio edition programs. Due to the network latency, data transmission through direct wire is usually better than streaming. A FireWire cable [36] could be used for high-speed and synchronous real-time data transfer; this also would separate the storage into different files.

Storage

In order to reduce the risk of data loss, a redundancy in the data collection system is advisable. During the test sessions, one incident resulted in 10 minutes of footage loss due to a recording software error. In that case, the portable camera provided an additional recording that made the analysis possible without repeating the task. Test repetitions should be avoided when possible, because of the risk of biasing the data collection when repeating the same task and the inherent difficulty of recruiting visually impaired participants. An additional solution would be to record the data gathered in two independent hard disk drives from two different computers. This alternative solution has been implemented into the technical infrastructure of the laboratory after the incident. A high level of quality of the recordings is generally recommended when a sufficient storage space is available. In other case, a trade-off between space and video quality should be made in advance.

Discussion

This paper has presented a technical and physical infrastructure to carry out evaluations of mobile assistive technology with visually impaired users. The preparation and the execution of the laboratory test led to a series of reflections and lessons learned by the research team that are considered useful for future usability and accessibility research with visually impaired users. In addition, several lessons can be inclusively applied when testing touch interaction with able-bodied users.

An infrastructure suitable for the evaluation of touchscreen assistive technology with disabled users (RQ1) would be one that firstly optimises data collection; secondly, allows the research team to do an effective retrospective analysis under different and more demanding conditions than when testing able-bodied users; and thirdly does not interfere or trouble the comfortability, safety and trust of the users. Having in mind that sensory-limited users do not have the same level of access to information, leaving aside that not all information channels are designed with this type of users in mind, their
comfort and tranquillity are crucial to avoid interference and distortion of the test and results.

The proposed infrastructure contributes to a controlled scenario for evaluation; however, it is not exempted of potential improvements that can qualitatively benefit future tests and be applied to other mobile technologies and able-bodied users (RQ2). For instance, to evaluate the accessibility of touchscreens and the choreography of gestures associated, the video recordings require a sufficient quality that allows zooming in with great detail and professional software video visualisation to substantially reduce the speed for optimal viewing. In addition, the data should be collected through multimodal channels (e.g., video and audio), having the necessary tools to synchronise audio and video signals, which, if streamed over a network, usually incorporate latency. This synchronisation is the key to detect and understand the correlation between the sounds of the interface related to participant’s touch on the screen. Finally, due to the inherent difficulties of recruiting disabled users and the discomfort of having to unnecessarily repeat tasks and test sessions, redundancy in data collection is strongly advise through the use of two or more independent sources of data storage, i.e., two different computers.

Conclusion

Mobile assistive technology for touchscreens is widely used by multiple user groups. When designing, testing and evaluating technology with sensory-limited users, there is a specific need to balance the interface design and functionality on the one hand and the usability and accessibility of mobile assistive technology on the other. Accessibility and usability evaluations are essential in order to improve not only the interface design of the mobile assistive technology, but also the interactions between devices and users. These evaluations are enabled by a laboratory environment, where the research team has full control over all steps of the test scenario, including tasks and interactions between the test participants and the technology used. In particular, for mobile assistive technology that involves visually impaired users, accessibility and usability evaluation aids to identify interaction issues that lead to uncover design flaws, obstacles to successfully use the device and potential adjustments of the system to accommodate user sensory limitations. This paper has analysed the physical and technical infrastructure used for evaluating a mobile user interface using a gesture-based speech-assisted interface navigation system, Apple Inc. VoiceOver., within the research project “Visually impaired users touching the screen - A user evaluation”. The test infrastructure provided sufficient control over the factors involved in the test at the same time that brought the flexibility to dynamically adjust the environment for adequate data collection.

Empirical research data obtained from the usability and accessibility evaluation using the infrastructure described in this paper will be published and available for the research community. Future research in the agenda of the authors includes the test of the infrastructure including the technical improvements proposed in this paper with other user groups, including other vendors and solutions of assistive technology for operating mobile user interfaces.

Acknowledgments

The research team thanks all test participants for their disinterested contribution. Thanks also to the Centre for eHealth and Healthcare Technology of the University of Agder, Norway, for lending and adapting their test infrastructure facilities.

References


Address for correspondence

Berglind Smaradottir
PhD Research Fellow/ Project Manager
Department of Information and Communication Technology
Faculty of Engineering and Science
University of Agder
Post Box 422, N-4604 Kristiansand, Norway
E-mail: berglind.smaradottir@uia.no
Phone: +47 37 23 37 78
Mobile phone: +47 95 23 66 25

46