MASTER THESIS

in

Universal Design of ICT

May 2017

Investigating the Lydhør application from the NLB: An Accessibility and usability assessment in the context of Universal Design

Thomas Godwin

Department of Computer Science

Faculty of Technology, Art and Design

OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

1 Preface

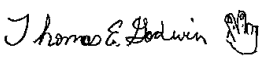
I would like to thank my wife Anne Cathrine, and my sons, Lucas and Ethan for understanding that conducting this research sometimes required “holing up” in a quiet cave, alone, to get stuff done. Sorry that it stole some of our together time.

To my advisor, Norun Christine Sanderson, many thanks. Had I listened to your words of wisdom earlier, this project would have gone much smoother, but sometimes I had to learn the hard way. I can certainly say that there was some luck involved when you were assigned as my advisor. Good luck on my part, and bad luck on yours, for having to deal with me. All jokes aside; your calm, systematic, “no BS” approach, helped to recover from or avert multiple crises, and without it, we both know that this project would never have gotten back on schedule, and that the organization of the thesis would have been pure chaos.

I would also like to wish Siri Kessel a happy retirement, and thank her, and the many other staff members at HIOA, who took time out of their chaotic schedules, to offer advice (or constructive criticism).

Finally, I would like to take the opportunity to thank, Arne Kyrkjebø at the NLB, for letting Lydhør be available to for dissection, and for taking the time to meet me on many occasions, answering 1000 emails, and for the book tip (I should have the time to finish reading it soon). I hope that you guys can use my findings to make Lydhør the “model” application of its’ type, and make all of your sister organizations jealous, and eager to make their application as good as Lydhør will be.

Oslo and Akershus University College of Applied Sciences
Oslo, Norway
15.05.2017
Thomas E. Godwin II



2 Summary

The research in this project was an exploration of Universal Design principles, applied to an existing mobile application. The developers of the application (named Lydhør) have a challenge in ensuring that the application is accessible and usable, due to the unique group of users that utilize the application. The application, simply put, is an e-reader, that reads aloud books with a synthetic voice and books that are in audiobook format (a human has read the book aloud and recorded it).

To qualify to use Lydhør, one must have what is loosely termed as a "print disability." This means that users must have problems with reading "conventional" books. This can stem from blindness or limited vision, or an inability to turn pages in a book, or many other reasons.

Universal design (in the context of IT), in short, is the process of designing an artifact to be as accessible and usable to as many users as possible, within a single design. This represents a unique challenge for the Norsk lyd- og blindeskriftbibliotek[1] (NLB- The Norwegian audiobook and braille book Library), in ensuring that their application is "universally designed."

For this research, the Lydhør application was examined and assessed to determine if accessibility and usability problems were present in the application, with a higher priority placed on ensuring accessibility. The process often involved various inputs from users by design, which improved the quality of the output. For problems that were identified, solutions were created and then tested, by members of the print disability group.

Midway through the research, only had a few minor accessibility transgressions had been identified (the highest priority goal of RQ1), and the focus of the research shifted towards identifying and testing solutions to usability problems.

The design of the project was intentionally dynamic, to be responsive to information learned from user inputs from the earliest phases. It was impossible to know which problems existed before performing earlier research activities. Although the final research activities had been planned, their precise contents were intentionally "left open."

This research is carried out in the context of universal design focus, so needless to say, some design activities took place, which increased the quality of the final output.

Because of the role of design in this project, a quasi-SCRUM approach was loosely followed (for design related activities), which allowed the designs solutions to be responsive to problems actually identified, not predetermined activities, which might have served to ignore the user inputs. The advantage of doing so, is most eloquently

[1] www.nlb.no

described by Ken Schwaber. Scrum allows researchers/designers “to devise the most ingenious solutions throughout the project, as learning occurs and the environment changes (Schwaber, 1997).”

To accomplish the goal of answering the research questions, a host of research methods/activities were used. One of these was a survey that gained input from over a hundred users. From information gathered in the survey, a prototype application was created that attempted to address problems that were identified in the survey. The prototype was then tested (with participants from the print disability group), to learn if perceived solutions, were in fact improvements.

There were many insights learned that should serve to improve not only Lydhør, but can also improve accessibility and usability in mobile applications in general, and with mobile e-readers, and will add actionable knowledge to the limited research existing that is specific to “print disabilities”.

Table of Contents

3 Table of Figures

4 Introduction

This masters project research assessed the accessibility and usability of the Lydhør application, in the context of Universal Design principles. Lydhør is an application that patrons at the NLB use to access various book titles on their smartphone. The application was researched to determine if there were accessibility or usability problems were occurring. When possible solutions were created, they were tested, to see if they did indeed address the issue as intended.

Lydhør users must qualify to use the application, and the requirement is that they must have a print disability (this is discussed further in the literature review). The term "print disability" is not a particularly positive term, and the hope is that a new term will evolve that is less demeaning. However, it is used in this thesis, because it is likely the term researchers will use when searching for articles, particularly those in the "librarian" field, who use the term often.

The project served to fill a void in scholarly research, as little exists that specifically address print disabilities in the context of smartphone use and in the context e books (utilized on mobile devices). This group (as a group), have received little focus (in the context of universal design and e-readers for mobile devices), yet they have the need to access books like all other people. The likelihood that e-readers will be more frequently accessed via mobile devices is high, especially those that read the text aloud.

Researchers that have performed user testing that pertained to e-readers and accessibility performed in 2013, noted the scarcity of these articles, and pointed out the need for more research in this area. The authors also noted that the growth of digital reading devices threatened accessibility for the print disability group, because the devices and books themselves were not being designed in a proper manner. Notably, even their research did not include users that were "print disabled" (L. Maatta Stephanie, 2014)

Another study echoes the need for this type of research. The authors of an empirical literature review of research performed in 2015, that is related to the same subject area, urged for research in this field. They pointed out that the existing research in this area is not proportional to the population of users in the print disability category (Lundh, 2015).

Another article encourages more research, this time particularly in the context of mobile use, because of newer applications that allow articles to be accessed by phone that are being created are nearly unusable to the print disability group (Blechner, 2015).

4.1 The Norsk lyd- og blindeskriftbibliotek

The NLB, (which would translate roughly to the Norwegian braille and audiobook library), is a state maintained organization, located in Oslo. Their purpose is similar to the mission of regular libraries. However, the format of their books is different from those offered in typical libraries, because their mission is to ensure that everyone (in this country) can enjoy reading.

The library is not open to the general public. It is only available to citizens of Norway (and is digitally accessible to residents of other countries that have agreements with the NLB, and typically offer similar types of apps).

The NLB does not disclose demographic details about the precise totals of which specific disabilities make up their patron group, likely because of ethical implications. They do state on their website that to qualify to use their services, one must be considered to have a "print difficulty"[2].

An application named Lydhør (translates roughly as audible), is the application that the NLB offers for mobile access to their services. With the proliferation of smartphone use, the importance of this application is growing. This service could possibly replace other services such as braille books, and books that could formerly be "read aloud" via pc, because of the obvious convenience that comes with the convenience of being able to access books via mobile phone. The books that are in text format that they offer (not the audiobooks), can, for example, be output to braille keyboards, which removes the need to print these books in braille.

4.2 Problem statement

The Lydhør application, is a mobile app, provided by the NLB, to its patrons. The application is used for listening to books, that are read aloud to users who cannot utilize traditional books for various reasons. Lydhør is utilized by a diverse population made up strictly of members that are "print disabled." This creates a possibility that issues can exist, specifically pertaining to accessibility and usability for this group.

Investigating Lydhør in a universal design context, with frequent input from Lydhør users, can improve the experience for them, potentially making it both more accessible and enjoyable. Potential solutions will be created to address any issues noticed during early phases, and then tested with users, to assess their viability. Any improvements that are made to Lydhør, can also be projected and applied to other research/development that involve print disabilities, mobile use with e-readers and other mobile applications, and universal design topics.

4.3 Research questions

RQ1 –What problems are Lydhør users experiencing, specifically related to accessibility and usability of the application?

RQ2 – What are feasible improvements that can be made to increase accessibility and usability, that can address the items identified in RQ1?

RQ1 offers a broad way of identifying possible issues stated in the problem statement. It will be answered by a variety of research methods. RQ2 is directly dependent on what is learned from RQ1. Items identified in RQ1 will be attempted to be resolved, and these

[2] http://www.nlb.no/bli-laner/hvem-kan-laane

resolutions will be tested (with Lydhør users). The final output of this research will be the improvements spoken of in the problem statement.

4.4 Direction of the research

The core components of this research are: The Lydhør application, and the fact that this research is being performed in the context of a universal design of IT focus, and mobile accessibility/usability. The information will also be useful to researchers/developers interested in the print disability category, particularly in the context of mobile use in general, and also specifically in the context of e-readers (on mobile devices). Universal design principles, when applied, also often serve to benefit all mobile users, not just the print disability group.

One of the initial decisions made regarding this research project, was in fact, not scientific in nature. But nevertheless, it played an important role when making future decisions, so it is included here.

This decision was that the result of the research should, in some fashion, benefit or attempt to benefit the Lydhør users. Although this decision was made prior to learning about the existence of action research, it was following the same line of thought of one of the core facets of action research. In particular, the aim was similar to McNiff's description of action research "A main aim of action research is to generate knowledge that can lead to improved understanding and experience for social and environmental benefit (McNiff, 2013)." Note that a strict action research method was not followed for this project, but the goal of producing beneficial research was strived for.

4.5 Thesis structure/reader guidance

The structure of this thesis may be somewhat atypical when compared with equivalent academic documents. Two factors precipitate the abnormal structure, the first being the presence of design related activities. In the context of a universal design (of ICT) concentration, design discussions enhance the outputs of the research, and demonstrate mastery of the subject. They are a way to demonstrate a practical application of data analysis. Finding the proper placement of these discussions can be problematic, however.

Secondly, the research was intentionally designed to be *dynamic*, and based on data gathered from various iterative user input techniques (this is discussed in the introduction to the research methods chapter). One goal of doing research is to learn answers to something. In this case, RQ1 had to first be "answered" before the specific details of the final activities could be determined.

Trying to find the proper locations for the survey design and prototype/user testing design discussions proved to be problematic. Because of the iterations, these activities are based on results from prior activities. They also occurred in a linear fashion. Discussing these items before addressing the results that they were built upon, or "out of order" detracts from their quality and understandability.

In order to address the abnormal demands, the results discussions are split. There were many research activities, each with their own set of results. After analyzing the data, and practically applying it, a new activity began.

Therefore, after the research method and literature review sections of the thesis, the results are presented linearly, but split to accommodate discussion of design activities, to preserve the linear fashion in which these activities occurred. This also meant that some of the research method discussion had to occur in the design sections, to preserve linearity. The discussion and conclusion sections follow. Hopefully, the need for deviating the structure, will be obvious, and appreciated, after the document is reviewed in totality.

5 Literature review

5.1 Literature review structure

The beginning of the literature review focuses on laws/regulations that affect newer technologies and this research. Then, some information is provided about the varying opinions about what universal design (of IT) entails, which has implications that affect the regulations. This is followed by a short description of print disabilities, which is the demographic group that is the focus of this research. This is followed by a section about varying characteristics of the people with vision limitations and a section focusing on dyslexic users.

The remaining sections are dedicated to other members of the print disability group, such as those with cognitive impairments other than dyslexia and people with motor ability impairments. These last sections contain less information than the blind and dyslexic user sections, because there was less research available that was applicable and *of reputable quality*. Most of this information had to be extracted from research focused on accessibility topics that focus on computer interactions, so that they could be extended to a mobile context.

5.2 Applicable laws and regulations

Laws that mandate that ICT artifacts are universally designed or accessible (or both) are becoming increasingly common. For example, the US 508 regulations require that employers use and purchase accessible technologies and that government websites are also accessible[3]. Note that this is not an exhaustive description of the regulations.

This thesis was composed in the country of Norway, which has passed some of the most progressive laws in the world regarding accessibility. The regulations contained in the Norwegian anti-discrimination and accessibility act, regulate not only government sector but also the private sector and are enforceable[4]. This law is particularly interesting because it specifically uses the term “universally designed”.

The codification of the term “universal design” into law will not be all inclusive or perfect. These laws lack coverage for some users, some of which are in the “print disability” group. Particularly those with cognitive impairments. The WCAG guidelines, which are often referenced by accessibility laws, are also lacking in regard to cognitive impairments, and do little to address them in their documentation[5]. The US regulations have only recently (early in 2017, enforceable in 2018) began to address cognitive issues explicitly[6].

In the future, these regulations will be applicable to mobile applications in some

[3] https://www.section508.gov/content/learn/laws-and-policies
[4] https://lovdata.no/dokument/LTI/lov/2013-06-21-61
[5] https://www.w3.org/TR/WCAG20/
[6] https://www.access-board.gov/guidelines-and-standards/communications-and-it/about-the-ict-refresh/final-rule/iii-major-issues

instances, and could have implications that affect their design. However, the requirements that these laws insist on, have yet to be fully developed, and are therefore possibly subject to abuse via poor interpretation or misinterpretation.

5.3 Defining universal design

There is no singular, established definition of universal design when the concept is applied to technology or IT. Universal Design is a concept coined by Ron Mace, he defined it as "… the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design[7]."

This definition was intended for the physical environment, such as buildings and street crossings. When the concept was later applied to technological artifacts, there was no unified definition. Rather, there was a conglomeration of interpretations. Not all of which agree on the finer details.

The dispute generally revolves around two aspects of universal design, accessibility and usability. The W3C states "that Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web[8]". For the purposes of universal design (of IT), which encompasses more than the internet, this description can be extended and interpreted as "IT artifact accessibility means that people with disabilities can perceive, understand, navigate, and interact with the artifact."

Usability on the other hand, is defined in one article as "the set of factors that qualify the interaction between the user and the software (Casare, Silva, Martins, & Moraes, 2016)."

The definition universal design is not only vague, in terms of the regulations, it is also frequently interpreted differently in scholarly literature. One researcher, who has done extensive research in this country (related to universal design), interprets some wording changes that have occurred in evolutions of applicable laws in this country, as a shift from "'usable for all people, to the greatest extent possible' to 'can be used by as many as possible,'" (Fuglerud, 2014), which seems to place emphasis on accessibility.

Additionally, competing terms/approaches have emerged, each varying somewhat, sometimes only in minor facets and sometimes varying quite substantially. Although the goal is similar, other approaches such as accessible design, design-for-all, and inclusive design must be noted. The discussion is now complex enough that there is a lengthy scholarly article dedicated to a discussion of the goals and differences between these methodologies/approaches. Notably, this discussion focuses heavily on the idea of accessibility (Persson, Åhman, Yngling, & Gulliksen, 2015).

On the other hand, one study, that focused on usability of legal research databases notes that information in legal research is oftentimes not usable, and thereby

[7] https://www.ncsu.edu/ncsu/design/cud/about_ud/about_ud.htm
[8] https://www.w3.org/WAI/intro/accessibility.php

inaccessible to the print disability community (Blechner, 2015). They used the terms universal design and usability interchangeably, and consider usability as a core component of universal design. Some researchers that tested the effectiveness of the WCAG guidelines (1.0 & 2.0) advocate that the definition of accessibility be expanded to include the wording ‘‘usability for all’’ (Rømen & Svanæs, 2012).

The DOIT group, a group focused on ensuring that impaired individuals achieve the same fulfilment from life that non-impaired individuals receive, point out that disabled users are often excluded from usability testing. They note from experience that increasing usability also serves to increase accessibility[9], and they have collected empirical data to back up this claim.

In addition, in some articles there is even language such as classifying “accessibility as a sub-class of usability (Casare et al., 2016)”.

Note, that for the purposes of this research, universal design is interpreted as containing both accessibility and usability, and that the implied priority is on accessibility. This does not, however, mean that usability should be disregarded, it should be viewed as a factor that improves the quality of accessibility.

5.4 Defining print disabilities

A short discussion of the print disability group is presented here. This group is the “focus” or “target group of this research.

Potential patrons of the NLB are generally people that would have some form of print disability. The examples given as to whom qualifies for NLBs’ services that are given on the NLB website[10] are as follows:

- Impaired vision
- Dyslexia or other reading difficulties
- ADHD with difficulty concentrating
- Problems holding a book
- Other cognitive challenges and language difficulties that have occurred due to sickness or physical damage

This definition is broad, likely with the intention of ensuring access to anyone in need. Users can qualify for a reason that is not on the list, if they can demonstrate something that impairs their ability to read “traditional paper” books.

Various definitions of the term “print disability” are used in academic research. Beyene, a researcher whose focus in on print disabilities related topics, describes it as “difficulty associated with effectively utilizing print text due to visual impairment, physical disabilities, and some forms of learning disabilities (2016).” Another definition is similar but includes additional language such as “developmental” and “learning” impairments

[9] http://www.washington.edu/doit/what-difference-between-accessible-usable-and-universal-design
[10] http://www.nlb.no/bli-laner/hvem-kan-laane

(L. Maatta Stephanie, 2014).

Despite not being able to obtain precise demographic information about the composure of the Lydhør user group, an interview with Arne Kyrkjebø, a director at the NLB, did provide some broad indications about the Lydhør user group. He stated that the population of users is diverse, and that a large portion of the patronage consists of blind or visually impaired members. There are many elderly users, who will qualify for use because of various reasons, which is often declining vision. There is also a group from a younger spectrum who receive their textbooks in audio format from the NLB (although the NLB is not the only institution that has responsibility for this). Although it was not stated implicitly, comments during the interview left the impression that vision impairments of all ages would represent the majority of users, many of which are elderly, and that the next largest group would be students affected by dyslexia (Kyrkjebø, 2016).

The authors of one usability research project involving the print disability group note that the definition of "print disability" has evolved, mainly because of legal implications, and has only recently began to include cognitively impaired users (Blechner, 2015). This change can also be noted within the Daisy consortium[11], which in the past focused mainly on blind/visually impaired users, as they have begun to add support for books that in some cases can be both heard, and the words can be followed along with visually in some of their newer formats (Lundh, 2015).

5.5 Research pertaining to blind and limited vision users

Most likely, the largest portion of NLB patrons are blind or have severely limited vision. Some of the recent research regarding blind users relates to memory and learning styles. These two factors have a substantial impact on the learning process itself and are explored in detail below.

5.5.1 Memory and Learning Issues for Blind Users

Some Lydhør users are visually impaired students that use the app to "read" textbooks and for general reading. In addition, visually impaired people typically need to "learn" or memorize functions of an application to be able to use it effectively. The learning process for blind users is different than that of sighted learners. This is due to the loss of the ability to learn visually, which is a common tool used by teachers to improve the learning outcome (Izzo & Bauer, 2015).

One study that revolved around using multimodal inputs and outputs for blind users set out to identify limitations experienced by blind computer users. It also tried to determine which of various modes could be utilized by blind users and which of these modalities are best for various types of tasks. They referenced a substantial amount of prior research that showed that different types of inputs (i.e. auditory or visual), utilize

[11] http://www.daisy.org/

different portions of the brain. When different inputs are utilized, the workload of the task at hand can be distributed to different parts of the brain (Shimomura, Hvannberg, & Hafsteinsson, 2010).

Prior research had suggested that using a combination of sensory cues increases the efficiency of the learning process, and is often a design consideration with multimedia material used for learning. It had also shown that providing too much input information to the brain is detrimental because of the limited amount of information the brain can consume (Shimomura et al., 2010).

They recommended “small, reversible incremental steps” for blind and vision impaired users, presented in hierarchical design where users could quickly sort their options to increase efficiency. Efficiency is a problem with this group because they often have to process much more information than a visual user to achieve the same goal (Shimomura et al., 2010).

For visually impaired users, the mental workload is a factor. Interfaces and/or menus/lists of information must be memorized in many cases, to *efficiently* use an application. In one study, researchers found that despite the efforts put forth so far in regard to improving the different generations of DAISY players, there are still notable issues with the mental workload required when operating the device (Huang & Chiu, 2016).

Another study which tested preferences echoed the idea that memory can be quickly overloaded. In that study, when blind users were introduced to complex user interfaces, they preferred the tactile paper prototypes versus the mobile application version. With the paper (braille) prototypes they could get a better “quick overview” of the information presented. They also felt that “navigating” was simpler using the paper versions, due to the heavy memory load required to try to memorize identical menus presented in the mobile application, in order to avoid having to rescan it (Miao, Pham, Friebe, & Weber, 2016).

In an educational environment, the idea that all people learn differently, i.e., pictures are great for visual learners, but may be less helpful for people who learn more by reflecting on examples, is a common idea. The more ways the material is presented, the higher the likelihood that for example the most possible students in a class of diverse individuals will learn the materials. One study noted and explored the various methods used and reasons for using them by college professors (Izzo & Bauer, 2015). They noted the fact that they are increasingly using media and technology with the intention of presenting a concept in various formats to increase the likelihood that a concept is learned. They also advocate that this method helps to accommodate various learning deficiencies.

One of the recommendations in the findings of the Shimomura study was that when researching accessibility issues or technologies that will be utilized to render something accessible for people with vision deficiencies, that a substantial portion of the research

should be dedicated to the methods in which blind users learn (2010).

They asserted that different methods of learning are more efficient for different types of tasks and provided an applicable example. Blind test subjects, when having to solve complicated, information laden math tasks, must either try to memorize and retain large amounts of data when they receive the information via text-to-speech, or note down the information. Similar to the way that sighted users reference the information visually on the page, blind users preferred to “scroll” back to the information via the braille reader (Shimomura et al., 2010).

5.5.2 Feedback, critical for blind users

Feedback is especially critical for blind-users, because in the absence of it they cannot ascertain what is occurring with software. One group of researchers focused on trying to ascertain the optimal mix between methods of auditory feedback for blind users. They found that feedback presented as redundant speech tended to offer the proper information to the users, but was slightly annoying and required a heavier focus cognitively. The other form of feedback was various short sounds that had a certain meaning (in this case directions where to walk). These short cues were more pleasant to the user but typically do not contain enough information when detailed information is necessary (Ibrar Hussain & Gencai Chen, 2014).

These researchers proposed and implemented a “hybrid” system in which the blind users were presented with a mix of instructional signals that helped them (physically) navigate. During their research, they found that for an application giving walking directions, the users were uncomfortable with the pure non-speech directions. They were satisfied with a carefully blended mix of auditory sounds and verbal directions. Their final conclusion was that a carefully blended mix of verbal commands and short auditory cues provided a more “efficient” method of conveying navigation information (Ibrar Hussain & Gencai Chen, 2014).

Another study found that for blind users, continuous feedback increased effectiveness levels when they operate controls or navigate through software interfaces (Panchanathan & McDaniel, 2015).

Shimomura et al. found that in order for software to replace the human assistants that are presently often utilized in all-blind schools, that it not only should provide continuous feedback, but the feedback should be provided reactively, as well as proactively (2010).

Findings from another study found that blind users could navigate graphical webpages, (with a haptic mouse), somewhat efficiently. Haptic feedback from the mouse alerted these users that they had encountered an actionable page element, and where the borders of it were located (Charoenchaimonkon & Janecek, 2015). The blind users preferred the “rapid and continuous control feedback” versus the slow, sequential, procedural feedback that is typical of applications that tailor to blind users.

In another study, researchers asked blind users to create gestures that they would like to use. They found that these gestures were more effective when auditory or haptic feedback was provided that confirmed that the gestures had worked (Dim & Ren, 2014).

At the conclusion of their study they established guidelines for designing motion gestures used to interact with smartphone interfaces. Their first and most important recommendation was to provide feedback for every action that was performed (Dim & Ren, 2014).

Dim and Ren found that intricate information could be relayed to sightless users via haptic signals. And, when these signals are thoughtfully designed, certain feedback with certain gestures felt almost natural or expected, with little or no training (2014).

While haptic feedback can improve the experience, it should not be the only source of feedback. One study found that haptic feedback may be much less effective for elderly users. This is due to a scientifically proven decline in tactile sensitivity that occurs with aging (Benedito, Guerruero, Nicolau, & Goncalvez, 2010). The same study had pointed out that over 82% of the blind population is older than age fifty.

5.5.3 Navigation problems for blind users

One study noted the problems that have the most detrimental effect on blind software users. The worst problem (per the blind users), was controls which were not labeled or poorly labeled for screen readers. The next worst problem was when content and navigation elements were poorly structured. These two problems were most detrimental in terms of the “extra time” that was added to tasks due to these problems (Miao et al., 2016).

Other research that focused on the usability of VoiceOver by blind users identified problems that hinder navigation. VoiceOver could sometimes not provide enough information about interactive elements. “Expanding” content sometimes confused navigation efforts. They also noted that there is often focus handling problems with text input and forms (Leporini, Buzzi, & Buzzi, 2012).

5.6 Elderly users

Per the director of the department at the NLB that has responsibility for Lydhør, a substantial percentage of their patronage falls in the elderly category (Kyrkjebø, 2016). This of course includes a broad spectrum ranging from completely blind elderly users to elderly that have perfect vision and qualify for membership there for some other reason.

However, there are other concerns with elderly users not related to vision. The WAI lists the typical impairment groups suffered by elderly that affect their use of technology. The four broad groups are reduced vision, diminishing physical ability, declining hearing, and shortcomings in cognitive abilities [12].

One study’s findings noted that many elderly users had lost the motor skills required to

[12] https://www.w3.org/WAI/older-users/Overview.php

successfully perform gestures typical with mobile phone use consistently (Dim & Ren, 2014). Additionally, another researcher noted when reviewing previous studies that there is a "wide gap" between experience levels (relating to computer based technology) when comparing elderly users and younger users (W. Chen, 2013).

However, elderly users, despite having less exposure to technology, can learn and utilize technological apparatuses. One group of elderly test subjects previously unfamiliar with the involved mobile technologies did in fact easily learn swipe to scroll, and pinch/spread to zoom. They had also easily grasped the back button, and could navigate fairly effectively, even if they had to return to a home page or start page of an app in worst case scenarios (Barros, Leitão, & Ribeiro, 2014).

However, knowing what to do versus being physically able to do the action can be problematic. Researchers in a study involving elderly testers in a mobile context noted that the "swipe" gesture was problematic for them (Teixeira et al., 2016).

Barros et al. recommend not placing soft "or software" buttons at the bottom of the screen on mobile applications, because this induces many accidental presses of the hard buttons on the bottom of the phone. (Barros et al., 2014). They also noted that elderly users will often not use software/hardware designed specifically for elderly users, so to increase acceptance, designers should exclude the "elderly" designation.

The previously mentioned research that had noted the swipe performance capability problems, also noted some other key interaction problems for elderly for their mobile prototype. The highlights of these are that the testers found the menu structure to be complex (even though it consisted mainly of six tiles with icons occupying the entire screen). Users also complained of having difficulty returning to the main menu, difficulty with the small font size, and about confusion caused by lack of feedback (Teixeira et al., 2016).

Another study found that elderly users were more efficient with touchscreen input than with hardware style inputs such as physical keyboard (for example a Bluetooth keyboard), and they tended to prefer this style of input (Oehl, Dahlmanns, & Sutter, 2013).

One group of researchers developed a set of best practices for web accessibility (Zaphiris, Kurniawan, & Ghiawadwala, 2006). Notably some of the content had become dated due to advancements in technology, or newer research that improved upon it.

A newer study took this set of guidelines, refined it where it was applicable in a mobile context, and tested it. There were many which were applicable to this research. Such as the recommendation to use "large targets," although no specifics details were given. They noted that feedback should be obvious and that feedback methods should not consist of minute changes that are hard to notice. Additionally, designers should avoid requiring input in the form of a double click, and avoid using scroll bars (Díaz-Bossini & Moreno, 2014).

Other research that specifically focused on gesture completion by elderly users, also found that the double tap gesture was difficult for seniors to successfully perform along with the long tap (Cáliz et al., 2016). They noted that other gestures such as the swipe were easier to use, and that many seniors are capable of completing them.

They also recommended using one "screen" or "window" at a time and the use of simpler layouts. This was complemented by the recommendation to provide fewer choices to users to reduce the demand on memory (Díaz-Bossini & Moreno, 2014).

Another study that obtained qualitative input from elderly users echoed some of the previously mentioned findings/recommendations, such as large screens and simplicity. They noted that testers identified the "swipe" gesture as being problematic, which contradicts the Cáliz et al research (Van Biljon & Renaud, 2016).

And finally, one group of researchers noted that elderly users have a preference for a "qwerty" type keyboard for touch screen devices. They are familiar with that layout, and the location of the keys matches their expectations (Rodrigues, Carreira, & Gonçalves, 2014).

5.7 Dyslexia

The International Dyslexia Association[13] defines dyslexia as being "characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities." One study points out that dyslexia, is often accompanied by other simultaneously existing conditions, that impair writing, attention, and concentration (Rello, 2015). Note that others consider these conditions as a part of dyslexia.

Many of the articles in this section refer to dyslexia as a cognitive impairment, which it is. The section following this one is dedicated to cognitive impairments, but dyslexia is presented first, because it has characteristics that distinguish it from the others discussed later, and more research was located about this topic versus other cognitive impairments, warranting its own section.

The authors of an older web navigation accessibility study, pointed out long ago that the WCAG guidelines nearly neglect all cognitive impairments, and noted a void in research dedicated to cognitive impairments and accessibility (Small, Schallau, Brown, & Appleyard, 2005).

Rello notes that many studies related to dyslexia and accessibility focus on large text sizing and sans serif fonts, and results of his research advocates this as well. Additionally, they advocate using even larger text, and excluding both serif fonts and italic fonts to improve reading speed, accuracy, and comprehension for dyslexics (2015).

Other prior research related to dyslexics and internet use suggest a design that presents as little text as possible, and make several situationally dependent suggestions about the

[13] https://dyslexiaida.org/definition-of-dyslexia/

most effective background colors in various contexts (C. J. Chen & Keong, 2016).

Another web related accessibility study identified the key problems dyslexic users face when using the internet. These were text sizing and color problems, and they also noted that some complicated webpage layouts could be confusing. Complex language caused a problem for dyslexic users, and they that sometimes finding the proper way to navigate could be problematic for them (McCarthy & Swierenga, 2010). Note that this research is a bit dated, and the navigational problems may have been related to the "text intense" navigation structures that were used more frequently then. They also found that screen readers that highlighted text as it was read were beneficial, along with the use of plain language.

Researchers that used brain scanning, noticed that there was mounting evidence that dyslexic readers were dependent on memorization of words, rather than sounding the words out (Shaywitz & Shaywitz, 2007).

Other research that set out to "measure the effectiveness" of the WCAG guidelines (versions 1.0 & 2.0) noted that "links" should not be placed close to the edge of screens, and that "crowded" webpages were difficult to read and navigate for dyslexic users. They also found that when the interface changed drastically, it was confusing for them (Rømen & Svanæs, 2012).

Berget compiled and published several articles related to dyslexic users (and technology use) (oftentimes with other researchers as well), when composing her doctoral thesis. They contained several useful findings.

They found that presenting icons in search results did not decrease search times for dyslexic users, and may even be counterproductive (Gerd Berget & Frode Eika Sandnes, 2015a). It is important to note that this is in the context of search results.

Another of her studies suggested using an autocomplete function that makes suggestions as the search terms are being entered, and that the autocomplete should have tolerance for errors that dyslexic users typically make (Gerd Berget & Frode Eika Sandnes, 2015). Search engines could also suggest other searches for parts of the terms that have no results.

They had found in the subsequent study that autocomplete may improve search performance and reduce the effect of dyslexia on search behavior, when it was substantially error tolerant towards dyslexic typical data entry mistakes. They also found that the autocomplete language should be modifiable to match the search language. Finally, they noted that wordy search result lists were troubling for dyslexic users (Berget & Sandnes, 2016).

Some of Beyene's research focused on search results with dyslexic users. He had noted multiple comments from them explaining that search results were too lengthy, which reduced their usability, as Berget and Sandnes had found (2016).

Another of his observations of interest was that users of the same impairment, sometimes had different and opposing preferences. Many testers indicated that "too much" information can be overwhelming (the testing included dyslexic and blind users). There were also dyslexic participants that preferred having icons in search results. This preference was limited to a color coded icon that displayed availability of a book or article, that could be quickly scanned to learn the status (Beyene, 2016).

5.8 Cognitive impairments

One systematic review of web related accessibility issues provided a list of "best practices" from previous studies related to cognitive impairments and accessibility. Some of these advocated using simplicity when designing, such as the use of simple language and intuitive navigation controls, limiting the "layers of a site," and using a "simple" design. They also mentioned grouping similar controls and menu content. There were many other recommendations, but most of the others were so simplistic that they were just "common sense," or standard good design practices (Bernard, Sabariego, & Cieza, 2016).

Findings from another study involving young adults with cognitive impairments noted that software and hardware use could be challenging for these users. Again, overall simplicity was recommended, and it was noted that these users typically avoid complex functions, some even only used their smartphones to make and receive calls. One of which had impaired memory, and could only make calls with from stored "speed dial" numbers, and could not remember the number to dial it. Many of these users found the menus and structures of typical smartphone software complex and intimidating, but note that these were an older generation of smartphones as this research is a bit dated (Dawe, 2007).

Another study observed navigation and search task performance for intellectually challenged individuals. They found that the simple layout of the Google[14] search engine page was preferable to these users for searching, versus a Portuguese language search engine/news website, Sapo[15] that uses a crowded and distracting page. The testers seemed to try to memorize the position of the search text box on the Sapo page for multiple searching tasks, and they had difficulty locating the search box with each new search task. Finally, they recommended using buttons with images instead of text, notably they meant images, not icons. They also recommended presenting search results as images, but notably, supplied no further information about how to accomplish this (Rocha, Carvalho, Bessa, Reis, & Magalhães, 2016).

An application that simplified Facebook content and the process of using the Facebook mobile application was tested with cognitively impaired users. Some of their

[14] www.google.com

[15] http://www.sapo.pt/

recommendations were more fitting for use in specialized adaptive software, such as using familiar voices with audio prompts. But they like others, advocated simplicity, and use of larger buttons. They also noted that consistent design features are critical along with enhanced “error forgiveness (Davies, 2015).”

5.9 Motor impairments

One study that researched touchscreen use by people with motor impairments noted that the needs of these users vary widely, because of the broad range of possible impairments that exist, with varying intensity. To address the various needs, they recommended use of flexible interfaces that are adaptable with various input possibilities (which both android and IOS include now). These users needed input mechanisms that are less reliant on physical accuracy (Chourasia, Sesto, Kuehn, & Wiegmann, 2017).

The need for modifications was echoed in another similarly study of touchscreen gestures, where the needs were dependent on the intensity of the impairments. They also recommend modifiable inputs to address these needs (Trewin, Swart, & Pettick, 2013).

5.10 Similar research

At the conclusion of the literature review, only one project similar to this research had been located. They too conducted user testing with members of the print disability group with e-readers, however they were testing variations between various dedicated physical e-reader devices (Maatta & Bonnici, 2014).

Their tester demographics varied as well, because they intentionally excluded users with severe vision limitations from their testing, to avoid, as they stated the “confounding issues” this would add to the testing process. Instead, they asked the testers such as dyslexic participants to assess how they “felt” the experience would be for those with vision limitations. They did however, note that the physical e-reader devices were completely inaccessible to blind users and some with vision limitations, with exception to the IPad which had VoiceOver (L. Maatta Stephanie, 2014).

6 Research Methods

6.1 A word about user involvement/iteration

One of the key intentions of the research, was to get input from end users, often. This shaped the choice of research methods, so it is discussed first.

The need for end-user involvement performed iteratively is one that is often repeated in academic literature related to universal design. One example of this comes from a classic article from 1985 about designing for usability, that stressed three key principles. The first is an emphasis on involvement of users, and the need to understand users. Secondly, “intended users should actually use simulations and prototypes to carry out real work, and their performance and reactions should be observed, recorded, and

analyzed." The third principle recommends iteration or more specifically "there must be a cycle of design, test and measure, and redesign, repeated as often as necessary." (Gould & Lewis, 1985).

These ideas are repeated again in an article about "user sensitive inclusive design," where the authors state that universal design and design for all "recommend that the needs of disabled people should be considered throughout all phases of product development (Newell, Gregor, Morgan, Pullin, & Macaulay, 2011)."

Nielsen places emphasis on awareness of "categories of users and individual user differences. (Nielsen, 1993)."

Gulliksen et al. recommended (and utilized) an approach to user centered design that had several similar features. Such as "early focus on users" which highlights an understanding of the users. Additionally, they call for user participation throughout all phases. They also call for development of prototypes (that are to be tested with users), and "continuous iteration" wherein design models are tested with users, redesigned based on findings from user testing, and reassessed with users (Gulliksen et al., 2003).

Several other similar ideologies that closely mimic these recommendations are discussed by Fuglerud, such as human centered design and participatory design (Fuglerud, 2014).

Kenneth Lazar, who has decades of experience in HCI and accessibility, particularly with blind users, echoes this sentiment in one of his books. He states that is important to test using participants with a diverse group of disabilities. In addition, he recommends designing an interface "that works well for a majority of users with impairments, especially perceptual and motor impairments"(Lazar, 2010).

6.2 Overview of research methods for the project

A combination of research methods was chosen for the project, rather than a single method. This combination would facilitate the techniques addressed in the previous section (namely gather user input, correct based on input, test if the input was properly corrected).

The methods were intentionally designed to be interdependent, so that outcomes from earlier research activities dictated the content and direction of later ones. This was to allow user input to dictate the research (and design) activities, in line with strategies employed with user centered design or participatory design for example.

The methods chosen were designed to gather qualitative data, because most of the useful design related information was predicted to be subjective in nature.

The methods chosen are shown below, in order of occurrence:

- Review of user ratings
- Functional testing of the application (cognitive walkthrough and a guidelines review)
- An informal interview
- A user survey (of users of the Lydhør application)

- User testing

When developing the project plan, application testing, followed by the survey user, and later user testing (with a prototype developed from input from the survey), was believed to be a feasible way to be able obtain input from users, in an iterative manner. The user testing would allow any design assumptions/corrections made after the survey data was processed to be vetted.

Originally, the ratings review and interview were not included in the project plan, but when the opportunities arose, these activities were included because they could potentially provide key user inputs (which they ultimately did).

Two design activities were also performed and discussed in this thesis. The design of the survey is discussed after the results of the first three activities listed above, because it was heavily dependent on them. The prototype/user testing design is discussed after the survey results, because it was dependent on them.

These design activities could be described as research methods and included here. However, doing so would detract from the thesis, and it would create a general feel of disorganization, so they are discussed after the results that they were dependent on.

6.3 User rating review

When downloading an application in the Google Play Store (in this instance but also true if the app were used on an apple device), one can access the applications' ratings[16]. In the Google ratings, users have typically given the application one to five stars, and make comments about the app. There were several existing ratings for Lydhør.

When downloading Lydhør for the cognitive walkthrough and guidelines review, the decision was made to access the ratings for Lydhør. This would allow an early opportunity to gather user sentiments about Lydhør, and identify reasons users liked Lydhør, or potential problems.

This process was not sophisticated, the existing app ratings were translated and assessed. Because ratings are sometimes removed when an app is updated, screenshots were taken and stored. Data gathered during this exercise was used to guide portions of the cognitive walkthrough, and was used to help develop the subject matter of the survey. This process was one of the procedures used to work towards answering RQ1. Screenshots of these reviews are included in the appendices.

6.4 Application testing

Testing Lydhør was a core operation of the research. It was performed to gain an intimate knowledge of how Lydhør worked and to allow a subjective assessment of the app's design, and to simultaneously do an initial accessibility/usability assessment.

A hybrid testing approach was used, which combined two forms of testing that were

[16] https://play.google.com/store/apps/details?id=com.nlb.android&hl=en#details-reviews

described by Lazar et al. as a “cognitive walkthrough” and a “guidelines review”(Lazar, 2010).

The cognitive review portion of the testing was simply operating the application, and using all available functions, in various contexts. This portion of the process was used to learn the application. In addition, issues that were mentioned in the user ratings were also investigated, and attention was paid to potentially problematically situations.

The guidelines review was performed to ascertain if the application conformed to the principles set forth in the Web Content Accessibility Guidelines (WCAG) 2.0[17] and the mobile version of the guidelines (that is still a work in progress); Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile[18]. For brevity, when these items are referred to at later times in this text, they will be referred to as simply WCAG guidelines or the WCAG mobile recommendations.

Lydhør was tested on two mobile phones, both with and without utilizing a screen reader (on multiple occasions). First a Samsung A3 GT-i9300 with an older operating system (Android version 4.4.4) was used. Later, a Samsung S7 SM-G930F (Android version 6.0.1) was used. TalkBack was the screen reader on both phones. There was no specific reasoning for using these two particular phones, other than the fact that they were the only ones available. The IOS (apple) operating system was not examined.

The results of the testing were recorded as testing was performed. Screen shots were taken when necessary and notes were taken as issues of interest were encountered. The results of this testing were used to develop the survey, the prototype, and influenced some of the design of the user testing. This procedure was used to gather more information to answer RQ1, and possible solutions that would address RQ2 began to be explored here.

6.5 Informal Interview

Despite the fact that the interview was not a part of the project plan, when the opportunity to obtain an interview user became available, the opportunity was capitalized on. This particular interviewee could offer unique insight, because this person not only was an experienced Lydhør user, but they also had substantial expertise in the discipline of universal design.

Questions and tactics for performing the interview, although developed on short notice, were designed based on tactics from the book entitled User Centered Design, which was being researched in preparation for the user survey (Lowdermilk, 2013).

The topics of some questions that focused on the Lydhør experience were planned beforehand. They focused mainly on potential problems that were noted during the ratings review and subsequent application testing. Time did not allow “scripting” of

[17] https://www.w3.org/TR/WCAG20/

[18] https://www.w3.org/TR/mobile-accessibility-mapping/

these questions, but they were asked based on a short checklist of keywords.

Simultaneously, the questions that were to be used for the next research method/activity, the user survey were being developed. The interview offered an excellent opportunity to "pre-test" the survey questions.

Pre-testing the survey questions before the survey is "officially" released, is recommended by Lazar in his book, Research Methods (Lazar, 2010), and has utility for various reasons. This book was also being researched in preparation for the user survey, but offered applicable information for the interview.

The rest of the interview time was intentionally left open. Letting the interviewee speak openly about the most important aspects from their perspective, would ultimately add more input for addressing both of the research questions, because of the unique skill set the interviewee possessed.

The interviewee was asked beforehand if they desired to choose the interview location, which they did. The survey questions were printed and taken (they are available in the appendices). The interviewee was asked to bring the device that they accessed Lydhør with, with batteries charged. Notes were taken as the questions were being answered. Directly after the interview, these notes were reviewed while still fresh, to assure that they would be understood later, because of concerns about poor handwriting and the use of a somewhat "cryptic" shorthand system.

6.6 User survey

During the early project planning phases, a user survey was envisioned as being a useful method to quickly gather user input, from many users. This subjective opinion was formed mainly from the frequency of recommendations for surveys (for quick, extensive input from many users) that was seen often during the compilation of the literature review.

There were multiple reasons for using a survey. First, it allowed information that directly addressed both of the research questions, to be gathered, from, importantly, a large number of actual Lydhør users. It also offered an opportunity to assess the intensity of problems identified from the ratings review, and application testing. It also had potential to uncover yet unidentified problems, and to learn which features users were dependent on or had a preference for.

The interview had a strong impact on the direction of the research, so the survey would now also offer a chance to balance unintentional bias that may have developed after the interview. Most importantly, the design of the prototype that would be utilized during user testing, and the design of the user testing itself, was intentionally dependent on the survey results. These were not developed until after the survey results were fully assessed.

Because it was ultimately going to be one of very few opportunities to gather user input,

substantial effort was implemented into the survey design. In particular, the accessibility of the survey tool itself, and the design of the questions were critical factors, that if done improperly, could nullify the usefulness of the survey.

During design of the survey, two books were consulted frequently, and offered nearly step-by-step instructions for the entire survey process, User Centered Design and Research Methods (Lazar, 2010; Lowdermilk, 2013). The interview, user ratings review, and the research questions themselves affected the design of the survey.

The survey software chosen was Nettskjema[19] software. Other survey software was researched, but ultimately not used, due to concerns about accessibility gaps in the software and other technical matters. A fellow researcher had been surprised to learn that access to results from one particular "free" survey software was in fact not free, after a certain threshold of total respondents was exceeded. To be safe, the software provided by the University was utilized.

The intended method of circulating the survey invitation was to have the NLB circulate an invitation to participate and a link, to all Lydhør user's email addresses. At the last moment, they instead decided to post a link on their Facebook site, which netted less than 40 responses. When prompted they reposted the Facebook notice, but participation did change substantially.

In order to increase survey responses, several organizations that would have members that qualify as "print disabled" were contacted. One organization, Dysleksi Norge (a Norwegian Dyslexic association), did respond and had forwarded the invitation to their entire email list. Five other organizations such as Norges Blindeforbund (the Norwegian version of the Federation of the Blind), an association for the elderly, and various others did not respond, but may have possibly spread information about the survey.

As a result, respondents to the survey grew by almost a hundred additional participants. Notably, the number of respondents indicating that they *did not* use Lydhør also grew.

6.6.1 Survey accessibility design considerations

The Nettskjema software was checked for accessibility, by performing a guideline review with WCAG 2 guidelines. The review was not exhaustive. Some failures to conform with guidelines were noted.

A researcher familiar with the process of making surveys accessible advised that custom modifications had to be made by programmers to make survey software fully accessible. This process was expensive and was not tangible for a project of this scale, so the attempt at making the survey fully accessible had to be scrapped.

In any survey that uses any software, the use of scale questions (such as Likert scale) reduce accessibility/usability for screen reader users. They are essentially forced to tab

[19] https://nettskjema.uio.no/

through each potential answer, for example 1 to 10, so 4 10-scale questions can create a 40-step process. Qualitative "type" answers offered better potential for useful information, so Likert scale questions were avoided.

6.7 User testing

Whether in a professional business environment or in an academic research environment, a product can be designed and created to incredible standards, only to fail its core purpose, to be a useful and usable tool for the end users. If the final product is "functional" but is confusing or unpleasant to use for the intended audience, or both, then it is has failed to meet an implicit core requirement.

In this case, the "product" mentioned in the teaser above is the information gathered from all of the previous research activities. Although substantial information would have been gathered during the earlier research activities, it risked being misinterpreted. These activities would have all gathered qualitative and subjective inputs, from various, diverse sources. An improvement for one user could detrimentally affect another user.

The decision to perform user testing was made early in the research process, based on information learned while studying "user centric" techniques. It offered the ultimate test for subjective assumptions made. The vision, when developing the project plan, was to create a prototype, with its structure being determined from earlier user inputs, and then to test the structural decisions/assumptions, with participants from the print disability group.

The design of the prototype is complex enough to warrant a discussion of its own, in another section of the thesis. However, in short, it was designed to attempt to correct problems with Lydhør that were to be identified during the ratings review, app testing, interview, and survey portions. It also intended to preserve desirable aspects.

The prototype supplied a tool for user testing, which was influenced by the same factors. In the case of this research, user testing offers the final concrete measurement of whether or not of research questions were answered. It also, metaphorically speaking, improves the quality of the "answers" of the research questions, by refining the data yet another time.

Much of the design of the user testing was influenced by the book Praktisk Bruker-Testing (Practical User-Testing). This was in essence, a start-to-finish guide written by authors familiar with user testing related to HCI. This book offered procedural guidance and wisdom about things that could go wrong. Although rarely cited in this text, this "manual" was often consulted and influenced the design and flow of the testing (Eli Toftøy-Andersen, 2011).

Participants were given tasks to perform, in two broad categories. The tasks were to utilize various functions within the app, and to perform searching tasks, using both the Lydhør application and the prototype. The reasoning behind the choice of activities that were performed during user testing, will be clarified as the results of the all of prior

research methods are unveiled, because they are intensely dependent on the results from each of these. Full appreciation of the design of the user testing is only achievable after the results are discussed.

Volunteers were offered to their choice of the location of the testing. Suggestions were made that they could visit the university and participate in testing at a lab there. They could also choose to do testing at home or at a quiet location of their choice, such as a library. They used their own mobile device, and were asked to have one book already on their phone. One of the tasks in the user testing involved comprehension of the downloading versus streaming process, so this request was intentionally vague to accommodate the testing.

The tests themselves were pre-tested three times before the first official test began. The original intention was to time the (identical) tasks, and compare them. This would provide concrete information as to whether or not the prototype improved performance versus Lydhør. During pre-testing, the realization was made that timing tasks was intensive, and it stole focus from the main goal of testing, getting input from the users.

Timing the tasks was dropped. The users instead performed identical tasks on with both applications. They were observed performing the tasks, to gain insight about the process. Various qualitative inputs were solicited.

The participants were spoken to in their native language, and the prototype and user testing tasks were also presented in Norwegian. Participants were asked to indicate if they misunderstood (spoken information), because this was not the researcher's native language. The reasoning for not conducting testing in English, which is the researcher's primary language, (most of the residents of Norway are fluent English speakers), was that they might feel less comfortable during testing. If they were not eager to speak English, they most certainly would offer less input.

Participants were told at the beginning of testing what would occur and why, along with some background information about the research. They were also told that they could stop at any time they wanted, and additionally, that the information from the testing session could be destroyed immediately if they so desired.

They were encouraged to think out loud during testing. One researcher asserted the usefulness of this technique, particularly when trying to ascertain qualitative information about personal preference differences in task performance. She also asserts that she frequently had attained more details from open conversations versus structured questions, although researchers should be aware that this technique can create a heavier cognitive demand (Charters, 2003). To avoid cognivite "overload," they were directed only to think aloud when they encountered something frustrating or something that they really liked.

During testing, if participants appeared to be encountering problems when completing a task, they were often asked at that moment what was happening/why it was

problematic. Notes were taken typically at the end of each task completion, which often gave users short breaks. The path that users chose to get to a certain function was written down for the tasks that were designed to learn navigation and menu preferences.

There were mild variations in the tasks performed as the testing progressed, as testing had shown that a test was flawed or unnecessary. The reasoning for the changes is discussed in the results section. The "script" of task questions is in the appendices.

Note that testing never utilized all of these tasks in one testing session. There are several search tasks in this script, but typically only three or four were performed. During pre-testing, books that had been in the database in prior "pre-tests" were suddenly not there, which created an unsuccessful search, even if the terms were entered perfectly. Therefore several alternates were available, in case this occurred during actual testing (which it did).

After the tasks were performed for both apps, the users were questioned. The questions typically sought more clarity to problems that were observed and asked the testers to compare the experience with both Lydhør and the prototype. Then participants were asked general like/dislike questions regarding both Lydhør and the application. During the final portion of testing, participants were given a blank slate, and the opportunity to express whatever they wanted to say about either application.

6.8 Data analysis methods

Analyzing the data accrued during the various phases of the research activities, required multiple techniques. Each individual situation required a unique solution, to interpret the data gathered, most of the time with the purpose of applying the data later, to move towards satisfying RQ2.

Finding a one-size-fits-all method for this entire project, in the existing research base, was impossible. Universal design of IT is a young concept, and researchers are "feeling" their way through the process. This is exacerbated by the fact that this project contained several data collection processes (and some design processes).

To address this dilemma, various techniques were applied at different times. Because this project is inherently design related, design evaluation methods were used in some portions. A static analysis, which was created as a technique of data analysis in the larger design science research method, was used for the application testing. Hevner et al., describe this simply as a "structural analysis of an artifact to determine static characteristics" (Hevner, March, Park, & Ram, 2004). Simultaneously, during application testing, the data from the ratings review was "re-analyzed" to provide a visual representation, but this was a simplistic process that needs no data analysis method "tag" other than a simple; read, assess, identify.

For analysis of the qualitative survey data, the grounded theory, which Lazar describes as an "inductive research method" was used (Lazar, 2010). However, rather than gathering

data to create a “theory” as the technique is typically used for, many mini “theories” were formed to be tested in the prototype design/user testing phases. One example of a mini theory is: “It appears that screenreader users do not encounter a higher proportion of X problems than those that don’t use one, with widget X.”

Functional black box testing described in the design science research method was used for data analysis of the user testing results. The authors describe it as “Execute artifact interfaces to discover failures and identify defects (Hevner et al., 2004).” However, rather than simply executing the program, this was modified to, “observe a very specific group of users utilizing the application.” The data analyzed in this case was the interactions of users with the software, and their subjective assessments.

6.9 Ethical considerations

The organization, in charge of assuring that research is performed according to ethical standards, in the country where this document was composed, is the Norwegian Centre for Research Data, or the NSD[20].

Research that collects personal data, or data, that when combined can identify participants, must be approved beforehand with the NSD. The survey results were however, anonymous, so the survey did not require an application be filed with the NSD. The survey did solicit participants for the user testing, which would necessitate gathering of contact information.

However, the contact information was not tied to the survey results. In order to gather the contact information, participants had to first submit their survey answers, and were then redirected to another website, which was not linked to the survey data.

During user testing, no video or audio recording was used, only notes were taken. No personally identifying information was recorded on the notes. This process, when assessed with the NSD’s online tool[21] that determines whether an applying with them is necessary, again, did not require applying.

Consent was gathered from user testing participants, as required by NSD, but also as an opportunity to inform them that they could stop testing at any moment for any reason, and that they could additionally have data collected during their testing session destroyed, if they so desired.

Because this research involved humans with physical and mental capability limitations, the activities were designed in a manner that preserved anonymity. The survey did not question which “print disability” a participant had. In addition, the gender of the participants in the user testing and interview, may, or may not, have been switched in the reporting. The gender of the participants has no significance in the context of this research.

[20] http://www.nsd.uib.no/nsd/english/index.html
[21] http://www.nsd.uib.no/personvernombud/en/help/research_methods/online_surveys.html

7 The Lydhør app

Lydhør is an e-reader application, which is used to listen to books. It was specially designed to be used by patrons of the NLB, which must have some form of print disability to qualify to use any of their services. The app was designed to be usable by all of the NLB patrons, although it would unfortunately exclude patrons that are both blind and deaf, from utilizing audiobooks. Text-to-speech titles could still be accessed via a braille printer.

Lydhør can be used to listen to books, or to find new books. Books can be checked out, and immediately accessed. In addition, books can be placed on a wishlist, as users can only check out a finite number of books. The book files can be quite large, which could also be another reason to use the wishlist function.

There are limitations as to how many "copies" of a book can be loaned out at a time, so sometimes books can only be reserved and then users must wait until a copy becomes available.

The app is only available in the Google Play Store[22] and Apple's iTunes.[23] This means that Lydhør can only be accessed via mobile phone or tablets (android or iPad). Lydhør cannot be used on personal computers (books at the NLB can be accessed via PC, but with different software).

Some of the titles available have been read aloud by a person and recorded, and come in audiobook format. Others are read via synthetic text to speech, which reads text format aloud. When synthetic speech is used, the words are displayed on the screen as they are being read, and can be highlighted, and displayed using various customizations.

Generally, the structure of the books is preserved, even when a book is in audiobook format, to assist navigation within the books. There are multiple ways to navigate within a title available for users.

Since the application is available on different platforms (Apple IOS and android), there are different variations of Lydhør. The most significant difference between the versions are the presence of onscreen button on some versions versus a physical button used with the other version. The versions addressed in this research were Android versions.

The interface can be explained in different fashions. For this discussion, the different screens or "views" are discussed, although the information could be presented in a different manner, such as by function. The discussion starts with the bookshelf screen, and functions that directly affect the screen are discussed afterwards, such as how the books are added to the screen.

7.1.1 The bookshelf screen

Lydhør uses one central screen to access most features of the application. These features

[22] https://play.google.com/store/apps/details?id=com.nlb.android&hl=en
[23] https://itunes.apple.com/no/app/lydh%C3%B8r/id603112511?l=nb&mt=8

include the various settings the application has, the function that reads the books, and the search interface. For the purposes of this discussion, this “main” screen will be called the bookshelf screen, because that is essentially what the screen is, a location for displaying the books that users have checked out, reserved, or added to the wishlist.

The elements on the bookshelf screen are four tabs, a search button (icon), and another button that serves the purpose of a menu (in some versions this button is accessed via a physical button on the phone). The bookshelf screen is shown below in Figure 7.1.

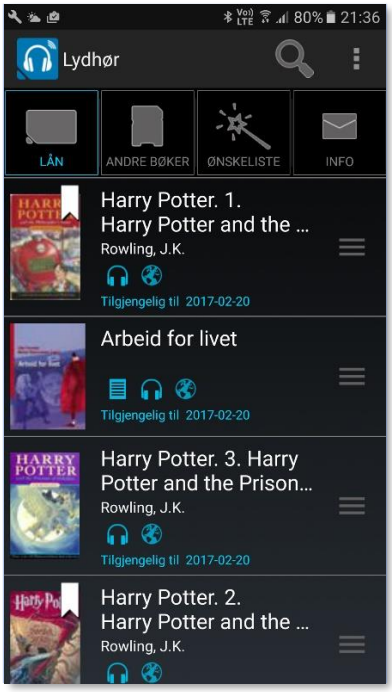


Figure 7.1: Bookshelf Screen

The four “tabs” seen near the top of the screen in Figure 7.1 serve the purpose of dividing the bookshelf information for display. The Lån (Loans) tab is the first tab in this section of the application, and it shows titles that have been checked out. The second tab, “andre bøker,” which translates to other books, which was originally thought to have no purpose, is the location that the user manual is supposed to be stored. The third tab “ønskeliste,” (wishlist) is where books that have been added to the wishlist are stored, and where reserved books are stored. No purpose has been found for the fourth, “info” tab, but it may have a purpose.

Books that have been downloaded, reserved, or added to the wishlist are displayed in large panes, where up to four books can be seen per screen. Each book “pane” has different information about the book. The cover is shown when available, or an icon that says no cover is shown when the cover art is unavailable. The title and author (when available) are printed in text there. Under these are icons which give information about the book’s format (synthetic text to speech or audiobook), and information about whether the book has been downloaded (if not there is an icon to indicate that the book can be streamed). There is also a “hamburger menu” in the right side of the book

pane. This function of this is discussed at the end of the search results section because of other related information there.

The icons seen above in Figure 7.1 are enlarged below in Figure 7.2 and Figure 7.3. The first icon, that appears to be a sheet of paper, indicates that the book is a text-to-speech book. This icon is simply removed when a book is in audiobook format. The headphone icon appears with every book. It probably indicates that the book can be listened to in audio format (although this was never confirmed). The last two icons are mutually exclusive, the globe in Figure 7.2 indicates that the book can be streamed. If the book has been downloaded, the globe is replaced with a downwards facing arrow, shown in Figure 7.3 . While a book is downloading, the display changes, as shown in Figure 7.4, to show the progress of the download.

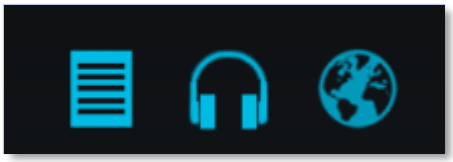


Figure 7.2: Book that can be streamed

Figure 7.3: Book that has been downloaded

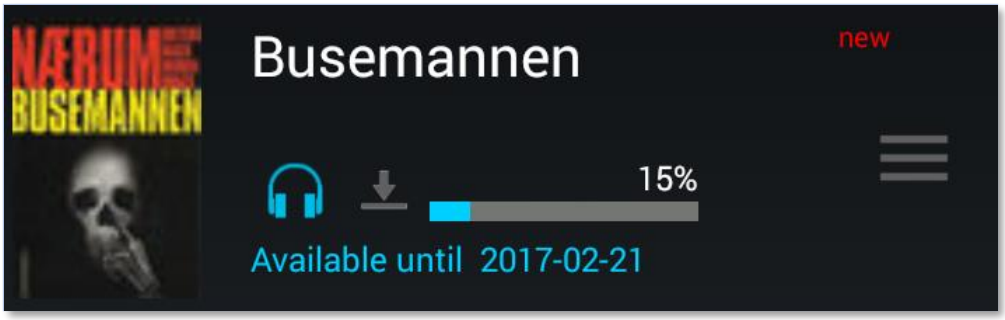


Figure 7.4: Download Progress indicator

7.1.2 The search screen

There is only one way to find books to read in the Lydhør app, and that is via selecting the magnifying glass icon that is visible in the top right of Figure 7.1. There is no other method to activate the search process within the app. This icon is only visible when the user has the Loans tab activated, and disappears when the user has activated any of the other three tabs, as shown below in Figure 7.5.

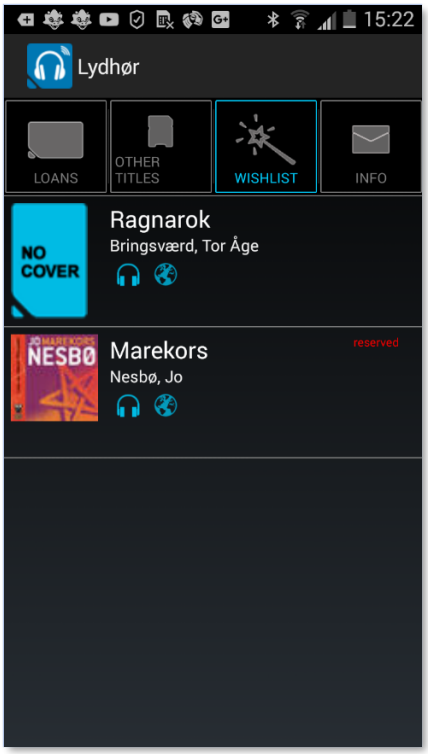


Figure 7.5: Wishlist tab

Note that the page elements in this picture are in English (versus Norwegian shown in Figure 7.1). The application apparently sets it language based on the language of the phone. Two different phones were used during testing, one where the default language is English, and the default language on the second phone is Norwegian. The screenshots were generally taken during application testing on either phone and that is why the application is shown with two different languages.

The Lydhør search function is a basic search function. There is no potential to “browse” the library’s selection in the app, in any manner. Activating the magnifying glass brings up the phone’s keyboard, and the search terms can be entered into the small box at the top of the, as seen in Figure 7.6 below. The search terms are committed via the magnifying glass, that is visible in the bottom right hand corner of the keyboard.

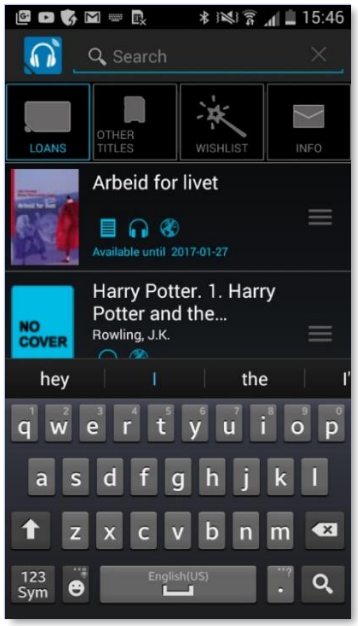


Figure 7.6: Search Screen

When a search returns no results, a small message stating “Displaying 0 search results” appears under the search field and the keyboard disappears, as shown in Figure 7.7 below.

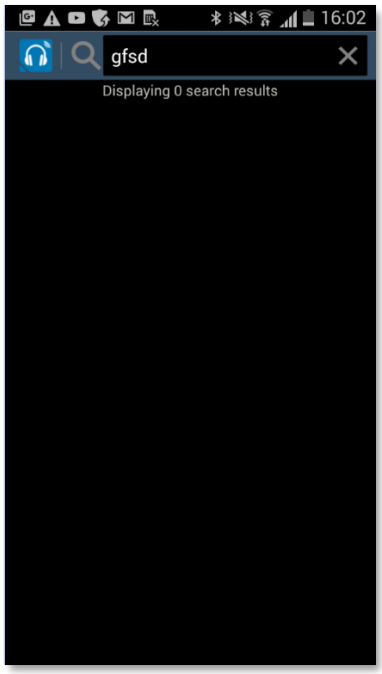


Figure 7.7: Null search results

Searches that do return results lists them in “panes” on the screen. Each pane contains a small picture of the book cover (when available), and the title and the author (again, when available). The search results are displayed in a nearly identical manner as the bookshelf screen. The tabs that are atop the bookshelf screen are not shown. In their place is the box with search terms and a small message displaying how many results the

search found. This is followed by the books that were found during the search, as seen in Figure 7.8 below.

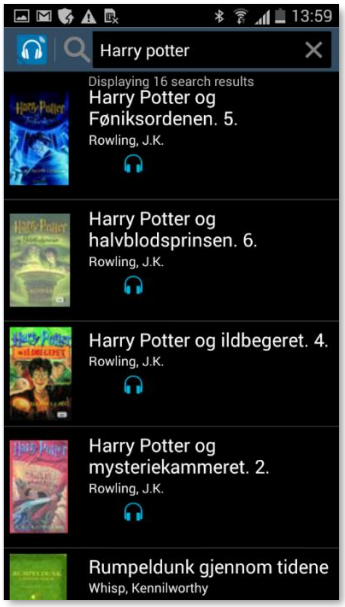


Figure 7.8: Search results

A desired book can be selected by pressing on the corresponding book pane. The majority of the screen is darkened and three buttons are shown. When the book is available for checkout the first option available is, to check out the book. This text on this button is changed to “reserve this book” when the book is unavailable. The second button allows the book to be added to the wishlist. And, the third button is only shown when a preview is available (a snippet of the book is read when this button is activated).

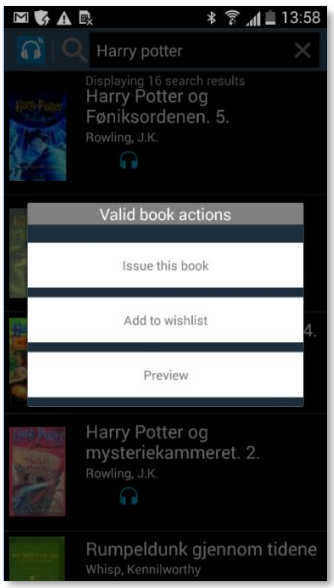


Figure 7.9: Book actions

One item of note is that after a book is issued or checked out via use of the checkout button shown in Figure 7.9 above, the book is not physically downloaded to the phone.

It is available for streaming, but the files are not stored locally, and one must have internet access to stream a book from the NLB servers.

In order to be able to listen to a book, without access to internet, the book must be downloaded. This is accomplished via the hamburger menu, that is available on the individual book panes on the bookshelf screen. Activating the hamburger menu gives an option to download a book (if it has not been downloaded), and the option to return a book when users are finished with it as seen below in Figure 7.10.

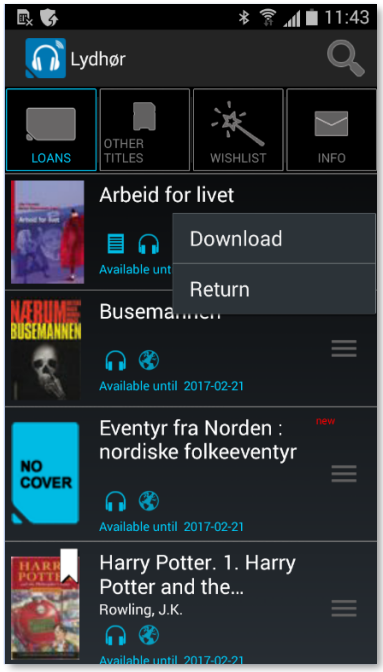


Figure 7.10: Hamburger Menu

7.1.3 Play screen

In order to listen to a book, one can “select” a book by touching or tapping a book pane in the Loans tab. A new screen is then opened, which is displayed while books are being listened to (or paused). The book is automatically read, from the beginning. If the book was created with a human voice, the screen will display some chapters and highlight the chapter that is being read as shown in Figure 7.11 below.

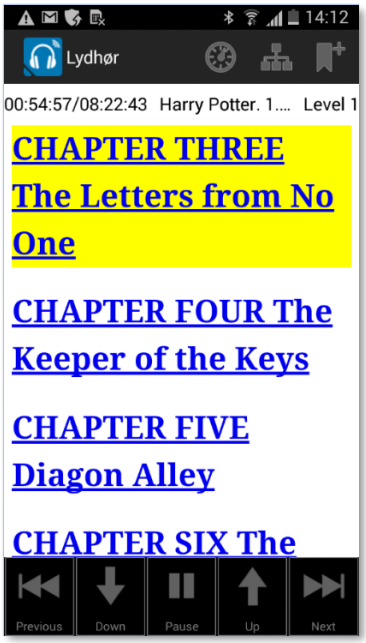


Figure 7.11: Play screen audiobooks

If the book is being read by a synthetic text-to-speech book, the screen will display text and the text that is being read is highlighted as seen below in Figure 7.12. Note that the text size, highlighting colors, background colors and more be can be modified in the application settings, which can improve accessibility in certain situations.

Figure 7.12: Play screen synthetic text-to-voice

The two play screens contain the same set of controls. The controls on this page are likely to be the controls most used in Lydhør, because most of the time spent using the app likely occurs here, as some books take hours to listen to from start to finish. Since many of the controls are “navigation” related a short discussion about in book navigation is presented before explaining their function.

Many books are read from beginning to end, in a linear fashion, such as suspense fiction book. However, many books are never read from beginning to end, and only portions of the book will ever be accessed. This is often the case with school textbooks, and the NLB is one of the government organizations that share the responsibility of producing textbooks for blind/limited vision users.

Therefore, being able to “navigate” effectively within a book can sometimes be critical. As discussed in the literature review, applying a proper structure to a book increases accessibility for blind/limited vision users. It is also helpful for all users, and the reasoning is simple. If a user must access chapter 10 in a complicated physics textbook, and has only fast forward and rewind controls, the process can be lengthy.

All of the controls seen in Figure 7.12 above, except the first one that resembles a car speedometer, are navigation related (the first button adjusts the speed at which the screenreader reads the books). To avoid redundancy, they are discussed later in detail in the testing results section.

The white stripe that is located below the top row of dark buttons provides some information about the book. The first set of numbers on the left display the current “time” position with a book, and the total length of the book. In the center, the title is displayed (the portion that fits in the space). The last item visible that says “jump 30 seconds” is the skip interval, which will be explored in the cognitive walkthrough results.

7.1.4 Login and settings menu

In addition to the screens already mentioned, Lydhør has a login screen, that is self-explanatory, and only requires discussion in the results, related to problems there.

There are also two separate settings menus in Lydhør. One is a “global” settings menu, and the other is a menu available only when listening to books. The main or global menu has many features that are geared towards enhancing accessibility, like changing the text size or changing the background or highlight colors. There are also some practical settings that allow users to choose whether for example they want to stream books using their mobile data, or auto-download books.

The other settings menu, is only accessible via the book play screen. It contains a link to the main settings menu. There is also a search function that can search contents of a book. This is less effective with audiobooks as the only “contents” available are usually just chapter headings, not the book text. The last feature of this settings menu is the sleep timer, which can be set to stop the book from reading after a certain amount of time.

8 Results phase 1-Ratings review, application testing, informal interview

8.1 Ratings review results

This simple exercise provided key information about user sentiment, both what users liked about the application, and potential problems with the application. For example, users were having problems finding their way to the position where they had stopped reading a book (bookmarks were disappearing). One user would like to have a sleep timer (the app has a sleep timer, but the user was unaware because they had never found it).

Some users were left wondering what had happened when there was no feedback for an action. Technical problems were causing users to have to uninstall and reinstall the app and download the books again. When searching for books, users were provided only with the title and author in the search results, and would like to have more. One user thought that the "skip" function was broken, but it was apparent by his answer that he didn't understand that the skipping can be modified to skip at different intervals.

This activity supplied an invaluable reference of factors to be aware of and look for when testing the application, which the results are described in the next section, and when constructing the survey.

8.2 Application testing results

Two distinct tests were performed on the Lydhør application, a guidelines review, and a cognitive walkthrough, as discussed in the research methods section. The purposes for doing this were: learning the application, identifying accessibility failures, and identifying usability problems. The results of the guidelines review are discussed first, followed by the results of the cognitive walkthrough.

8.2.1 Guidelines review

The sets of guidelines used to assess accessibility were the aforementioned WCAG guidelines and the WCAG mobile recommendations. The WCAG guidelines, which were designed with personal computers in mind, had to interpreted as to how they would apply to mobile applications, when the mobile recommendations were silent about a particular topic.

Although these are not perfect circumstances in which to guarantee accessibility, these are the only "rules" that exist. On a positive note, since the mobile guidelines are being developed, some of the information gathered in this testing could serve to improve them.

The guidelines review resulted in identification of very few accessibility issues, which was a positive finding, because it means that the current version of Lydhør has a "high score" for accessibility, which is great for Lydhør users.

8.2.1.1 An important guideline breach

There is one issue where the design of Lydhør *appears* not to meet WCAG guidelines. Because it affects the entire application, it is presented first. WCAG guidelines 1.4.4 state "Except for captions and images of text, text can be resized without assistive technology up to 200 percent without loss of content or functionality[24]."

At this point, the mobile guidelines only address ensuring that operating system level accessibility features are not blocked. Users should be able to select system wide text size settings and zoom by triple tapping[25]. These features are not blocked in Lydhør.

In fairness, most of the text in the Lydhør application can be enlarged via the settings menu. But text in the Lydhør application menus and on some interface components cannot be scaled, pinch zoomed, or enlarged. The mobile recommendations also address the pinch-zoom gesture. However, the applicable rule states not to block the gesture in mobile *browsers* and is silent about mobile *applications*[26].

The pinch zoom gesture is technically not "blocked" to be fair, it is in this case not enabled. However, because it results in the same lack of accessibility, it is considered an accessibility breach. This combined with the first statement about zooming text to 200%, means that users will not be able to see some of the text in the application.

The reason the wording "appears not to meet guidelines" was used instead of "does not meet," is because there was no known method to measure scalability on mobile devices, when testing was performed. Instead, the Lydhør text was visually compared against a Microsoft word document and a Google Chrome browser zoomed at 200%, the only known option available.

8.2.1.2 Login screen accessibility

The first time Lydhør users access the Lydhør application (or if they have to uninstall/reinstall the app) they must log in. The required credentials are simply a username and password that is issued by the NLB. It is simple, and is shown below in Figure 8.1.

[24] https://www.w3.org/TR/WCAG20/
[25] https://www.w3.org/TR/mobile-accessibility-mapping/
[26] https://www.w3.org/TR/mobile-accessibility-mapping/

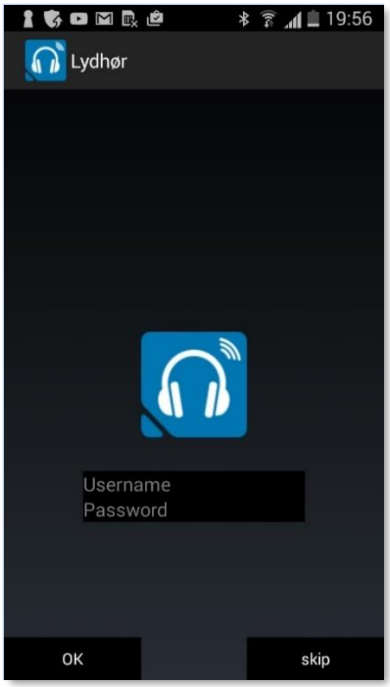


Figure 8.1: Login screen (vertical orientation)

Practically speaking, users would typically only encounter this feature a limited number of times. However, particular attention was paid to this screen, because of the potential breach of accessibility that would arise. In essence, the application is not accessible for use, when the login process cannot be successfully completed.

Multiple points from the WCAG guidelines or WCAG mobile recommendations were applicable. When a guideline was extended, or applied to mobile in the mobile document, that item was typically given predominance.

The contrast ratio of the login screen appeared to be non-compliant with WCAG text contrast requirements, because of the appearance of the username and password textboxes. Currently, there is no software available that checks the contrast ratio on mobile screens, except for the possibility of checking the contrast ratios of elements of websites that are accessed via mobile.

In order to test the contrast ratio, screenshots were taken, and they were assessed on a laptop, utilizing the color contrast analyzer application (Vision-Australia, 2016). Note that the testing did not give consistent results, because the testing software acquires a color sample size that is approximately a pixel. Although difficult to see in Figure 8.2 below, the text sampled from the screenshot from the mobile phone, was slightly pixelated (with mild variations in the color of the pixels), and therefore gave the varying results.

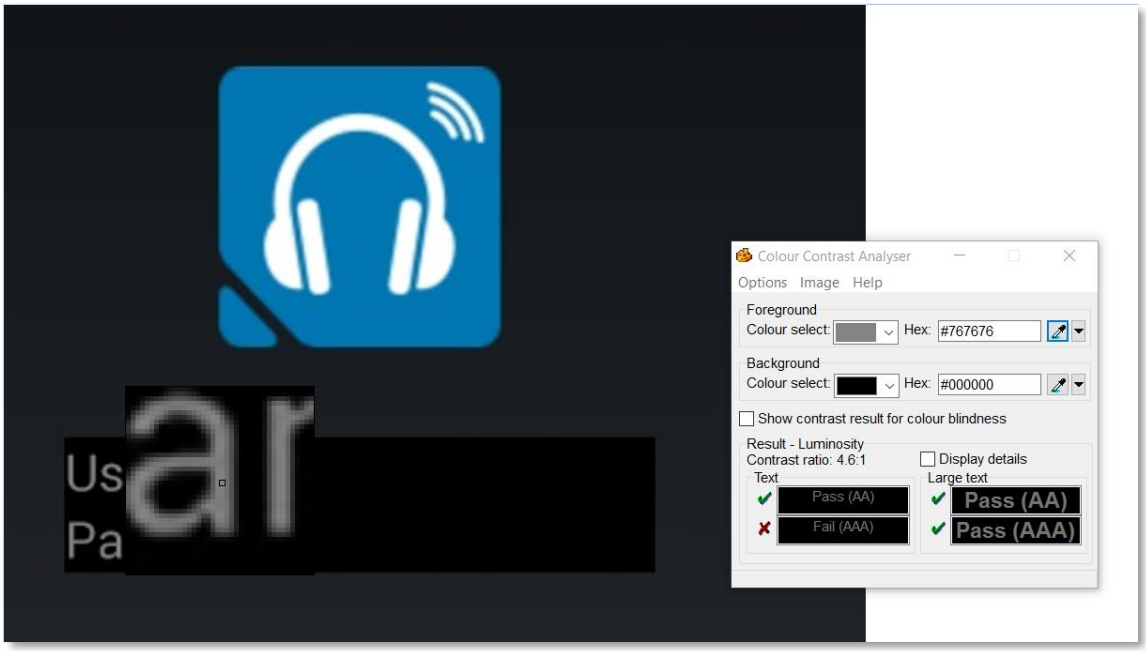


Figure 8.2: Contrast Check

The weakest contrast ratio result was 4.6:1 and the strongest was 6:1. Both measurements passed AA Level contrast requirements and did not meet AAA level requirements (W3C, 2008). The WCAG mobile recommendations expand on this point, in relation to mobile devices (Guideline 2.3 Contrast). This particular guideline points out that mobile phones have a smaller screen size (which typically have smaller text sizes, and smaller text would require a higher contrast ratio), and the fact that mobile devices are frequently accessed in lower visibility scenarios, such as outdoor use[27].

The recommendation in this section only refers to whether the lower standard contrast ratio of 3:1 should apply, and does not recommend different contrast ratios. Since the mobile guidelines are a work in progress, there is a possibility that the two factors mentioned in 2.3 may culminate in the recommendation of higher contrast ratios in the future. In the case of Lydhør, there should be no drawbacks from using a high contrast ratio, and utilizing a higher ratio only serves to increase visibility.

Another, possibly even more significant problem, was noted in relation to contrast. This regards the transition from the textboxes to the surrounding background. The contrast ratio between the textboxes and background is 1.3:1.

WCAG guideline 1.4.1 is an applicable guideline regarding these elements. The guideline states that “Color is not used as the only visual means of conveying information, … or distinguishing a visual element.” In fairness, screenreaders are properly presented information about the textboxes (they have been properly labeled so that screenreaders announce that the element is a textbox, password input). Another means of identifying the element is available, therefore the guideline is properly satisfied.

[27] https://www.w3.org/TR/mobile-accessibility-mapping/

However, for users that are not utilizing screenreaders, the guideline is not fulfilled. Those relying on vision to navigate through the application have only one method of identifying the textboxes, and color, is, in this case, the only means of distinguishing the visual elements (textboxes). You can see in Figure 8.1 that there is very little visual information to show that the text boxes are there.

The WCAG mobile guidelines expanded on the WCAG guidelines by adding “touchscreen specific” guidance (3.2 Touch Target Size and Spacing). As of now they are not requirements and are labeled as “best practices.” (W3C, 2015).There are two recommendations that are applicable regarding the text boxes on the Login screen.

One is that the width and height of “touch targets” are a minimum of 9 mm. The other recommendation states that when targets are sized close to the minimum requirements, that they be padded “ by a small amount of inactive space” (W3C, 2015). The textboxes on the login screen are 4mm high, and have zero “padding” between them. This is only a variance from a recommendation but can reduce accessibility, and is therefore included in these results.

8.2.1.3 Accessibility Issues: buttons

There are three buttons that allow a book to be “issued,” added to the wishlist, or previewed. Identical buttons are also used with the wishlist section for choosing actions for the books that can be reserve. The discussion below applies to both. The buttons can be viewed in Figure 8.3 below.

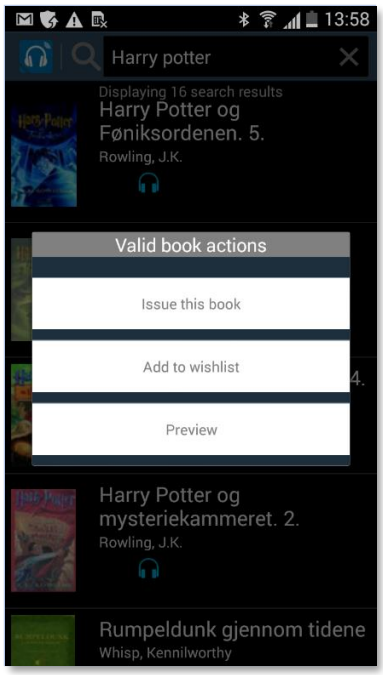


Figure 8.3: Inaccessible buttons

The text on the white buttons shown in Figure 8.3 does not meet WCAG guideline 1.4.3. The contrast ratio measured varied between 3.2 and 3.5, neither of which meet the minimum contrast requirement. The “padding” recommendation discussed previously in

reference to the login elements, is also applicable here.

8.2.1.4 More contrast related issues

One additional compliance failure related to contrast was identified on another set of buttons. You can see in Figure 8.4that the text on the bottom of the screen on the buttons has a low contrast ratio.

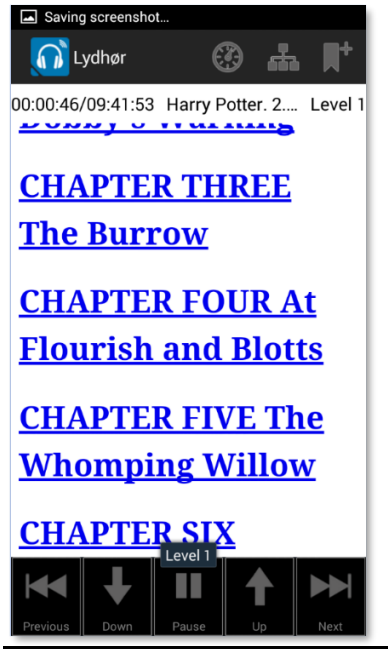


Figure 8.4: More Low Contrast Buttons

The text on the buttons does not meet WCAG guidelines of 4.5:1 relating to contrast 1.4.3 (W3C, 2008). The contrast ratio is quite low 3.4:1 for these buttons.

There is another item of interest here, which is the controls/icons that are at the top and bottom of the screen. These controls have an identical contrast ratio as text, but the guidelines are not applicable because they apply to text. If an application user has limited vision, but does not use a screenreader, identifying the controls can be problematic. In the discussion section of the thesis, this topic will be explored further.

8.2.1.5 An opportunity to increase accessibility

Figure 8.5 below shows a small portion of the search results screen, which is similar to the bookshelf screen. The designers have surrounded each individual book with a thin, grey border, to visually separate each individual book.

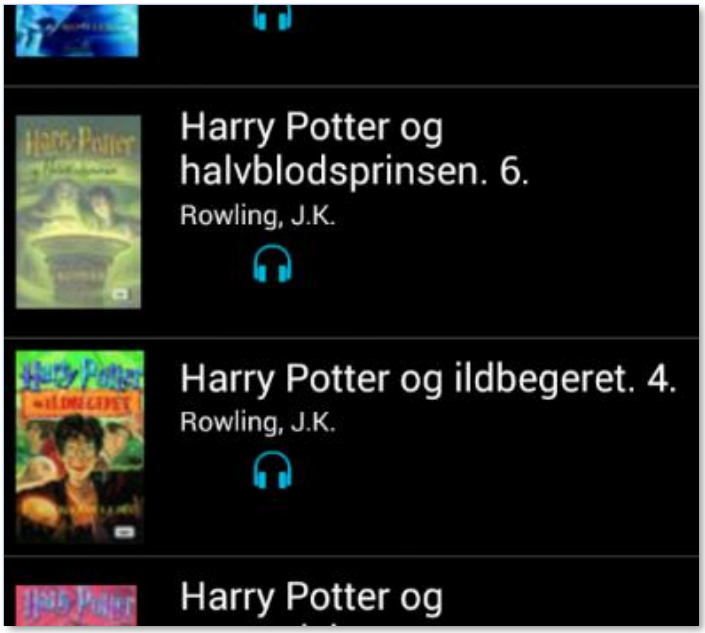


Figure 8.5: Bookshelf closeup

This does not violate any guidelines, but accessibility can nevertheless be increased, if the contrast ratio of the border-to-background is increased (by using a different color), and the thickness of the line is increased.

8.2.2 Cognitive walkthrough of Lydhør

During the guidelines review, only a small number of accessibility problems were identified. Accessibility had been prioritized, but had the thesis given most of the focus to these items alone, this would have resulted in an extremely thin thesis.

This was noted at an early point in thesis development, after the guidelines review was completed, so attention was shifted to usability. The cognitive review addresses several usability issue results.

8.2.2.1.1 Login screen

Returning to the login screen, a problem occurred when accessing the application in landscape mode. When Lydhør was accessed in landscape mode on either phone, the buttons from the bottom of the screen cover the textboxes as seen in Figure 8.6.

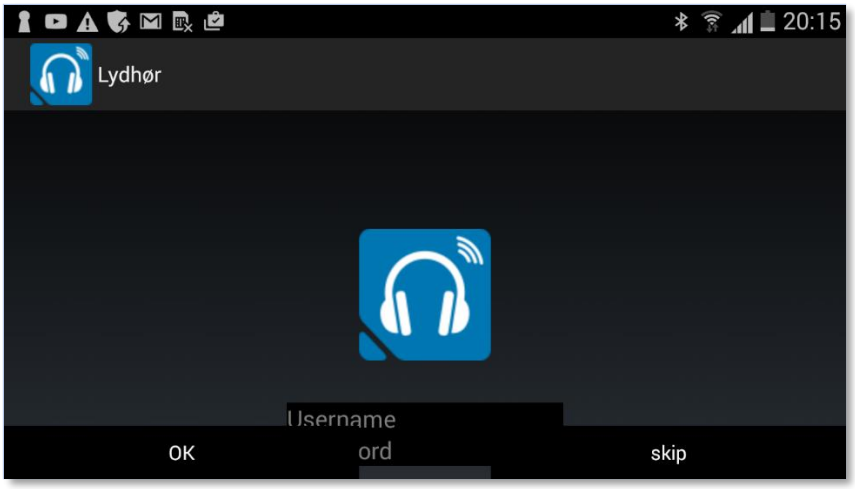


Figure 8.6: Login Landscape mode

Note also that the skip button is unnecessary here. Skipping the login process will allow users to open the application, but not use it.

The large icon in the screen can also be problematic for users with limited vision. It is not functional, but its size could draw attention to it, and when it cannot be seen clearly, users may believe that it is a control that is not unannounced by the screen reader.

8.2.2.2 Bookshelf screen

During the cognitive walkthrough of this screen, the purpose and meaning of the tiny blue icons seen in Figure 8.7 was not readily apparent. Even after substantial investigation, they were still somewhat mysterious.

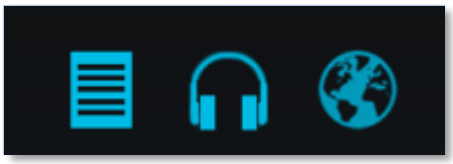


Figure 8.7: Indicator icons

Even after the cognitive walkthrough, there was still substantial confusion surrounding the last icon, the globe. Notably, it was during a subsquent session with a screenreader (returning to an item for clarity) that the meaning of the icons was discovered. In this case, the screenreader experience is more informative than the visual experience. The distinction between reserving and checking out a book was also not understood at the conclusion of the cognitive walkthrough. Some of the user rating review comments also expressed confusion about these items as well, which is another reason why extra scrutiny was applied here.

Although it is technically a component of the settings menu, an attempt to utilize the user manual for clarity about the icons, revealed that the user manual, although offered, either does not work, or has not been created.

8.2.2.3 *Search function Interface and functionality*

When testing the search interface some items were noticed that can improve usability, that were not quite egregious enough to be considered accessibility transgression.

The first item noticed when accessing the search function from the magnifying screen icon on the bookshelf screen, was what occurs after it is pressed. There is substantial feedback because the onscreen keyboard appears.

However, the fact that the search field has appeared might not be fully evident to users, especially for those with limited vision. This change with the header is subtle, a small search field appears on the top of the screen, replacing the text the “Lydhør” header text and two small controls, and is shown in Figure 8.8.

Figure 8.8: Pre-search & activated search

Another confusing situation was identified involving when a search returns no results. A message stating “Displaying 0 search results,” that is barely visible, appears under the search field and the keyboard disappears. Notably there are no buttons on screen to “return” to the previous screen. The only possible way to return to a previous screen is via the “back” button on the phone. The only other action possible at this point is to “push” the X in the search box, which will clear the search text, and returns the keyboard, which does allow for a new search. See Figure 8.9 below.

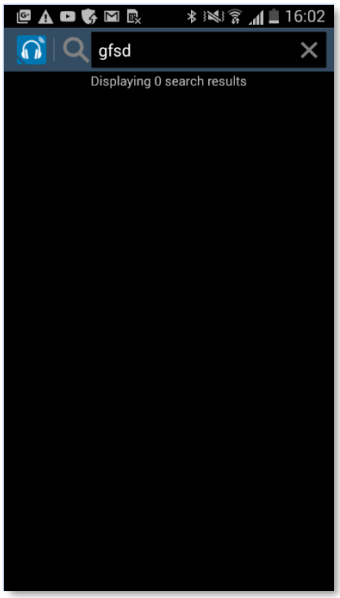


Figure 8.9: Null Results

The actual search function itself was also tested, promptly party due to multiple indicators of dissatisfaction read in the rating review. These comments were generally pertaining to dissatisfaction about the spelling precision required to find a book, and they were accurate.

In fairness, the NLB website recommends using their website for browsing, and even states that the website has a "better search function[28]," which it does. If logged in, the books that are checked out or reserved, are automatically synchronized to Lydhør, from the site.

8.2.2.4 Play Screen

Books that use synthetic text-to-speech (not audiobooks) will display and highlight text as it is being read. With many books this function works properly. There are occasions where the display is very erratic, see the examples, Figure 8.10 and Figure 8.11 below.

[28] www.nlb.no

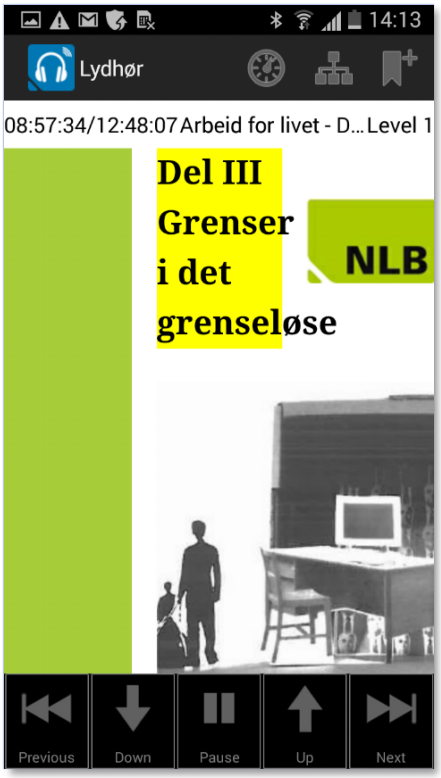


Figure 8.10: Text displayed as it is read

Figure 8.11: Text displayed as it is read (2)

When testing with a screenreader, there were several instances where the screenreader speech and the book speech overlapped one another, which is the same as having two

people talk to you at the same time.

There are four ways to navigate through book content in Lydhør. When a book is opened, there are five buttons on the bottom of the screen for “skipping.” There is also a branching icon that allows users to pick from a variety of options that go directly to different portions of the book. And the last bookmark icon allows users to return to previously made bookmark locations, that can even be individually named. All of these elements are visible in Figure 8.11 above.

In addition, the table of contents elements in books tested are also functional hyperlinks that allow the user to go directly to a section by activating (touching) the link. These links are presented similar to links on webpages, they are blue and underlined, and are shown in Figure 8.12 below.

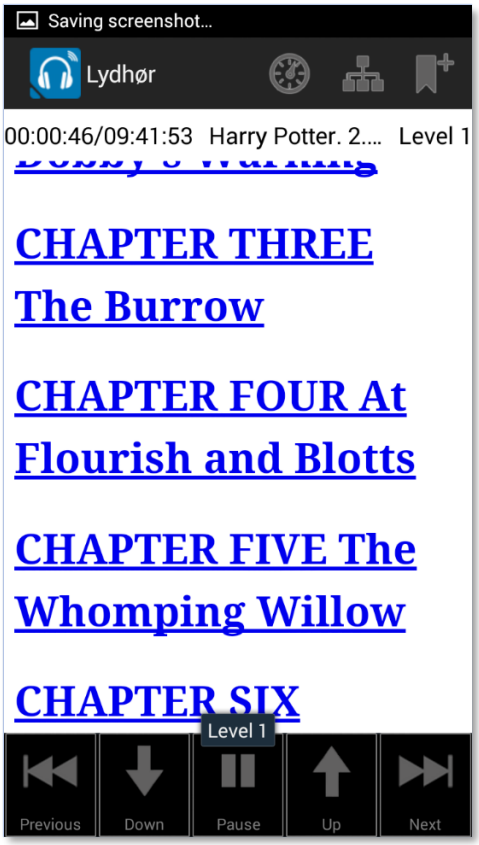


Figure 8.12: Navigation Hyperlinks

The skip buttons on the bottom of the screen are not particularly intuitive, and it took some time to learn that the up and down buttons adjust the skipping interval (down arrow for the rewind interval and up arrow for the fast forward interval).

As the interval buttons adjust the intervals, this is displayed in two locations on the screen. Both of which are not prominent and were barely noticeable. You can see in Figure 8.12 above that “level 1” is visible in a black box at the bottom of the screen and in the top right corner.

These bottom buttons can be modified from the settings to something that is called

“click-zone” navigation. This purpose and meaning of this setting was never understood during testing, and again, the user manual was unavailable.

The “speedometer” icon that is visible at the top of the screen in Figure 8.12 can be used to change the speed at which books are read. This section offers both plus and minus buttons, and a “slider” to adjust the speed at which books are read.

During screen reader testing, the fact that the buttons do not give feedback when activated was noticed. Talkback does “ding” to let the user know that they have successfully activated a button. However, although the speed was being modified, users that cannot see the screen have no idea to know which increment the speed is being changed at, and the speed at which the book is being read is now set to.

The branching icon that is visible at the top of the screen in Figure 8.12, is used to navigate to specific portions of books. Users can navigate to bookmarks (which would have been created before by the user, a specific page, or sections of books. The section option lists all chapters and makes it possible to navigate to a certain chapter or subchapter. This button functioned properly.

The only item of note was that the option that lets users navigate to a certain page, lists the individual pages, which could be tedious for users. For example, getting to page 500 would require scrolling through a list of 500 pages. In fairness, this list is searchable, which is indicated by a small magnifying glass icon at the start of the list.

The last icon on the top right of Figure 8.12, the bookmark, was investigated thoroughly because responses in the user ratings had mentioned the fact that bookmarks sometimes “disappear.” This could not be replicated during testing.

The only questionable design aspect, in relation to the two settings menus, was to question the need for two menus instead of one. The “settings” menu accessed via the book play screen does not contain settings. This may not be the optimal location for the “in-book search function” and the sleep timer. During the cognitive walkthrough, the items in this list seemed out of place (which is a purely subjective assessment).

8.3 Informal interview results

Two WCAG guideline violations were discussed in the interview. The scaling problem that was mentioned in the application testing section of this paper was one of these (The screen and buttons cannot be zoomed in on).

One positive experience from the interview, was that it allowed a new context in which to view guideline failure. Instead of simply viewing the technical description in the guideline, the extra work created and problems that this failure created, were observed.

The interviewee could utilize the app visually, but needed text that was presented with a heavy zoom, with very few letters at a time occupying the entire screen, due to vision limitations. This user sometimes preferred zoom versus using a screen reader, because it is faster than activating the screenreader, and because the screenreader voice and book

reading voice often overlap, as discussed previously.

The interviewee had to utilize a loop for magnification, for items that were not scalable or were presented in low contrast. If the user did not perfectly "mark" the place on the screen that they had identified with the loop, by holding their finger there, they would lose their place on the screen, and the process had to be repeated. The location of the text on the screen was important because it was needed to know which button to press. Because the buttons were placed closely together on the screen, the problem was further exacerbated.

The failing contrast ratio of the text on the buttons in Figure 8.3 was also discussed.

Several usability/user experience items were also discussed during the interview. One example is the problem that sometimes occurs, where the phone must be connected to the internet to access titles that have already been downloaded.

The fact that books were dated on the day they were "transcribed" by the NLB instead of the book's true publishing date was noted as a problem. And the fact that the date on newspapers and magazines is not provided in the search results was also discussed.

Some other issues were also discussed, such as the fact that in one instance, three identical searches revealed no results until the third try in one instance. Or the fact that the background changed colors spontaneously after searching for a title.

This interviewee noted the placement of the sleep timer and demonstrated that placing the sleep timer on the screen where it was typically needed (the play screen), would help. Displaying the timer on screen would reduce the substantial time and effort the simple act of finding and activating the timer currently takes.

The search process was also discussed extensively.

9 Survey question design

The discussion of results is temporarily halted to discuss the design of the survey questions. This section is a somewhat a mix of research methods discussion, but also a practical application of the results that have been discussed already.

The survey was designed to gather information to answer the research questions (identify issues and possible solutions), and to help make choices about the design of the prototype and the user testing scenarios. Many of the questions were intentionally designed to allow users to offer specific insight about their experiences with the Lydhør application, and provide their personal preferences regarding mobile applications. They will be discussed here. The original survey questions are available in the appendices, but they are written in Norwegian.

The first question asks users if they use Lydhør, in order to filter out non-users. Respondents that answered no would not be included in the final results, but the information that they provided was to be scrutinized, in case they offered features that they preferred from other applications.

Next users were asked if they use apps that are similar to Lydhør, and if yes, to state the name. This question was included to learn of other similar applications that were being used. The features of some of these applications were examined later.

There were various reasons for inquiring whether users had preferences/features they preferred in Lydhør or similar applications. First, it served to set a positive tone in general, for the survey. Because later questions would need to inquire into problems encountered, the aim was to keep the “feel” of the survey as positive as possible. This information might possibly also offer potential “upgrades/improvements” to add to the prototype, so that they could be assessed during user testing.

The next question simply asks which device was used to access Lydhør (ie: iPad, smartphone, etc.). The menus and way that the application is presented varies based on the device used. Different screenreaders are used. The question was important to identify whether there were any device specific problems being encountered.

The next two questions were intended to identify potential differences or preferences between the two screenreaders, and general comparison of them. The first question was flawed (which rendered data gathered there useless).

The second set of the screenreader questions, asked users if they used a screenreader when not listening to books. If they answered yes, there was an open-ended follow-up question asking to “essentially” describe the experience. The intention of the questions was that these answers would allow a comparison of visual use of the application versus non-visual use.

It is important to note that use of a screenreader would not necessarily imply that users had a vision deficiency. For example, some dyslexic users might prefer using a

screenreader for navigation, in addition to when they listened to books. However, not using a screenreader would imply that users, although they might have a vision deficiency, were navigating visually. This would allow comparison of many of the survey responses in a visual versus non-visual frame. In addition, answers to the follow-up question could give insight about how the “navigation” experience was when a screenreader was not being used.

The next two questions addressed usability issues that had been identified during earlier phases (review of the ratings, informal interview, and functional testing). The first of which asked if problems were experienced when searching for books, and the follow-up asked for details.

The next question set out to identify what information was important for users searching for books, and asked that, in an open-ended format.

The next question asked users if they experienced problems navigating within books, and if so, what they were. “Navigation” within Lydhør, consists of two distinct activities, navigation within the app and navigation within books. “Application navigation” had already been inquired about. As previously explained, within books there are different ways to “move” through the content (skip buttons, TOCs, and a button that provides the entire structure of the book). This would allow both identification of problems and show which set of controls they involved.

The next question was a question listing several known problems (based on what had been learned up to that point from the combination of the results of the ratings review, interview and application testing). Some of the problems listed were outside of the scope of the research, but they had been identified during the ratings review. Since a survey was being performed anyways, the thought was that these results would be given to the NLB, and they could prioritize any future modifications to Lydhør. The other problems listed were related to usability and accessibility. This was designed to show which proportion of users were encountering the various problems.

And, as is typical with surveys, the last question was an open-ended question designed to elicit responders to provide more details about anything they wanted to, such as features that they liked or disliked about Lydhør.

The survey was the only tool used for recruiting participants for user testing. Participants were presented with information about user testing in the conclusion section of the survey. In order to maintain the anonymity of the survey participants, and to ensure that they first submitted the survey results, a link was presented after they submitted their survey answers which directed users to another site that gave more information about user testing and solicited their participation (and guaranteed that they would first submit their survey responses).

10 Results phase 2-survey results

At the conclusion of the survey, there were 129 respondents. Of those respondents, answers from only 109 respondents were used because twenty participants indicated that they did not use the Lydhør app. Answers from respondents that indicated that they did not use Lydhør were still investigated to learn if data contained in their responses was useful (such as if they mentioned a feature of an app that is similar to Lydhør that they preferred and why).

The survey answers generated a substantial amount of insight into how users felt about the Lydhør application.

10.1.1.1 Proportion IOS to Android

The question asking about devices used was intended to learn an approximate proportion of Lydhør users that used android versus IOS or apple devices, and if certain problems were device specific. The answers showed that many Lydhør users have multiple devices, and often these were of both operating systems. For example, a user might have a Samsung telephone and an iPad. There was a strong preference for apple products, but the question was constructed in a flawed manner, and many answers could not be counted.

The flaw in the questioning was that the word “tablet” was used in the example. Generally, when people refer to a tablet from apple, they call it an IPad, which made it possible to determine which OS they used. However, tablets that are not apple are generally just referred to as tablets. When the answer was simply “tablet,” that could not be counted, because they could have been iPads or android devices. When cross referenced, there was no strong indication that any certain problem occurred more on a particular device, but the accuracy was corrupted by the tablet answers.

10.1.1.2 Using screenreader for other purposes

The question that queried whether users used a screenreader for other purposes than listening to books and identified visual users found that, of the qualified respondents that answered, 107 did not use a screenreader and 33 did. This proportion is not particularly useful for the research. But other question results (like the problem grid question) were filtered by this question, then viewed separately to see if the screenreader use had introduced additional accessibility or usability problems.

The answer to this, in a broad sense, and after many comparisons of various answers filtered by screenreader use, was no, for the most part. Most problems occurring were relatively proportional, meaning that problems mentioned were similarly proportional between screenreader users and non-users. One exception to that is that screenreader users more often mentioned searching problems or the desire for having a better searching method. Oddly enough, other responses from this same group also mentioned more frequently (than non-screenreader users) that they liked the Lydhør search function. Screenreader users also had a slightly higher ratio of problems related to in-

book navigation.

10.1.1.3 Similar apps used

Regarding the question about similar apps used, the most commonly occurring answer was Storytel. This was not extremely surprising because the application is heavily advertised in this country.

One surprising result from this question was that school related applications were mentioned frequently. Lingspeak, Lingdys, and Brettbook are applications that are provided on a national level to students with "print disabilities." These applications, (when combined) were mentioned nearly as often as Storytel.

10.1.1.4 Preferred function/properties

The question asking preferred functions or properties did generate a substantial amount of information to use for the prototype design. These answers required sorting, between those that liked some aspect of Lydhør versus some aspect of another similar application.

Sometimes the answers were in direct contrast, for example, one user preferred the five button navigation used with the Lydhør play screen, and thought the Storytel method of three buttons was terrible. Another user preferred Storytel's three button spooling and felt that the five-button method didn't work properly. This second answer also seemed to indicate that this user did not understand that the skip increments could be changed.

Because the responses contained so much qualitative information, most of it will not be detailed here. However, some of the highlights will be presented. Although the information is not presented in entirety, this answer set was referred to often during prototype design.

Amongst those answers that indicated they preferred features from another application, there were some reoccurring themes, better stability being the predominant one. However, many users preferred certain aspects of finding new books to read in other applications. Generalized, they preferred browsing for new books in other applications because they had options to learn more about a book than the title and author, although the way the information was presented varied widely depending on their preferred application.

There was also a theme amongst these answers that bookmarks functioned better in other apps, and that books started precisely where they stopped listening to them. Technically this should happen with Lydhør, and during testing of Lydhør for this research, it worked as it should and never failed. However, judging by the answers in the survey, the Lydhør bookmark function can be both unstable and, for a few users difficult to understand.

For the users that had a preference with Lydhør, there was some diversity in the answers. For example, many mentioned that they liked that they could download books (for offline use), while others liked that they didn't have to download books.

There was one very interesting set of answers from this category. They were interesting when combined with the results from the question that asked if users used a screenreader. There were several responses indicating that they preferred the “in-book” navigation. One of those responses even gave a clue that they liked that it functioned liked the DAISY reader. Those that preferred the existing skip and level buttons, overwhelmingly used screen readers to navigate through the application.

So, it appeared from this set of answers, that those that were used to DAISY functionality, were used to how it operates and prefer this method. Other items of interest from those that liked Lydhør functions were the sleep timer and the possibility to change the reading speed.

10.1.1.5 Problems when searching for books

Because this had been mentioned often in the user ratings for the app (and during the interview), participants were directly asked if they encountered problems or misunderstandings when searching for books. Those that answered yes were asked to elaborate. Roughly 40% of responders indicated that they did indeed have problems or misunderstandings.

The majority of those stating problems, mentioned something about the need to have search phrases spelled perfectly to find what they were looking for. Many respondents wanted more browsing capabilities, i.e. they might not know exactly which book they were looking for, and wanted to be able to peruse the collection of available books. Others liked how other apps suggested books based on books they had already accessed.

Many Lydhør users indicated frustrations revolving around books in series. Testing the app had shown that the NLB did usually add series numbers to the titles of books. However, many series have different individual titles. For example, the popular series that starts with Divergent (1), is followed by Insurgent (2), and Allegiant (3). So, if a user has read Divergent, it is likely that they must search elsewhere on the internet to find out which title to access next. The numbers are that the NLB uses are not worthless, however, because if someone accessed Allegiant, the 3 would hint that it is not the first book in a series, if that person was familiar with the way the NLB notates serial books.

10.1.1.6 Search result improvement

Previous research activities had found unhappiness with search results. Users wanted more information than the title and author, that had been made clear. But what exactly they wanted was not clear, so this question was asked directly in the survey.

The answers here were diverse, and obviously, users had various individual preferences. Some of the frequently occurring answers were genre and the date published. Again here, users wanted an improved series notation. The age group the book was written for and a short description of the book was also mentioned often.

Some respondents in the user survey indicated that they never use the Lydhør app to

search for books, only to listen to them. They instead use the NLB website, which has some capabilities to browse the collection without searching for exact titles.

10.1.1.7 In-book navigation

The user ratings review had shown that there was some displeasure with options to navigate within a book. Even though there are many available options, the ratings review had shown that multiple users had problems getting to the places they wanted to in books. The survey questions(s) tried to learn why.

Of the group that mentioned problems, the screenreader group was a little higher proportionally. Many of the navigation complaints, were due to bookmark failures, such as “it’s hard to find where I was when the bookmark glitches.”

There were a few that concerned navigating to the middle of chapters, and to subchapters. And finally, some respondents stated simply that the navigation controls were hard to understand.

10.1.1.8 The grid

The “grid” question that asked that participants check off problems that have encountered, was intended to be used to see the “weights” of the problems occurring. Of the “problems” related to this research, searching was by far the most frequently mentioned problem that users mentioned. The second most frequent answer was bookmark related issues. This was followed by the fact that there was no current page display, or no way to know how much was left of a book. The next two had an equal number of results, and they were the problems finding the sleep button and missing feedback/feedback that caused uncertainty. Only two respondents indicated displeasure about not being able to scale or “zoom in” the screen.

10.1.1.9 Catchall question

The last question, which was to be a “catchall” and gave respondents an open slate to talk about whatever they felt was not covered in the survey, did not yield the anticipated results. This question was meant to uncover any items not addressed in the survey, particularly problems with accessibility. No new accessibility problems were mentioned, and many answers concerned stability issues.

However, the sleep timer placement showed up a few times. The most common answer here had something to do with search related problems/wishes. The problematic comments focused mainly on the “precision” needed to find books. The wishes in general echoed answers from other questions. Like wanting more information about a book than the title and author.

The fact that books that were “loaned” were not available to use was mentioned a couple of times. This could be related to a misunderstanding the book reservation process, or it could be a glitch.

Many users wanted a tutorial or a way to learn how to use the various functions of Lydhør. One for example knew that bookmarks existed but had no idea how to use them.

One respondent wanted more school textbooks to be available in Lydhør. And finally, many respondents wanted a spool (like a fast forward or rewind) possibility instead of skipping by increments in books.

11 Prototype and user testing tasks design

11.1 Prototype design

At this point in the research, many possible solutions or improvements were being conceived to address the various problem points already identified (RQ1). To test these "solutions", a tool, that had changed the problematic aspects of Lydhør was needed, to see it they were indeed viable, or still flawed. A prototype application, although not elaborate, could be utilized for this purpose. To be an effective testing mechanism, this would need to be provided to testers on their own devices, to observe their use in a natural setting, especially for those that used assistive technologies.

The majority of the user testing tasks focused on the two most frequently mentioned problem "categories" that were mentioned in the user survey. These were various problems related to searching and navigation, navigation being a multifaceted problem. They are listed below.

Searching problems

- Search terms
    - precision required
    - suboptimal input method
    - search feature was not easily found and difficult to see
- Search results
    - Applicability/precision in relation to the search terms
    - Results display

Navigation problems

- Finding search button
- Finding particular sections inside of books
- Finding various other functions

Other problems noted

- General understandability --- button purpose ---- download vs. stream
- Confusion/lack of feedback

Potential solutions were to be created based mainly on knowledge learned from previous research and from knowledge acquired in the earlier research method phases, and from acquired knowledge of universal design principles (meaning the time spent studying the various courses in this degree focus). Many of these proposed solutions, had to be created solely based on subjective assessments, and would be tested to learn if they indeed solved the problem.

All of the questions have a common thread, and should all end with the phrase "for the print disability group." To avoid redundancy, that phrase is only asserted here but belongs at the end of all the following questions.

Will a simple bar (navbar) with the same three buttons, (which are core functions), that

is always present in the application, improve various issues/be the solution to the questions that follow?

Is a navigation bar simplistic (meaning an organized location place to place some of the most common functions/assistance) and familiar to most users?

Would the use of several “screens” (ie: the screen progresses to a new screen or displays noticeable new information) resolve some of the concerns/confusion about lack of feedback?

Where would the majority of users expect the sleep function to be located?

Because many in the group have difficulties with text, could visual indicators (icons) replace text descriptions when possible (that have been properly labeled for screen readers)?

Would the majority of visual users accept/understand that icons/buttons are actionable?

Would a help menu be a useful manner to assist users when they were stuck, and has the help menu question mark become an accepted, understood icon for users?

Testing would utilize both the real Lydhør application, and the new prototype. There were originally two purposes for doing this. The most important of which was that users could discuss their preferences between the two versions and would have both apps fresh in their mind. The other reason, was to time and compare efficiencies on the same task in both versions, but this was ultimately dropped.

11.2 Prototype design, technical considerations

The first parameter to consider was the range of devices that the user testing would occur on. In line with universal design principles, the aim was to design an environment, that, offered the same experience on all of the devices that Lydhør users typically used. One studies’ findings suggested this practice, and in their experience claimed to have obtained higher quality qualitative data as a result (Craven & Booth, 2006).

Because the Lydhør application is only available via two sources, the Google Play Store[29] for android devices and the App Store[30] for apple devices, personal computer presentation/layout did not need to be considered.

This would still require that the user testing occur on an elaborate array of devices, with different operating systems and screen sizes. There are software development tools that exist that remove the hassle of coding for both operating systems. However, the time frame that had been allotted for developing a tool for user testing was insufficient to both learn a new programming language, and implement the code.

However, a fully functional prototype was not the intention of the research. The use of representative prototypes is commonplace when testing design of IT artifacts. Many

[29] https://play.google.com/store?hl=en
[30] https://itunes.apple.com/app/apple-store/id375380948?mt=8

examples that have delivered fruitful testing results are offered by Buxton in his discussion of the "Wizard of Oz" testing technique (Buxton, 2007). The techniques described there range from use of paper prototypes to test menu decisions to having a human "type out" the computer's response as testing occurs.

Recommendations for improvement opportunities could be developed, and the intended testing could be performed with a far less sophisticated tool, even with paper based prototypes. But paper based prototypes would not capture the various techniques in which users typically went about using their devices. For example, navigating to a particular button with a paper prototype can be a quite different endeavor for a user that relies on a screen reader, versus a dyslexic user, and could even necessitate braille printing to accomplish the testing.

However, preparing a relatively hearty prototype that achieved *most* (not all) of the desired testing functions could be designed to be accessed by a browser. The three core technologies involved in creating a browser based prototype (HTML 5, Cascading Style Sheets, and some mild JavaScript) would need to be learned rapidly. Using a browser type solution would allow users to do testing on the users' own devices.

Some care was given to the fact that browsers interpret and visibly display elements differently, even on identical screen sizes. Additionally, screen readers, when used, announce elements differently.

For example, during testing of the prototype, although the prototype was not to be used on computers, NVDA screen reader was used for quick testing of compliance of the elements (to learn if the elements were announced properly). This screen reader did not recognize a button, which would need to be announced to screen readers during user testing. However, TalkBack and VoiceOver did properly announce the element.

And finally, Lydhør is accessed on several different screen sizes, so consideration as to how the elements were visibly presented to visual users had to be accounted for, for each possible device (group).

Many developers in the business sector are using various "frameworks" to address the visual variations mentioned before. There is of course a strong desire by clients that their websites are consistently presented on all devices they are accessed on. Note that consistent, in this case does not mean identical, it means optimized for each particular screen size. One of the frameworks making this possible is called Bootstrap[31], which is likely also the most popular framework.

Bootstrap provides higher odds of visual consistency no matter which device or which browser is utilized. In addition, the design allows for rapid development, particularly when a website is intended to appear differently on different size screens.

Another bonus is that Bootstrap is designed with accessibility in mind. They provide

[31] http://getbootstrap.com/

extensive ARIA and role support, and simplify the process of implementation, and provide extensive documentation[32] to help understand their use (which is simplified in Bootstrap). ARIA and roles, very simplistically stated, is a way to ensure that elements such as buttons are properly announced to screenreaders (this is not an exhaustive description).

Although inclusion of Bootstrap would create the need to rapidly learn one more technology set to create the prototype, it made it faster to develop a responsive prototype.

11.3 Testing considerations

Aside from establishing which technology to create and distribute the prototype, the contents of the prototype also required substantial consideration. How does one measure whether the search process is improved? How does one measure improvements in navigation?

For measuring navigation improvements, comparing the times it took to complete navigation tasks on each version of the application at first seemed appropriate. The prototype was designed with an intention of speeding and simplifying the navigation process, and later user testing could confirm whether or not this had been achieved.

During the pre-testing (of the user tests), navigation tasks were timed and documented (search performance was also). The information was compared. This approach was problematic, however. Measuring task time was an intensive activity, and meant that less time was available to observe the test subjects.

This also seemed to make the participants somewhat uncomfortable, as they tried to complete tasks as fast as possible. Because the testing sessions were only a limited amount of time, it also reduced the time to gather equally (or later judged more) important data, such as why a decision to perform certain navigation actions occurred.

For example, substantial time was used to create a help menu, which was intended to be consulted by the participants, to learn if it would improve the navigation process. This help menu was created because of survey answers and some subjective ideas. The pre-testing participants, feeling the need to complete the tasks as fast as possible, never wanted to use the menu, because it took time to consult.

In addition, when comparing the initial times, the data seemed to be of little useful value. Sure, participants in pre-testing could and did achieve identical tasks faster in the prototype on average than in the real Lydhør application.

However, the value in presenting improved times in the recommendations to the NLB, seemed much less useful than presenting various opinions, such as why users tended to look for certain functions in a particular location. The pre-testing participants had offered insights to these decisions, and the insights were perceived as being much more

[32] https://v4-alpha.getbootstrap.com/getting-started/accessibility/

valuable than time measurements.

Because of this, timing was dropped completely from the testing. This decision was discussed with an experienced researcher, who also agreed that timing the tasks was of little value.

After pre-testing the user-testing, the design of the user testing ultimately involved testing some navigation tasks (focused mainly of logical placement of functions), and searching for various book titles or authors.

The design and testing of the prototype itself was a critical milestone of the research activities. Its intention was to attempt to address all issues learned leading up to its implementation. It should also include any practical items from the review of the literature, if they were utilized. And finally, it should display cumulative knowledge acquired during the process of earning a degree in Universal Design.

In this case, it is also the final quality control measure that would test and ensure that the above items were interpreted and implemented in a manner which considered the research activities that had already been completed. Since some of the prior information gained was input gathered from users, the prototype should fit the average preferences of the print disability group.

Accessibility of the prototype was “measured” in regard to compliance with WCAG 2.0 guidelines. These were “testable” because the prototype was browser based, and there are no known compliance failures in the prototype. While this set of guidelines offers a nice starting point, they are still quite insufficient, so other principles learned during the study process were implemented and some previous research was used, for example the recommendations about “icon sizing.” Usability, due to the subjectivity involved, would be judged based on response from user testing participants.

Because the application was a “brand new” application, the accessibility problems that are present in the real Lydhør application were addressed. Before being introduced to test subjects, a guideline review similar to the one performed with the Lydhør was performed on the prototype. This was supplemented with manual screenreader testing during the designing phase and after the first run prototype was completed.

Having addressed the accessibility component of universal design of the prototype, focus was placed on usability issues within the print disability group. To eliminate much of the subjective nature of this portion of the design, several applicable prior research findings were applied.

The most applicable and relevant set of guidelines that was found for these circumstances, are Nielsen’s “Heuristic evaluation of user interfaces” (Nielsen & Molich, 1990). However, later the decision was made to switch to a more recent version of this list, which offers two enhanced versions, that have been refined.

Nielsen had found a solution to identify the majority of usability problems, and had

empirical evidence to support these claims. But he had noted that this involved a substantial amount of work to find and address many minor problems. In the same article another list, the "top heuristics to explain *serious* usability problems," seemed to be a more efficient approach considering the scope of this research (Nielsen, 1994).

The heuristics were adhered to as closely as possible, as often as possible, and the list follows. The entire list is not purely applicable, because it was designed with desktop computers in mind, before the advent of mobile smartphones. However most of the list is still applicable to mobile use.

Top heuristics to explain the serious usability problems

- Seeing/pointing vs. remembering/typing
- Consistency: same thing looks the same
- Feedback timely and accurate
- Salient repertoire of available actions
- Forgiveness: reversible computer actions
- Familiar user's conceptual model
- Feedback: show receipt of user's input
- Prevent errors from occurring
- Easy to discriminate action alternatives
- Modeless interaction

Note that the project timeframe did not allow for a full "standard" heuristic evaluation of the prototype. Instead, the concepts from Nielsen's list, combined with the prior research discussed below, created two standards to follow for the prototype design, simplicity and familiarity, which are discussed below. As decisions for the prototype design were made, the decisions were checked to see if they were compliant with the list above, but again, not in an exhaustive manner.

Some of the principles found by Mi, Cavuoto et al. were used. Their study resulted in a "checklist" of accessibility features for smartphones. A need for feedback when "touching" or selecting an item was one of these. Robust feedback was given for any action that occurred in the prototype, whether it was a "physical" change of screens, or for example, more information presented. They also noted a need for time to "self-train (Mi, Cavuoto, Benson, Smith-Jackson, & Nussbaum, 2014)." This last point could not be practically applied in the prototype testing but would be one of the final recommendations, as some participants in the user survey had also suggested that some form of "training" for Lydhør use would be helpful.In lieu of training, the help menu was available.

The same authors also noted a need for more "intuitive" navigational cues. The prototype would be used to test if navigation was indeed intuitive. They also mentioned a need to know the current status (Mi et al., 2014), which had already been implemented well in Lydhør, particularly for screen reader users. The intention was to maintain the current standard in that regard.

One previous study stimulated substantial curiosity regarding the prototype design. The research questioned if visual context benefitted dyslexic users, however only in the context of searching. They had found that inclusion of visual cues could be beneficial to both dyslexic and non-disabled users (Berget, Mulvey, & Sandnes, 2016).

An idea stemmed from that concept where visual context would be incorporated as often as possible in the prototype design. This would allow testing an "extension" of the previous idea. So, the prototype consisted of variations of similar "pages" where text and visual cues were included in different concentrations. The testers would be observed and sometimes questioned about their preferences to see if any knowledge could be learned as to whether visual or textual presentations (i.e. buttons with only icons versus icons and text for example) were preferred.

Two of the authors mentioned above conducted another study where they tested comprehension of icons. This added another layer of thought to the design, specifically, careful consideration of the visual indicators used. They had found that icon recognition is not universal, and recommended user testing of the icons to test comprehension (Gerd Berget & Frode Eika Sandnes, 2015b). This fit well in the context of this research, because it was questionable whether some of the icons used in the Lydhør app conveyed the proper meaning, as previously discussed.

In addition, findings from another study regarding the role of icons in toolbars offered some suggestions for the icons to be used and tested in the prototype. This research had found that reducing the complexity of icons could increase their understandability (Passini, Strazzari, & Borghi, 2008).

And finally, one study conducted research with users that had cognitive impairments (due to brain injury). They had found that these users had significant problems when dealing with extensive (text-based) menus on mobile phones (Nandigam, Symonds, Kayes, & McPherson, 2010). To address this, without removing functionality, the prototype design was to be broken into logical portions, that could help to reduce the load. This had to be balanced however, because too many "portions" could lead to confusion.

Using ideas from the research above, when somewhat "summed together," and combined with other universal design principles and previous user inputs, two broad themes were created to use during the prototype design process, which are simplicity and familiarity.

11.3.1 Simplicity

During the literature review, the concept simplicity was often mentioned in the prior research recommendations. One often recurring theme of complaints in the user survey hinted at complexity, particularly with navigation and searching.

Every function used in the Lydhør application should be preserved, but presented in a manner that was as simple and logical as possible. One study involving users with

cognitive impairments attempted to gather requirements for making mobile phones more usable for this group, found that simplicity was *key* for achieving this (Dawe, 2007). Other previous research had also used the same strategy, with positive results. They referred to the strategy as “hiding of complexity” (Stevens, Edwards, & Harling, 1997).

This simple modification would serve to benefit users with cognitive impairments. It also stood to simplify the process for users that had to rely on screen readers. The gain from simplicity here is an easily understood assertion, because not following this principle forces users to, for example, listen to menu items that they do not need to access all the time, but must hear in order to access items they need; and it places more decisions on users with cognitive impairments.

Another subjective design assumption was made, and that was to place the most likely used and most important functions first in the menu “lists.” This would again save time for users that had to rely on screenreaders, and would be easier to find for all users in general, because the first list item should be the first item that visual users “snap to” (this would be tested in the user testing).

After a user is logged in the current version of Lydhør, users encounter the single main screen of Lydhør every time they use the app. This screen is the lone gateway to every function of Lydhør, and it contains a large portion of information that must be considered with each use.

The prototype would attempt to “break up” some of this large amount of information from the Lydhør “main screen.” The next design decision was simply to question, “How many core purposes/function or verbs does Lydhør have?” The reasoning was to divide the application into logical “rooms,” envisioning a separate location for each separate main activity. Changing “rooms” or screens for different functions was also an effective way to offer clear feedback for inputs.

Lydhør has one main function, listening to audio books, with other sub-functions that support the main function. The supporting functions are: finding a book, and browsing the collection of books that one has “checked out.” There are also several rarely used minor activities, which are generally settings modifications.

The goal was to divide each main function and sub-function into its own “screen” because many of the users would be accessing via mobile devices with small screens that can display very little information before becoming too crowded for visual users, and tedious for screenreader users.

There is well-established HCI research known as the depth vs. breadth tradeoff. This research established that menus were more usable and understandable when they had a “broad-shallow” choices versus “narrow-deep” structures (Kiger, 1984). Per lazar, there is a multitude of empirical research that has now proven the “broad-shallow” approach to be superior (Lazar, 2010).

This idea had to be somewhat extended because at the time it considered on screen (computers) multilevel menus (that were mainly text based). It was also being applied to the *overall structure* of the application in this case, *not just the menu structure*.

Smartphone operating systems have now firmly established the use of buttons or icons to replace many of the text-only lists. This practice would be adhered to in the prototype design, when possible.

Using the "broad-shallow" structure allowed the same meaning to be conveyed in a smaller space visually, versus text-only means, which again, helps address the small size available on mobile screens.

Based on that line of thought the prototype was given an entry portal, that would allow access to all important functions and sub functions. This added a screen that was not present in the real Lydhør application. Even though it had added a screen, users were still usually only one step away from accessing any point in the application, and this screen offered very few decisions. The screen was simplistic, it contained two or three buttons, which was intentional to try to learn if two or three buttons were needed there.

The two "main screens" are shown in Figure 11.1 below. The three buttons, when translated, mean: listen to a book, my books, and find a book. The "listen to a book" and "my books" buttons both led to the same place, the screen showing the books users had checked out.

Listening to books is the main function of the application, and it was questionable whether users would be confused if they were not presented with that option. On the other hand, it might be logical for users to want to access their books first, to decide which book to listen to. This would be tested in every task, because these screens were varied with each task.

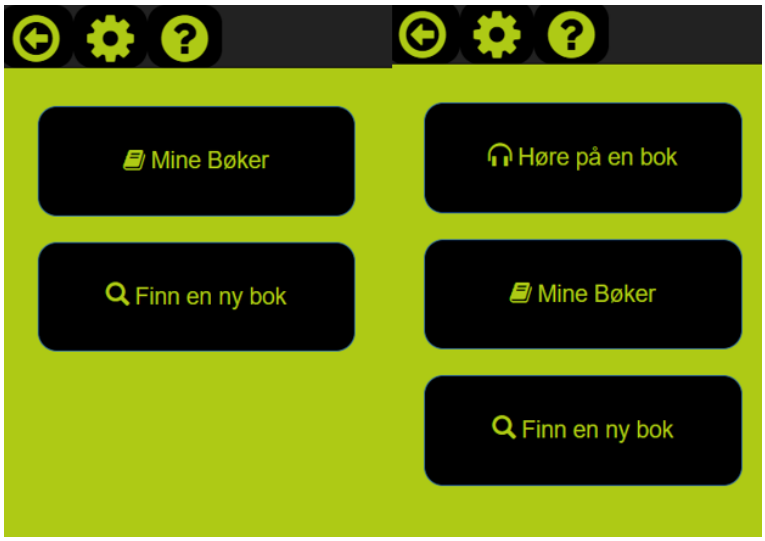


Figure 11.1: The two versions of the Prototype "Main Portal"

11.3.2 Familiarity

Some prior research conducted regarding icon recognition inspired some design

decisions that were to be tested in the prototype application, based around the concept of familiarity. This affected the choice of using icons, and in addition, the same principle was extended to the overall design, because navbars (which are discussed below), should now also be a familiar concept.

This prior research by Horton was an in-depth investigation of various aspects of icon recognition (Horton, 1994). Of particular interest in this research was the "model" of icon recognition that was used, that consisted of three steps. The first of which was what was termed "decoding," where users' prior experiences determined whether a user would "interpret" the meaning of an icon. Of particular interest was the idea that "past" experiences dictated the likelihood of successful interpretation.

This learning "effect" is echoed in the second "recall" phase, where the author asserts that icons that are "familiar" because of past interactions, are generally easily recalled. The last step asserts that learning whether or not there has been a successful "encoding" requires situational testing by human users.

Later research that used Horton's Framework, identified another item of interest. Guo found that younger users had much higher rates of icon recognition (Guo, 2016).

When combined, the Guo and Horton research (although they did not explicitly state it in this manner) offered an interesting question, that is applicable to this research. Did the younger users have higher recognition levels because of "past experience" with the icons? The implications, if correct, could open a new question based on those premises.

This question, developed during the context of this research (late in the process), was: Could user navigation be simplified and more understandable if familiar mobile application structures and elements were applied? By familiar the intended meaning is; a structure used often and often encountered by users on their mobile devices in other applications. And if so, would encountering familiar items make an application more usable?

One example of this that was used in the prototype is the settings gear icon. Some form of a gear cog is used by both android and IOS devices (the mobile phone main settings), and there is a high likelihood that visual users have encountered this icon frequently, because many important functions for these devices are stored there. This offered an opportunity to get a sort of "double" recognition value because of its typical meaning in applications, and, importantly, the type of content that these icons typically offer access to. For non-visual users, they should have had similar experiences from hearing "settings" announced.

This design assumption would be tested during user testing, so, in a subjective manner, all logical "main" or frequently used functions in the application that could be placed there were located in the navbar placed at the top.

Along this same line of thought, a question mark icon was used to house the help menu.

This icon is also now nearly universal in English and Germanic languages as a place to find answers. Help menus are typical in many mobile and PC applications, and users in the user survey also indicated that Lydhør would be easier to use if one were available.

Another example is navbars, which is increasingly being seen in many mobile applications. This is typically simply a bar that spans horizontally or vertically that two or more icons are typically installed in. Two examples of navbars are shown below in Figure 11.2. The gear and question mark icon are also shown there.

Figure 11.2:Prototype play screen

Another question that had to do with the concept of familiarity was the use of buttons, and icons, as actionable items. This would test whether sighted users in the print disability category were comfortable with button use, and or using icons as buttons. In addition, icons were used on all buttons, either as the only information or in tandem with text. This could potentially aid dyslexic users uncomfortable with reading text. Any experienced mobile phone user should have experience with using icons to “navigate” through their phone, so using them in application should be widely accepted and familiar, because all Lydhør users owned smartphones or tablets or both.

For the users with vision deficiencies, the previously mentioned mobile minimum size requirements (9 mm with 4 mm padding) were adhered to[33]. This could not be completely controlled because of the variance of the devices that would inevitably be used to access the prototype, but the aim was to meet this requirement at a minimum or larger.

11.4 User testing task design

The prototype was created to facilitate testing of mainly perceived improvements

[33] https://www.w3.org/TR/mobile-accessibility-mapping/

identified in the earlier phases of research. The following discussion continues to elaborate about aspects of the prototype design related to the tasks that users would do during testing.

The prototype was customized specifically for each testing task. To accomplish this, a portal was created, solely to be used during testing, that had a button for each individual task (with exception to searching tasks). A portion of this screen can be seen below in Figure 11.3, or at the website where it is hosted[34]. Details about the tasks are now provided.

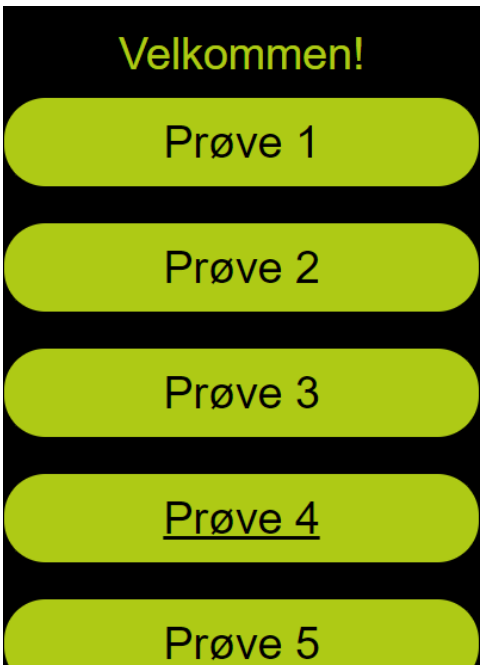


Figure 11.3: Testing portal

11.4.1 Understanding download task

Some of the survey respondents voiced displeasure about not being able to access titles that were downloaded to the phone. Some of these were undoubtedly related to a glitch in Lydhør, that is outside of the scope of this research. However, some of the answers indicated a misunderstanding as to whether a book was downloaded versus just checked out and could be streamed. This concept was also misunderstood during the cognitive walkthrough. The complexity is compounded by the fact that some books can only be reserved.

The prototype was a possible solution to increase understanding of the status of a book versus the way Lydhør approaches the same problem. This is why participants for user testing were told simply to make sure that they came to testing with at least book "in" Lydhør. This direction was specifically vague so as not to give away the purpose of this test.

Users would be asked to find a book already in the application (again intentionally

[34] http://nlblydhor.000webhostapp.com/

vague) for Lydhør and the prototype. They were then asked; “If you are not connected to Wi-Fi, can you listen to this book?” Because a bit of the surprise would be lost after the first time this question was asked, testers rotated, and some did (all of the) tasks first with Lydhør, and the other did the tasks first with the prototype.

For the prototype, the screen showed a tile with the title and author, a button, and below that a statement, that translated means “This book is saved on your telephone” or “This book is not saved on your phone, but can be streamed now.” The second statement was originally only “Can be streamed now” but was changed after both pre-testers tried to click the statement to test if they could indeed listen to a book.

For books that were not saved on the telephone, there is a button below that says, “download this book.” Figure 11.4 below shows three screenshots of evolutions of the bookshelf” screen.

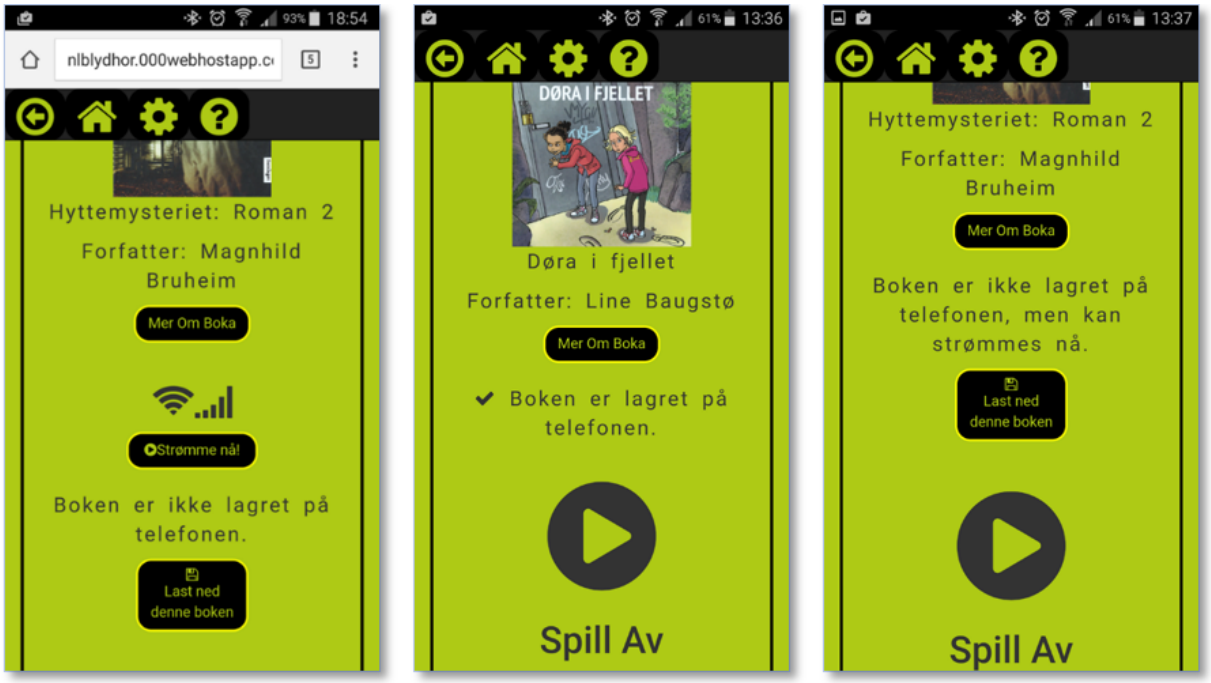


Figure 11.4: Evolutions of the prototype bookshelf

This screen was modeled after the actual search results from the NLB website, because it presented results in a manner that prior to testing, seemed to offer to reduce confusion. One particular feature there that was promising, was a button that allowed users to learn more information about a book, so it was replicated in the prototype.

At the conclusion of testing, testers were also asked their opinions as to whether this could be a possible way to design the search results screen (not just a bookshelf screen). One important note is that one book title now occupied more than one telephone screen whereas it was possible to see four or more on the real Lydhør app.

11.4.2 Sleep timer task

The sleep timer was apparently a popular feature with Lydhør users. Respondents typically mentioned that it was hard to find, and some mentioned that it would be nice

to have (because they were unaware that it was available in Lydhør).

A set of tests was devised to learn where the most logical “home” for the sleep timer would be. It also tested if the clock icon was appropriate or logical for the sleep function. The sleep timer was placed in various locations, and testers were asked to find the sleep timer and activate it in the regular Lydhør application, and in three different prototype screens (for three separate tasks) that were designed to learn if an optimal placement could be found.

This test was also part of another broader test involving various button placement strategies, where the idea was to learn the optimum placement and number of buttons on the book play screen.

Two of the screens had the sleep button included in toolbar on the bottom of the screen, and one on the side. The buttons were all icon based except for one set that had an icon and text on the buttons. The button was also available in the main settings menu, and information leading to how find the button was provided in the help menu, giving users multiple ways to locate it.

Figure 11.2

Figure 11.5: Play screen variations

The various placements of the sleep timer are shown above inFigure 11.5. It is the small clock icon on the buttons/icons on the first three screens. The timer was also available on every screen in the application via the settings gear (cog). The text on the screen was from a story on a Norwegian Folk tale website [35].

[35] http://folkeeventyr.no/tre_kongsdoetre_berget_blaa/

11.4.3 Bookmark task

Another task of finding and activating the bookmark followed. The intentions for testing are the same as with the sleep button (icon recognition and button placement). The third screen shown in Figure 11.5 was used for this test.

This variation of the screen design was responsive, and had a different appearance on a larger tablet screen, which had more space. The hope was that some testers would use a tablet device, and the layout shown below in Figure 11.6 could be discussed, but none did.

Figure 11.6:Sidebar play screen: tablet format

11.4.4 Changing reading speed

Being able to modify the speed at which the text is read aloud is an important feature for users that rely on screenreaders. Notably, in Lydhør the reading speed adjustment only affects the speed in which books are read, because the device's screenreader would be responsible for navigation or exploration of screen elements.

Lydhør uses a symbol that looks like a car speedometer, which many limited vision users may not recognize because they have never driven. For the prototype, no fitting icon could be conceived, and although important, it was thought that users would access this feature less than the other features. So, this function was only accessible via the settings menu (with no icon applied).

This time users were intentionally not directed to listen to a book, rather just to change the reading speed of the application. This was envisioned as something that users would modify before reading began, so the necessity of placing it on the play screen would be tested. Recognition of the icon that Lydhør uses would also be tested when testing the Lydhør version.

11.4.5 Skip to chapter 3

As mentioned before, Lydhør has a powerful function that allows users to get to certain places within a book, by chapter, or by searching for a word (or words), or by page number.

This is one of the three buttons atop the play screen, and it looks like a branching icon. It is shown below in Figure 11.7. One question that performing user testing intended to learn, was whether this icon conveyed the meaning of the button properly. The prototype used an icon very similar to the one Lydhør uses (but presented in higher contrast).

Figure 11.7: Branching icon

This task was also intended to learn the preferred skip method for users. So, the intention was to observe how testers skipped in Lydhør, because they might choose to utilize the five-button system on the bottom of the play screen, which was discussed in the play screen section.

Rather than using the five-button setup that Lydhør uses, the prototype instead offered options to change the interval when pressing the skip buttons. This was intentional to test if that might be a superior method, and a desire to free up space because of the limited space on mobile screens.

By observing users navigate with Lydhør, and the prototype, much could be learned about which method of in-book navigation was most popular with users. Testers were simply asked to open a book and skip to chapter three for this task.

11.4.6 Testing the search function

This portion of user testing was one area where the full testing "wishlist" was never fully realized. One recommendation that was implemented, was use of a search engine that offered search recommendations while the users input search terms, because of the frequent misspellings that are typical with dyslexic users (Berget & Sandnes, 2016).

When it became clear during development that coding a specialized search algorithm would not be achievable, the plan was to include Google's own search engine in the prototype, and have it search the NLB library website. The autosuggestions that this engine provides are accurate, usually even with spelling errors or letter transpositions. When testing this, the results were quite accurate despite many intentional spelling and letter inversion mistakes.

However, at that point, one test search returned something it had not returned before, advertisement links. In addition, the search engine was occasionally returning results in English, which was another aspect that Berget had found problematic for dyslexic users (Berget & Sandnes, 2016). Because removal advertisements would have involved

paying, and because of the results in English, plans to utilize the Google Search engine were abandoned.

You can see Figure 11.8 below showing how the Google Search Engine functioned with the NLB website. The three advertisement results that occur first show why this would not be practical for user testing. The results were also problematic because they were formatted in the typical white and blue Google results style.

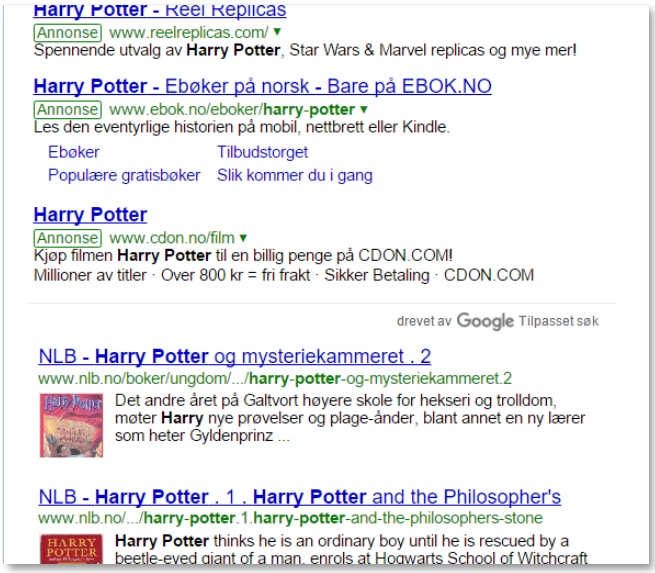


Figure 11.8:Google search with the NLB site

At the last minute, the decision was made to use the "backup plan." This involved using the actual NLB website as the search function in the prototype. When pre-tested it's search function was much less reliable than the Google search function, when identical misspellings and when identical letters were transposed (identical phrases used to test the Google engine). However, it did list recommendations as the search terms were entered, although they too, were far from perfect when mistakes were present.

The NLB webpage, like the prototype, had been responsively designed to varying layouts on different size screens. The display on small mobile screens was not perfect, mainly because of some of the elements that were now very prominent at the top of the screen.

The NLB offers the option to search the NLB website, or the database of books on its website. On the full-size screen, it is a small feature and contains the search box. But the website is responsive, and a small bar on the larger screen version becomes a large distracting element on small screens, which was difficult for the testers with limited vision to process. It also consumed more a large portion of the screen. In the Figure 11.9 below, you can see the large black feature atop the screen.

Figure 11.9: Prototype using the NLB search function

For the user testing, users did a few searches with both the Lydhør application, and the prototype. They searched for the same terms in both. Terms were chosen that had a high likelihood of being misspelled, or not being found without, for example having a colon in the title. This would allow observation of the various techniques employed by the users in this diverse group. And users were asked open ended questions after completing both, to see if any insight about preferences or difficulty could be learned.

The original prototype design used a simple, high contrast, black and white design as seen below in Figure 11.10. In a final check that occurred shortly before testing began, the change from the prototype pages, to the NLB website which was black, green, and white, seemed to be dramatic. It was possible that this dramatic color change could be disruptive for testers during user testing. To avoid this, the colors from the NLB website were used for the prototype, giving it the green and black design seen in the previous pictures.

Figure 11.10: The original prototype color scheme

12 (Final) results-phase 3 user testing

User testing took place with five participants that were qualified Lydhør users, meaning that they had a print disability. Three of which were females, one of which was > 60 years old and the other two were < 50. All of the females had limited vision capabilities, although the problems and severity of the problems were quite varied. One used VoiceOver most of the time while accessing the apps. The other used a loop that was installed in her glasses and occasionally used Voice Over, but preferred not to. And the last female participant only used Voice Over only if she absolutely had to. The other participants were males, the youngest was 12 years old and the other was in his mid-forties, both were dyslexic. All participants brought either iPhones or Samsung smartphones.

There were three additional participants (that were not print disabled), two of which helped to "pre-test" the user-testing. In order to assure that the user testing could deliver the best possible results, the user tests were pre-tested 3 times before the first official user tests began. There were two users that pre-tested (one tested twice). Both were females, one less than forty years old, and the other was more than seventy years old. The third individual had inquired about the prototype, out of curiosity (midway through the testing). This individual, a male aged 44 and a technology enthusiast, was easy to persuade into a practice testing round.

The users in pre-testing were not "print disabled," and that the intention of performing pre-testing was simply to ensure that user testing would go smoothly and identify problems in the testing process. However, some of the insights provided by them were worthy of including in this discussion of results.

For example, during the pre-testing rounds, the tasks were all individually timed. But as previously discussed, timing tasks was dropped in favor of gaining actionable, qualitative input from the users, and to encourage them to access the help menu (to learn if it was helpful).

Following the theme that the entire research project would be adapted and modified based on findings and needs, the user testing also evolved as it progressed. Adaptations were made before "official" testing, and in between rounds of "official" testing when they were felt necessary. The testing results are, for the most part, presented by each task performed during testing. The "speech" given to users before testing began, is in the appendices, and is named pre-testing speech. The transcript that was compiled for testing, which described the tasks is in the appendices, it is named testing plan. Both of these are written in Norwegian.

12.1 Understanding streaming versus downloading

Regarding the Lydhør app, two of the three visually limited users had very little problems understanding the small icons that the Lydhør application uses, and whether or not a book had been downloaded. One was a veteran Lydhør user that had long ago adjusted

the settings where Lydhør downloaded books automatically. The other one, which was using Voice Over showed me that when the title is read aloud it also announces "Downloaded."

The other visually impaired user had challenges, and didn't understand downloading versus streaming, or where in Lydhør she could go to find out, when we discussed it during testing. She had also never noticed reservation versus checkout. She said the fact that she sometimes checked out a book and have never received it, and just thought that it was a glitch, could have been in fact reserving books (the buttons are identical except for the text).

Both dyslexic testers had no problem with the current version of Lydhør, and were familiar with the downloading and streaming process, and the icons. They easily understood the prototype download status as well, and quickly identified the on screen elements that stated the status of a book.

The prototype screen was problematic for all the visually impaired testers because there were many elements on the screen that they had to slowly process, usually visually. One tester, activated Voice Over, and had to listen to several items before identifying the status of the book.

What was most enlightening was what happened when she tried to have Voice Over read aloud the picture of the book cover. She could tell that an object was there on screen, but couldn't see it enough to determine what it was, so she activated the screen reader.

However, this object had been set to be "skipped" by screenreaders. The reasoning was that it would cause screenreaders to read the author and title twice, once for the book cover, and once for the text. Since she had extremely limited vision, the book cover represented a frustrating object she could not learn more about.

The typical wisdom in the universal design community is to have screenreaders skip pictures that have no importance, and such would be the case with book covers. Alternate text on a book cover would likely normally be the title and author, causing the information to be heard twice. However, in this case, the "muted" book cover became a problematic object.

12.1.1 Finding the sleep timer

This set of tests showed the value of iterative user testing, as the testers pattern of finding this function were quite different than anticipated. The expectation was that users would find the sleep timer in various locations (in the prototype), and could be queried afterwards to learn why they went to a certain place.

Participants had been asked to start listening to a book (press play in the case of the prototype because this function was not activated in the prototype, which would take them to the book play screen).

The prototype results became one of the most interesting aspects of the entire user testing (when combined with the bookmark test, but this will be explained in the discussion). The testers unanimously favored finding the sleep button in the settings menu, which was present at the top of every page in the prototype, although one tester did access the help button to find it first.

Despite being asked to begin listening to a book, most participants went directly to settings on the first page of the application. It became clear that the testers would do this repetitively, and that testing the various positions of the on-screen button was ultimately unnecessary. So, after the second “official” test, when testers found the sleep timer via the settings menu (the pre-testers did this as well), the other two tests that were planned with the sleep timer were abandoned. This decision was confirmed as later testers typically followed the same actions.

The participants were asked oftentimes, why they chose to go directly to the settings menu to find the sleep timer, and why they would not expect it be onscreen somewhere when they were listening to books. Many indicated a “familiarity” with the settings cog icon, and stated that they expected the sleep timer to be in the settings menu, which is shown below in Figure 12.1.

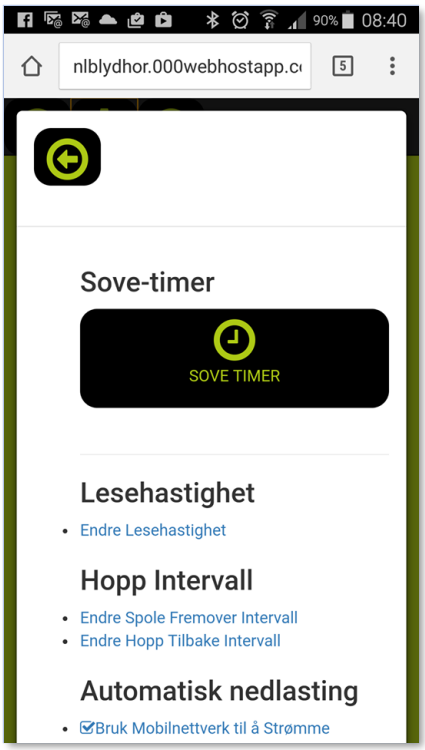


Figure 12.1: Prototype settings menu

The responses with the Lydhør app was varied. One participant knew the location because she used it frequently, but took the time to show me how much extra headache was involved with having to use the screen reader to read through the menu. Another remembered somewhat where it was, but again, used substantial time to listen to the

options. The other participants gave up on finding the sleep timer, or had substantial difficulty finding it. More than once they tried to access this via the control that looks like a speedometer in Lydhør.

12.1.2 Finding the bookmark

This task was identical to the sleep timer task, except for the target, and again asked users to begin listening to a book and activate a bookmark. In the Lydhør application, some users found the bookmark right away, but some of the users with impaired vision had to search around, and mentioned that the controls were difficult to see (these were the controls that had a low contrast ratio).

When using the prototype two users again went directly to settings menu, and wanted to activate the bookmark from there (the bookmark was not placed in the settings menu). When they didn't find the bookmark in the settings menu, one chose to access the help menu, which helped the user complete the task. The other user decided to try to start a book, and then saw the bookmark icon on the play screen.

However, three participants in this activity did indeed start a book right away, and activated the bookmark this time from the onscreen toolbar, instead of via the settings menu. Despite being higher contrast than Lydhør, users still had to "search" through the icons to find the bookmark, if they had a vision deficiency. All users recognized the icon, which is similar to the one used in the Lydhør app.

When these testers were questioned why they had chosen to access the button this time on the play screen (in contrast to accessing from the settings menu), one tester stated that the logical place for the bookmark was the location where you would be listening to a book.

When challenged that the same logic could be applied to the sleep timer, one participant noted that for them it was more logical to activate that function before starting a book. The others that were questioned about why they did not access via the settings menu this time, generally could not explain why; one had made the choice "on autopilot."

12.1.3 Changing the reading speed

Participants that had used the reading speed setting before in Lydhør generally remembered this components' location, although finding the precise control (out of the three low-contrast options) was tedious. Testers that were not familiar with the location of this function struggled in Lydhør. Some spent time searching the settings menu (it is not available in the settings menu), others just "tried" the various buttons, particularly the three icons at the top.

When doing the same action in the prototype, all of the users except one went to the settings cog, some accessed this from the entry screen, and others started a book and then went to settings. The other user started a book and tried the sleep timer button, then accessed the help menu. In the prototype, the only possible location to adjust this feature was the settings menu, no button was provided on screen. However, many of the

testers indicated that this is feature is frequently modified when listening to books.

12.1.4 Skip to chapter 3

The “skip to chapter 3” task, was another example that demonstrated why user input is often mentioned as a superior method in academic articles, because an “envisioned” solution can be faulty.

When doing this task in Lydhør, none of the users accessed the branching icon. Audiobooks do not display the book text on screen. They do however typically display a list of chapters, which are a familiar blue underlined “clickable link.” Both dyslexic users quickly capitalized on this option, before being asked to use another method. One of the limited vision participants also did this, while the others used the skip buttons on the screen, although they did have some problems setting the proper interval.

Two of the limited vision users, when changing the intervals, had to “experiment” somewhat. One, used a screenreader, and had to listen to several interval options. The other changed the interval and clicked skip, to test what the interval had been set at, and continued in that inefficient fashion until the proper interval was set. One of the visually impaired testers also wanted to access the links, but was asked to try use another method. She then also used the skip/interval buttons, which she was experienced with.

The small interval notification that is displayed on the screen in Lydhør, that was shown in Figure 8.12 , was never noticed. This feedback always went unnoticed.

With the prototype testing, some users were again tempted to just “scroll” down to chapter three, even though the prototype displayed the book text. When asked not to try that method, all the users except one tried the skip forward button, and set the interval to chapter. The lone exception went directly to the help menu, and read about how to skip chapters. The help menu guided users to use the branching icon, so that user did in fact use it.

After completing both tests, all participants were asked why they did not use the branching icon. The answers were overwhelmingly in agreement, that they were unaware of its purpose. No one had explored or encountered this button before. The branching icon served only to demonstrate what happens when an icon is not “familiar.” The icon was never used or explored, in either app, except for the one tester that gained information about it via the help menu in the prototype.

12.1.5 Search function testing: Lydhør

The search trials with Lydhør were problematic for testers. One took the opportunity to explain that he never used Lydhør nor the NLB website for searching for books. The internet contained more information, he explained, and it was generally easier to find out information about books that were part of a series there.

Despite being experienced with the Lydhør search function, many users still had

problems with the Lydhør search function. Finding, the magnifying glass for activating the search function (which was hard for some users to see because of the low contrast), was not easily accomplished. Twice during testing, users had activated a tab other than the Loans tab in previous tasks, so the search function was not available and finding the search became even more problematic.

Another yet undiscovered problem that surfaced, was the fact that search terms could only be submitted via the magnifying glass on the on-screen keyboard in Lydhør. Some of the limited/no vision testers had trouble figuring out how to submit searches despite having experience with the app. The magnifying glass was difficult for these testers to differentiate between it, and the letters on the keyboard, which is shown in Figure 12.2 below. One tester explained that each individual letter or element on the keyboard had to be visually scanned. One user also explained that she sometimes "scanned" the keyboard with Voice Over activated.

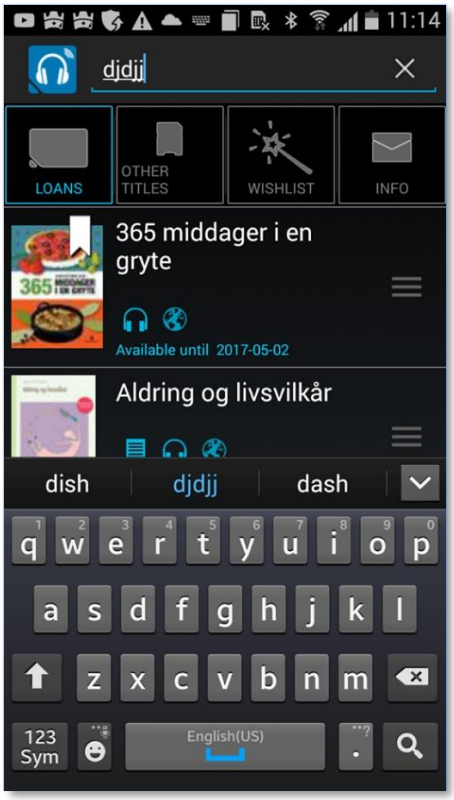


Figure 12.2:Lydhør search

Search results in the Lydhør application are presented with several results on screen, with only the author and title, and the book cover or a placeholder icon when the book cover picture is not available. You can see an example of this in Figure 8.8. When questioned during testing, many of the participants preferred this format of presenting results versus how they were presented in the prototype (they were asked if the prototype bookshelf could be a viable way to present search results).

12.1.6 Search function testing: prototype

The “anticipated improvements” that were used in the prototype, were much less helpful than envisioned. Testing of the search results autosuggestions beforehand had shown that they would not be read aloud *automatically* by screenreaders. This had opened a line of questioning as to whether it would be beneficial to be read autosuggestions aloud as one were entering search results.

However, if there were several suggestions, then a list of titles and authors would be read while the search terms were being entered, which could be quite distracting, especially for those who “listened” to the keyboard to find a particular letter.

What was not considered, but should have been, was whether or not the screenreader would read aloud autosuggestions, when they were “touched” or selected. When the testers with limited vision were entering a search, they oftentimes tried to get the screenreader to read aloud the suggestions (rather than entering the rest of the search), but this did not work with the screenreaders.

The navbar atop the prototype search screen (shown in Figure 11.9) proved to be both distracting and confusing. Users were “drawn” to this element, almost as though they expected this element to be the search tool, possibly due to the positioning at the top of the screen. It was especially problematic for limited vision users.

Two users after unsuccessful searches disliked the fact that the old search terms remained in the search field. However, another user, after an unsuccessful search, had searched the text that remained for errors, and found it useful that the text was still there.

On occasions, as testers were entering search terms, they noticed that the autosuggestions were disappearing. They were not being removed because they were irrelevant. For example, the word bacon might appear after the letters BA were entered, but disappear if BAC was entered.

Another unforeseen action occurred with text entry (this occurred during Lydhør testing also). Many times, users, particularly the two dyslexic users (but not exclusively), when entering a search term, would begin to type a word, and then use the autosuggested word that appears at the top of the phone keyboard (not the search box autosuggestions). Long search strings could be entered quickly via this trick, with much less typing. Both the dyslexic and visually impaired testers utilized this feature.

One of the testers with limited vision preferred to use the phone in landscape mode any time she needed to enter text from the keyboard. This is because the keyboard has larger buttons when the phone is turned sideways, which are easier to see. However, the keyboard occupies more than half of the screen, covering most of the page.

For this tester, the keyboard obstructed the search autosuggestions completely. This tester preferred to use search suggestions, because the process of entering text was

tedious, so the prototype version offered no real improvements for her. She did however use the keyboard autosuggestions sometimes to save time.

User testing showed that one feature, that was present on the NLB site (that was implemented into the prototype), that was envisioned as being a substantial improvement from Lydhør, did not live up to these expectations. The search box included a button that has the text “søk” (search), and a magnifying glass icon. This offered a non-textual method of identifying the search box. The thought was that the icon would make it easier to identify the search box itself as a search box.

However, it seemed to add more work for all the users with limited vision, because it took time to look closely and first identify the text, then the icon. In addition, the fact that it was a button was not immediately evident to them either. There were no problems for the dyslexic users identifying it. One problem that was not caught during pre-testing that was sometimes observed during testing, was that the search text input field, had very little space for text (because the search button occupies approximately a quarter of it).

As testers entered the longer searches, their search terms would disappear off the left side of the search box, and some users were disturbed by this. These users wanted to proofread their entry before submitting it, and two users (both with vision limitations) said after testing that they preferred to spend time to check that the search terms were right the first time.

One tester noted that it takes some time to “re-orient” themselves when search results were returned because she could only see one or two letters at a time. Reference Figure 12.3 below to see how the text would disappear from the search field.

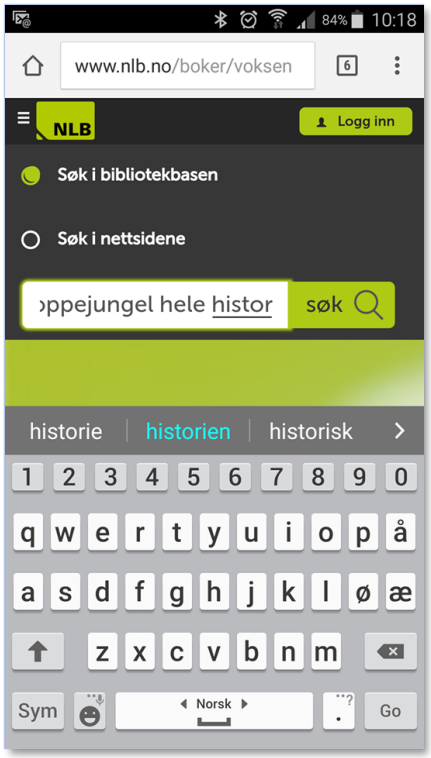


Figure 12.3:Small search field

12.1.7 Miscellaneous results

During the first “official” user test, one tester pointed out a visibility problem with the prototype. There were many modals, which are basically popups, which originally used a “windows style” X to close the modal. This X was barely visible, despite a high enough contrast, because of its’ small size. Initially, it was thought that users would just use their phone’s back button to get out of these menus, but iPhones do not have one.

This X was replaced on every modal page with a high contrast large arrow pointing to the left on the top of each page, as seen in Figure 12.4 below. This arrow had already been installed on all of the other pages of the prototype, due to comments from the survey, where users basically wanted an easy way to reverse incorrect button presses. This button helped the limited vision testers, if they were searching for something in the wrong place. All testers, once had they realized that they could just “go back” from any decision, also seemed more likely to explore.

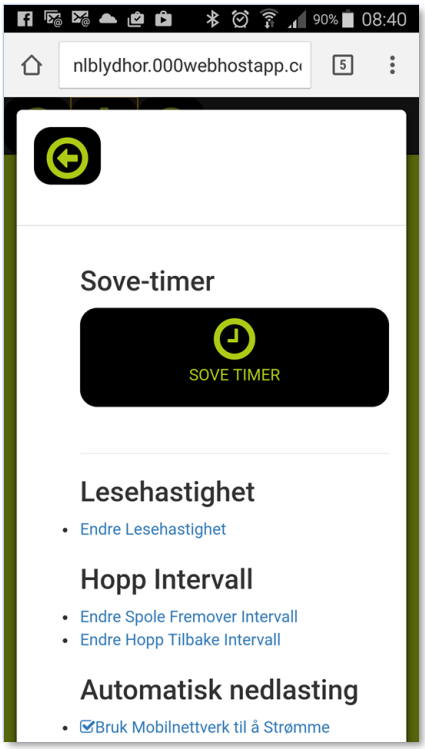


Figure 12.4:Prototype settings

12.1.8 General input from participants

At the conclusion of each user testing session, a general, open conversation was held where users could give any input about either version of the applications. These always started where the users could begin, and talk about whatever was most important to them. Sometimes, when the users were finished with input, questions were asked about particular actions observed during user testing, for further clarity about why they did something a certain way.

One of the topics that all participants were asked for their input about, was whether the way that information was presented in the prototype "bookshelf," would be a suitable way to display search results. This screen offered more information about books which respondents in the survey had frequently stated would be an improvement. The testers generally liked the idea of having more information, but did not like the fact that a single book occupied more than entire page. This would ultimately make sorting through search results even more time consuming.

There was one positive factor, which was the "mer om boka" (more about this book) button. The idea had been copied from the NLB website. Users liked the idea of having more information about a book available, but presented in a compact manner.

Most of the testers preferred sorting through search results that have only the title and author, like the Lydhør app presents it, but they were generally neutral about the picture of the book cover. Most had no preference as to whether it included the book cover, because it was so small on the phone screen anyways. However, if a small button (with

more details about a book) was presented with the title and author, it would be optimum, as long as it didn't occupy too much space and make fewer search results per page.

The button did distract some of the users with limited vision, and at first added more visual information that had to be processed. However, two of the users when asked, said that after the first time they learned what the button was, that it would likely not cause any further problems.

The help menu in the prototype was consulted often by testers when they were stuck. The testers generally liked the feature, because they could find the solution to any problems encountered during testing there. However, the menu was quite lengthy, and was tedious for some users to find what they were looking for. Throughout user testing this problem was observed, but no readily apparent solution came to mind.

One user did offer a great suggestion to improve it. She said it would help if the menu were searchable. Watching testers also inspired another idea. Rather than only explaining where a button was in the help menu, the button itself could be included there, (as was the case with the sleep timer), because users could just access it instead of having to find it.

Many took the time during the open discussion to address bookmarks. Bookmarks were obviously important, because it is difficult to find the place where you stopped reading, after the bookmark suddenly disappears. Many took the chance to voice their displeasure over this glitch, which apparently occurs quite often.

The fact that bookmarks could be named was also important for one user, because she often marked multiple places in books that she used when studying.

During testing, one tester was having noticeable trouble with the task of adjusting the speed at which books are read. She had found her way to the feature, but when attempting to adjust the speed, seemed confused. When questioned about whether or not she was having problems (she had successfully completed the task of changing the speed but kept pressing the plus button), she stated that she didn't know if it was working, because she was pressing the plus button but nothing was happening.

In fact, something was happening, but due to her vision limitations, she could not tell that the speed was changing, or what it was set to (reference Figure 12.5 below). The speed was displayed below, in AAA contrast but she didn't see it, so she couldn't see that it was changing. The text is prominent on the screen, but she was seeing a large grey block and had only "happened" on the plus button. She was also unaware that there was an OK button, because the thin outline was not distinguishable.

Figure 12.5: Speed adjustment

The younger dyslexic participant used the open discussion to voice his frustration with the play screen controls in the Lydhør application. It was also apparent that he didn't understand the difference between reserving a book and checking out a book. He also took the opportunity to ask why the Lydhør application had tabs, because he thought that only the Loans tab functioned and didn't understand the function of the other tabs. His mother that was with him during testing was also confused about this, and couldn't offer him any help. He also explained how much trouble disappearing bookmarks caused for him.

Three participants (one was not official, just the curious volunteer), mentioned that controls like those in Figure 12.6 below were preferred. The picture is a screenshot of the Podcast Republic[36] app, which uses one large play/stop button, and only two skip buttons, and their interval can be modified in the settings, or by holding the button down until the options appear. There is also a third skip button, but this is used to skip an entire podcast.

The reading speed is available and adjustable there also, and there is a slider display to indicate progress, which also can be used to "skip" within the content. These users all preferred this "style" of content navigation, and did not like the five-button system used in the Lydhør application. Notably, none of these three had vision limitations, and the reading speed button (1.0X) may be as difficult for some limited vision users to find or use as the speed display in Figure 12.5, because it is presented similarly.

[36] http://podcastrepublic.net/

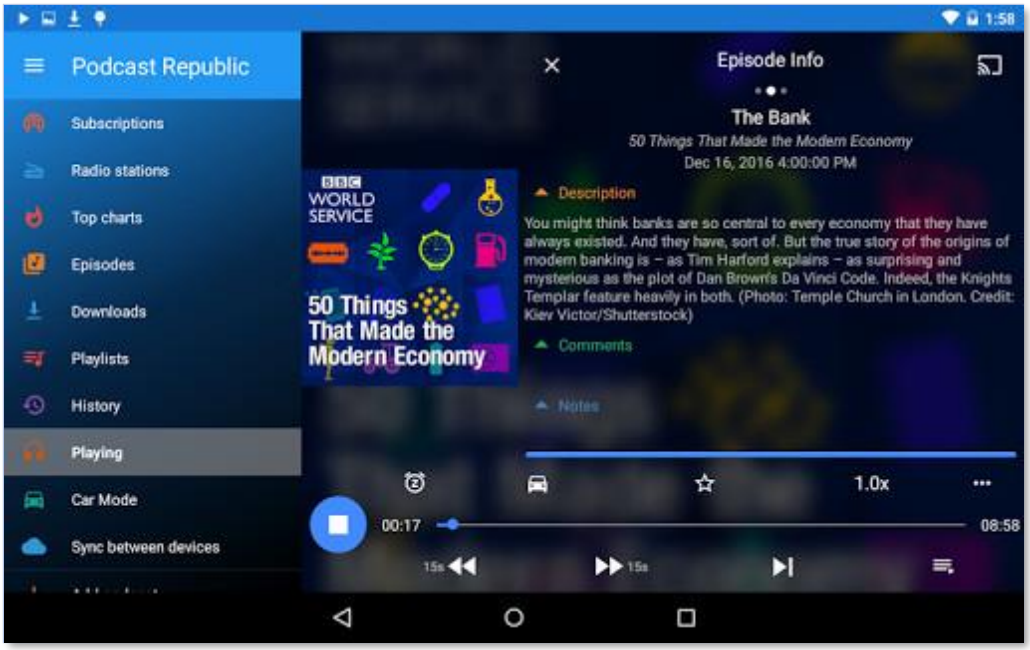


Figure 12.6: Podcast Republic controls

The three testers with vision limitations, when questioned about the Lydhør content navigation buttons (skip with increment system), all preferred this method. They were all experienced with it and preferred it, because it is what they had become accustomed to. They liked being able to easily modify the skip increments and did not like other applications that did not have this system.

One tester also pointed out that the way that the Storytel[37] application "systematizes" all of the available series (which a participant in the survey had also mentioned). She liked that they grouped their series, as can be seen in Figure 12.7 below. Once accessed the titles are shown in the order that they are written, which made it easy to access books in a series, without doing the tedious research required to find out the order of the series.

[37] https://www.storytel.com/

Figure 12.7: Storytel series presentation

Two users with vision limitations commented about the large play button used in the prototype. There were multiple tasks that required using the button to access a book. After they had identified it the first time, and were aware of it, it was an easy target to access visually, even for one user that mostly relied on the screenreader, because of its large size.

13 Discussion

13.1 Contribution to research base

There is substantial research dedicated to various aspects of human-technology-interaction related to vision limitations, much of which is dedicated to computer interactions, not mobile interactions. However, the other members of the print disability group (particularly with cognitive related issues such as injuries or ADHD or dyslexia) are underrepresented in research, and in the WCAG guidelines which make very few provisions for them. This research did serve to represent this underrepresented group.

The amount of research available related to print-disabilities in general, *with a focus on mobile applications* is rare. There is however, now a substantial conglomeration of research devoted to users within the print-disability group focused usually on one particular user-type and their interactions with technology, although again, typically with computer related interactions.

With the subject of this master's thesis tied to the Lydhør application, one of the general hopes was that the outcome of the research would benefit the print-disability group as a whole. This research began to explore the diverse and unique needs of the group, in the context of mobile use.

This research built a foundation for future work, which is necessary to ensure access to literature for all people. Being able to offer literary works via mobile technology, at this moment, could be considered a luxury. But in the future, most books will probably be found on mobile (or similar) technologies, as even trips to the bank are being replaced by access via mobile.

Through an iterative process that was driven by user input, the process of improving Lydhør (and the other similar applications offered by other nations to the print disability group), has been realized. Much like Brooks pointed out long ago in his now notorious article, a single solution or a "silver bullet" was not be realizable in this context, due to the complexity of the needs of users (Brooks, 1987). One neat, "catch-all" packaged solution is not presented.

However, application of universal design principles can serve to improve the accessibility and usability of the Lydhør application, mobile applications designed for the print disability group, and potentially mobile applications in general. Rather than identifying one broad solution, many small lessons or insights were gained as to how to follow a universal design approach that is considerate of the print disability group and mobile users in general.

13.2 The value of iterative user input for research

While reviewing literature (and studying a master's degree in universal design of ICT), the recurring theme of frequent user input is omnipresent, in many of the articles related to this research. Practical application of this advice, has shown the value of iterative user input. This may have been the one most valuable design lessons "learned"

during this research, that consulting with users most likely will show that some or many of your assumptions are faulty. For example, during this research, nearly ten days of effort, research, and testing, were nullified by a ten-minute session with a user.

These user interactions oftentimes do provide solutions, or at least a direction in which to aim possible solutions. So therefore, one of the most tangible "discoveries" from this research, is yet another validation of the power of consulting users, with design related task.

13.3 Lydhør and accessibility

For this research project, ensuring accessibility for Lydhør was the highest priority objective, and is mentioned first in the first research question. Lydhør was designed by an organization (or inputs from many organizations) that are hyper-focused on accessibility.

When they designed Lydhør (and the other similar apps), particular focus was likely given to users with no or limited vision, because the "print disability" group originally consisted of these members, and others, such as those with limited mobility were added later, upon recognition that their abilities to read "traditional" text were also impaired.

Despite that, some minor accessibility problems with Lydhør were identified during the research. The one that was probably the most problematic (during user testing) was the buttons/icons with a low contrast ratio. The WCAG guidelines (1.4.3) apply specifically to text on webpages [38].

Future revisions of WCAG guidelines, and the emerging guidelines being developed for mobile[39], should consider the contrast of elements such as buttons, and even check boxes, and *especially icons*, as icons are more often being used as functional buttons, particularly on mobile devices. For users that access the screen visually, but have a vision impairment, *being able to see and identify functional elements like buttons is equally as important as being able to read the text*.

The growing population of elderly in the world will mean an increase in vision deficiencies. With the growing need to access various business or government services via mobile (if one does not use a PC) the services will be more accessible and usable if the tiny functional elements on the screen are visible, via the use of a proper contrast ratio with icons/elements *against whatever background they are placed on*, not just with text. The need for a minimum icon size or "touch target sizing and spacing" laid out in the WCAG mobile guidelines, can also improve the experience for all users.

For now, the slate which designers have to work with, will be small screens that are typically less than ten square inches in area. Future designers should also try to de-clutter screens as much as possible, to allow for larger controls, by using several different

[38] https://www.w3.org/TR/WCAG20/
[39] https://www.w3.org/TR/mobile-accessibility-mapping/

"screens" to present their apps and webpages.

13.4 The idea of different "views" for different functions

One simple design consideration ultimately proved to be helpful for several reasons. This was the decision to change "screens" every time a new "function" began. The different functions were envisioned when designing the prototype. In the case of Lydhør, its' major functions were simply, finding a book and listening to a book. This was broken into three "screens," including a starting screen to allow users to choose between these two options, which worked well. The design intentionally utilized as many screens as possible.

The first reason for following this practice of changing screens with most actions, is that it is an excellent way to provide feedback for users. This was a problem identified in the survey, but it can also be helpful with many mobile applications being accessed on a bright sunny day, that are barely visible for all users. A new screen means that something has occurred, even if the wrong button has been pressed. Providing a return option, like the now classic "left arrow" made popular because of internet browsers, might ensure that they will explore more, and be less afraid to press a wrong button, because of the forgiveness that this gesture offers.

13.5 Most critical functions always available

More room allows designers the opportunity to add a consistent control, with the app's core functions, to every screen. This was something that all participants in testing enjoyed, and used.

A bar with three buttons was included on every page. The buttons offered an option to return (to a previous decision) to offer error forgiveness, and the help and settings portals. Users, during testing, frequently accessed all three of these options, which were envisioned beforehand as being important to always have available. The users learned the buttons quickly after exploring the first time, possibly because there was only three, which was easy to remember. The return (left arrow) button was used most frequently. But having a help menu also empowered users during this testing, giving them a solution to problems they would otherwise have to search for unassisted.

13.6 Put buttons in the help menus?

During testing, the help menu (except for the sleep timer button) was created in "classic" fashion. "To activate X, go to X and do X," which directed users how to find what was usually ultimately going to be a button. Accidentally, the sleep timer button, was place in the help menu without thought. During testing, after seeing that guidance to other buttons was inferior to the sleep timer button, which was in the sleep menu, the thought was "Why not just put all the buttons there?" This practice is already being exercised, although previously unnoticed, with the software (Word) that this document is being prepared in.

13.7 Is text needed on buttons?

Another accessibility problem found within the existing Lydhør application was failing contrast for text on buttons. The high contrast text used on buttons was a notable improvement from Lydhør. However, the users, during user testing, had no problem with buttons that displayed only an icon, and no text.

Designers should consider the need for the use of text on buttons. Many of the survey respondents indicated that they used iPhones, which generally favors use of textless icons as buttons (although there are some functions that are “buttons” in the form of actionable text with an arrow beside it).

13.8 Icons: recognition as buttons

Blind and limited vision users can receive information about *properly labeled* elements via screenreaders. Dyslexic users may prefer non-text navigation. And the population as a whole should be becoming more familiar with the use of icons as buttons, as every modern smartphone already uses icons as the sole method of accessing many of the applications. Use of icons to access functions within an application should not be that different than accessing the apps themselves. However, this practice should be approached with caution, because icon recognition is twofold; recognition of its purpose, and being able to visually recognize it.

13.9 Icon purpose recognition

This concept of icon purpose recognition was explored for this work, but not extensively. However, during the user testing, there was a small confirmation that misunderstood icons may never be used or even checked to learn their function.

13.10 Icon visual recognition

What was important for this research was whether icons could be visually recognized, meaning, could users tell what the icon was, particularly those who have vision deficiencies.

During user testing, even though the icons used were approximately 4mm tall with padding, some were still difficult for some users to see during testing. One which was only mildly complex, was removed after the first user test. This was a text sized play button (the same size of the font), which was a simple triangle in a circle. It took the first tester almost a minute to see the shape, because she was seeing only a small portion of it at a time. This tester said it easier for them to see had it only been a simple triangle.

Figure 13.1: Play button

Interestingly, this same icon was also present on the same screen, in a much larger button, which is nearly as large as it is shown in Figure 13.1, and the large one was not

problematic, and some testers made comments about the ease of activating and seeing the large one.

The decision to use simple silhouette type icons for the prototype, was made because they would hopefully be easier to process for those with limited vision.

Some prior research regarding icon recognition (purpose) had established that simpler icons were more easily recognized. However, their simpler outline icons could possibly create for problems for visually impaired users, so these icons were not utilized for this user testing. However, they also tested silhouette icons, and when they used a shape that closely resembled the intended item, these were also highly recognizable (Guo, 2016).

The statement they were more recognizable when their shape resembled the intended item is used, because, when examining their research, the “silhouettes” that were less recognized, did not resemble their matching real world item. A good example, is to check the camera icon that these researchers use, which absolutely does not resemble a camera. The silhouette icons that did have a strong resemblance of their corresponding item, were highly recognizable (Guo, 2016).

Because silhouette icons would be more recognizable for visually impaired users, they were used in the prototype. As soon as complexity was added to icons, they became hard to perceive. In general, for mobile screens, this testing showed that users of various levels of vision could both see and recognize simple, high contrast silhouetted icons, *when their shape resembled the shape of the object one was intending to portray*. This object must be familiar, for example because it is used often with many mobile applications and thereby recognizable, such as the settings gear (cog) or a play button triangle, or have a universal meaning.

Even icons that are not frequently used with mobile, that are recognizable objects tied to a meaning, can be understood readily if they convey meaning, such as the bookmark used in both the prototype and Lydhør.

The results of this study, when considered with the previous research that inspired the line of thought, would argue for use of simplistic icons, with little detail. Using easily understood silhouettes seemed to enhance perception of icons for visually impaired users, during testing.

For Lydhør, the recommendation would be to either replace the complex branching icon, or offer a tutorial or help menu that can educate users on the function of this icon.

13.11 Use a framework that supports scalability

The WCAG mobile recommendations specifically state that the pinch-zoom gesture in mobile browsers, should not be blocked. However, this new proposal is silent about mobile applications, which the new guidelines should include.

Lydhør was not compliant with WCAG requirement for scalability (although it is

technically not applicable). This made users that need to zoom to see the letters have to rely on the zoom that is a feature on both android and apple phones. The Android zoom only offers 3X scaling while the Apple products zoom at a higher rate (precisely how much is unknown).

The scaling feature was used often by users during user testing, including occasionally those that did not have vision problems. This gives more importance to the WCAG guideline 1.4.4, which says that developers should use only tools that allow scaling, for both internet connected apps and regular applications[40].

However, some discussions of meeting this requirement, such as G179 (A technique used to meet this guideline), may need some rethinking. This revolves around the idea that text should wrap as it is zoomed (which is a built-in feature of many responsive web design frameworks such as bootstrap). During user testing, some users, rather than activating the screenreader, preferred a very high level zoom, and were reading sometimes 2-3 letters at a time. If the text had "wrapped" at this point, the users would be forced to read one letter at a time, but the letters would now be vertically stacked. The users employing the zoom during user testing, and likely expected the letters to be horizontally oriented.

13.12 One area of caution with ARIA roles

One problem with ARIA labeling (via aria-roles) for screenreaders was identified during user testing. The icons used in the prototype were "hidden" so that screenreaders would not redundantly repeat the element title. The icons that were buttons themselves were contained within another element, and that element announced its purpose.

However, using this configuration can cause both elements to be hidden to VoiceOver (only), which was observed during user testing. This causes the elements to be completely useless for those not using vision to navigate. Designers are recommended to avoid this hiding an icon inside of "announced" button for web-based elements.

13.13 Most important items "first"

In scholarly articles, there is often a discussion about elements that are placed on the top and left of webpages being perceived as important by visual users. Typical screenreaders on PCs generally work from the top to bottom, left to right, when tab indexes have not been adjusted (although this can vary depending on the order webpage elements have been created and would not include for example Arabic languages, which are read in an opposite direction).

What was interesting during testing, was that the user relying on the screenreader also generally followed this trend on a mobile screen, with exploration beginning at the top left.

Mobile phone screenreaders work quite differently in that users with no/limited vision

[40] https://www.w3.org/TR/WCAG20/

"feel" their way around the different elements on the screen. As they touch the screen elements, the screenreader announces the information. This has somewhat streamlined the process of exploration, at least what was witnessed during user testing, because as the screenreader began to read something not of interest to the user, she could move quickly to the next element.

13.14 Searching/Browsing for books in Lydhør

One of the overwhelming responses in the user survey, was that if users wanted more information about books when searching for new books. Users often mentioned that they would like more information, or information that was presented in a better manner about series. (Note that these responses were prompted because one question asked specifically about this, and that there was no distinction between searching and browsing). They also stated often that they wanted to know more about a book, such as having the "back cover" short summary, genre, year published, even recommendations for similar books.

There are typically two methods that one uses to find new books, which is searching for a particular book, or browsing through a collection to find a one that they are interested in. It is only possible to search for specific titles in the Lydhør application (not browse). There were several responses in the user survey where respondents indicated that they never used the Lydhør app for finding new books, they used the NLB website instead. During research for the prototype development, the NLB website was consulted to learn what was different there.

Although it had some problems (particularly when scaled down to a mobile screen), the website had many good characteristics. More information about a book is provided there when searching. The way the search results are presented was also what users survey respondents wanted for Lydhør. One feature of interest was a button in each book result that when clicked gave more information about a book.

"Browsing" for books was also possible, via three age categories, and then several subcategories within the age groups, which were different dependent on the age group, and tailored to each group.

The recommendation for NLB for future versions of Lydhør, would be to consider using their website for searching for books, in the Lydhør application. Great detail has already been placed on guaranteeing accessibility on the site. Users indicated satisfaction with the site in the survey.

Rather than "reinvent the wheel," the NLB could make some modest alterations to their website, particularly the way the site is displayed on small screens, and use it for the Lydhør app. This offers economies of scale in that employees can focus all design efforts on one search apparatus. The site is already built in a responsive framework, which will make improving the small screen presentation even easier.

Adjusting for smaller screens should include hiding the large navbar that offers the

option to search within the NLB website or the list of books. This feature was often problematic during user testing. Users that are on the book search page were searching only for books so it is not necessary there, although searching the site may be useful for other pages on the website.

The search bar should also be modified for small screens. During user testing the smaller iPhone screens in normal vertical orientation allowed for only very few letters to be entered before filling the box. Here, either a smaller button or a separate button to activate the search would allow more space for the search terms (some of the users wanted to "double check" their search terms before committing, and panning the text was difficult to figure out how to do).

Finally, the buttons which allow users to search (or browse eventually) within categories, is a feature that users want. The initial choice that users are presented with are the children's, teenager, and adult book categories.

After one of these three choices have been decided, users will typically not need to decide these choices again. However, after one of the three have been selected, they are presented again, and occupy the entire bottom of the screen, as seen on the left side of Figure 13.2 below (Adults/Voksen has already been selected in this picture). This should be programmatically removed.

Removing these would make room for the other category buttons that are currently presented further down the page, as seen on the right side of Figure 13.2. It is worthy to note that no user testing was performed with the icons/text for these buttons, and they may be problematic for users, and require redesign. A layout that was quite similar to this, was problematic for elderly users and was discussed in the Elderly users section.

Figure 13.2: NLB browsing features

13.15 Improving search autosuggestions

During prototype development testing, the Google search engine performed well with the errors that print disabled users typically make. The cost of utilizing it (without ads) may be free or very little since the NLB is a non-profit organization. If possible, it could be incorporated into the site, as an easy fix to improve searching suggestions in a powerful manner.

This would likely require modification of the way the search results are displayed, which is the topic of the next subsection.

If it is not possible to include it, efforts should be focused on improving the existing search function, particularly the problem where autosuggestions disappear, and more effort should be applied in fixing how it reacts to typos and letter transpositions.

13.16 Displaying search results

Wordy results have negative ramifications for dyslexic users, and are not simplistic, which can affect users with cognitive impairments. Both assertions are backed by previous research.

Users overwhelmingly desired various information in addition to a book's title and authors. Despite the fact that there were many survey responses that indicated a desire for more information about a book, this was shown during user testing to be more complex. A possible solution was explored, the button that can be designed to occupy a small space, which only gives more information about a book when desired.

However, when the button idea was tested, more information other than the title and author was presented, which made one search result occupy the entire screen. Exploration of the search results was a tedious process. Users wanted to be able to quickly scan multiple search results.

The best possible combination of search result presentation format, (in Lydhør, not the website) seemed to be to provide the title and author, with the rest of the information stored in some form of collapsible format like the button.

In the case of the NLB, implementation of either the "more about this book" button would not require large effort by developers, because additional information is already stored about books on their website, and the button is already implemented there as well.

The need for displaying book covers is still debatable. They do enhance the experience for visual users. However, they reduce available space on the screen, which would detract from the experience for visual users.

And they cause the previously discussed dilemma of either being redundantly read aloud twice for screenreaders. If they are not "double labeled," they will be confusing for users with limited vision, that see the only that the item exists, but cannot get information about what the item is.

13.17 Books in series

To aid users with the problems associated with pairing and ordering books that are part of a series, some thought should be given as to the possibility of grouping these books, possibly in a manner similar to what was discussed before. This would reduce the research required for users that want to know which order the books were written in.

This would resolve an issue that was mentioned often by users. However, more consideration should be considered about placing entire series in search results, because of the overcrowding that will occur. One possibility is to encapsulate the information in the more about this book button.

13.18 Tutorial/help/FAQ

During user testing, the availability of a help menu proved to be useful. Because of the short duration of the user testing, “training” the users was not possible, but could also be a viable option. Training can be accomplished via pop-up tips (with the typical “do not show this again” checkboxes, although this concept would require further testing.

Many survey respondents indicated that they would like to have a tutorial. This option should be explored, and could create fluent users, faster. A help menu is certainly needed at a minimum.

This could help reduce the complexity surrounding the reserve versus checkout and the download versus streaming processes. It also helps the users find functions. A “searchable” help menu can reduce the necessary scanning of a large help menu, and placing the buttons in the search menu, (instead of directions to find a button as previously discussed) is almost a guaranteed improvement.

13.19 Skip button possibilities

User testing showed that removing the skip “increment” buttons from the Lydhør screen was detrimental for some users. Limited vision users that were accustomed to this arrangement, preferred to use it.

The other testers did not like it, and preferred a setup like Podcast Republic uses, as previously discussed. The contrasting desires of the different users can all be accommodated, by programming a choice of play button controls into Lydhør, that includes both styles, possibly even more.

Participants during user testing pointed out that (unlike previously assumed), the skip button is often accessed during reading, and with either configuration, should be readily available and easily modified on-screen.

13.20 Other play screen controls

The reading speed of books is another frequently used control. So, a “button” allowing easy access to change the reading speed should be available on the screen when users are listening to books. The speedometer icon that Lydhør uses was generally not recognized as the right button for this purpose.

The style of control that Podcast Republic and other players use, which some testers mentioned during user testing as a preferred mechanism, is shown in Figure 13.3 below.

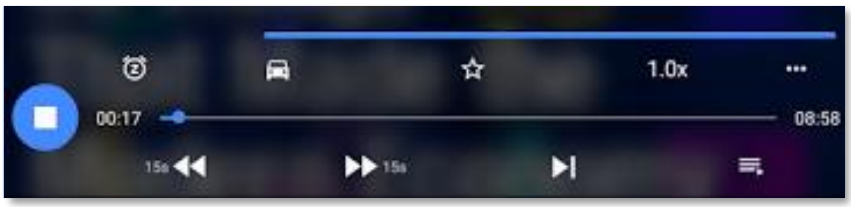


Figure 13.3: Podcast Republic controls (2)

Since, the speedometer icon was poorly understood by many users (and pre-testers), the 1.0X button that is seen in the picture above may offer an option for a better speed adjustment button. It should be "familiar" to some users already. A proper, functional help menu could also help lead users to this feature.

However, more thought should be given as to how to present it. Presenting it as shown in Figure 13.3 is problematic, because the control is isolated and not easily visible or identifiable, and could replicate the problem that is discussed directly before, and displayed in Figure 12.5.

There is one final note about the sleep timer (z in a clock) seen in Figure 13.3. This type of icon was desired for use with the prototype, because it seemed to convey the meaning of the sleep timer better than the clock used for the prototype icon (which some users mistook to be the reading speed adjustment). However, the selection of free icons available for use with the prototype was limited.

13.21 Text input on mobile devices

If it were possible to poll all mobile users worldwide, it is likely that an overwhelming majority of users would consider text entry on mobile phones less than optimal because of the tiny on-screen keys.

During the user testing, a realization was made that visual users generally do not need to see whatever current screen they are using when they are inputting text in their phones. They typically only need to see the keyboard buttons, and the text that they are inputting.

Future developers and researchers should consider this, as it stands to improve usability. When users begin a text input process, only the box where the text is entered and the keyboard keys are necessary. This makes it possible to have larger keyboard buttons, and, importantly, means that the text that is being entered can be displayed even larger.

However, consideration should be given as to where to place autosuggestions, such as on the side of the screen, or in a "reserved" place under the text box for example.

13.22 Dedicating research to physical e-reader devices

The lone research endeavor located, that was similar to this research, was focused mainly on noting differences between dedicated, physical e-reader devices, and

exploring possible adaptations to these devices. Notably, mobile phones are accomplishing e-book access, without requiring the purchase of a separate, sometimes expensive device. These days, many people already own a mobile phone. For those that don't, a mobile phone is cheaper to purchase than a dedicated e-reader device.

The mobile phone has already replaced made many things obsolete, since so many people own them. For example, home landline telephones are being unplugged, and calendar software is becoming obsolete, because their mobile phone equivalents are portable and are usually in range of their owners, even when they are not at home.

Rather than focusing research efforts on a device that will likely be made obsolete by mobile technology, research should instead be devoted to e-reader mobile technologies. Research related to e-book formats, especially the EPUB formats, should be continued, because even the DAISY consortium is moving in that direction[41]. They are moving efforts away from the dedicated, physical, DAISY e-book readers. Instead they are focusing on EPUB advancement, because they too, understand that this will be the trend in the future. The EPUB format offers the possibility for meaningful advancement of accessibility of e-books, but this discussion is complex enough to warrant its own dedicated Masters or even Doctoral research project ;).

13.23 Other miscellaneous recommendations

There are two recommendations not previously mentioned that add small usability or experience improvements with Lydhør that were requested by users. One would like to have a "push" notification when a reserved book becomes available, similar to the ones received when you get a text message on your phone.

The other improvement was to ensure that Lydhør removes books when finished, after confirming this with the user. Some of the more technically savvy users learned that the application rarely, if ever truly removes the book files from the phone (because they had investigated why their phone's memory was being overfilled). This is unconfirmed but was mentioned a few times in survey results.

Finally, although this last discussion is outside of the scope of this research some other items noted during this project with Lydhør deserve a short discussion. Improving the stability of the Lydhør application is imperative. Stability related problems of all varieties were by far the most frequently noted issues throughout the research. Bugs existed in all functions of the application.

The Lydhør application is quite complex programmatically, and it is built with technology that is now "dated" by today's mobile tech standards. The NLB could perform an inexpensive feasibility assessment to ascertain the costs of redesigning the application. There is a possibility that, comparatively, the costs addressing all of the bugs the

[41] This information was provided to me, by a member of the DAISY consortium board of directors, who wished to remain anonymous.

software has, may far exceed the cost of starting from scratch.

Starting from scratch also affords the possibility of sharing portions of code from applications with similar organizations around the world. Many of which many have, for some reason, created their own version of e-readers (rather than all of the organizations applying their cumulated efforts together towards creating one “sharable” piece of software). Reaching out to the open source community is also a viable option, as many open source technologies, offer free, stable, software, that brings together knowledge from around the world.

Also, a firm search technology, with very little need for modification required already exists, the NLB website. A new web app could easily be tied to this site. Newer text-to-speech components can be used for book reading, which were created for modern versions of phone OS’s (not for versions that are now several years old). And finally, this would allow for updating of the anti-piracy watermarking technologies, all of which become less secure as time passes.

Piracy is the key problem standing in the way of having publishers fully embrace accessible, electronic book formats. If book publishers are offered more secure protections against piracy, they will be more generous with libraries tailored to the print disability community.

If they indeed embrace an accessible e-book format, they can instantly reduce the incredible amount of effort that many organizations around the world exert to make books accessible, “with the flick of a switch”.

14 Limitations

The most troublesome limitation of this project was the lack of representation for the less frequently represented members of the print disability group. Finding “reputable” literature that addresses this group (in the context of this research), proved difficult. Although their voice may have been heard in the survey results, and even though all of the volunteers for user testing with this project were accepted, members of the print disability group such as those with cognitive impairments from ADHD or injury, were once again, underrepresented, even if they only make up the minority portion of the demographic. There were some findings that can be applied to improve Lydhør, print-disability research, and possibly applied in the WCAG guidelines, in reference to this minority of the demographic, but not as many as were intended at the outset of the project.

The relatively small number of participants involved with testing was another limitation. This group had responsibility for being the voice for a large, diverse group. With only four respondents (and a fifth Lydhør user that was recruited late in the testing phase), the diversity and size of the group was limited. Regarding academic research and research in general this is a small sample size.

However, Nielsen’s empirical study of usability research, where he suggests five users as an ideal sample size with usability research, somewhat mitigates the sample size deficiency. Typically, 80% of usability problems he says are discovered with this sample size. The research is a cost efficient and effective way to improve software when time and budget are limited (Nielsen & Landauer, 1993).

The five-user recommendation is not perfect, however, because Nielsen later states that this number is not optimum for “diverse” user groups. But the research did utilize an iterative process of gaining user inputs, which is a key for the “five user” recommendation (Nielsen, 2000).

Another likely criticism for this research would be its narrow scope. It does revolve around one application, created for a specific audience, that has one purpose, to make it possible to listen to books. To answer that type of criticism, it is important to note that for those who cannot read traditional books, having a mechanism to do so is critical. Having one that works well, is deserved. There are also many applications that serve the same purpose as Lydhør, typically offered by governments around the world. Lydhør is in fact a modified version of one of these applications, and this research can help to improve the others.

15 Conclusion

At the outset of this research project, a challenge was available, which offered the opportunity to apply universal design principles to potentially improve the Lydhør application (and other similar purpose applications). This project would also add to the scant amount of similar research available concerning the subject matter, which is a relatively new area of focus, academically speaking.

The first research question set out to learn if, first and foremost, there were any accessibility problems with the Lydhør app, and secondly if there were any usability problems that could be identified. The various research activities did indeed identify a few accessibility issues, and many usability related issues, which answered the first research question.

The second research question intended to find potential solutions to the problems identified in relation to RQ1. This was no simple task. In many cases, simply imagining a subjective solution would have been faulty, and the solutions may have done little to improve the original problems.

Actual members from the print disability group, were consulted several times. The practice of consulting with them *iteratively*, paid dividends. This practice was humbling, and showed that some of the perceived improvements created to answer RQ2, were in fact, not improvements. However, identifying flaws in the RQ "answers," should not be viewed negatively. These flaws, unearthed during the final user testing phase, actually strengthened the process of "answering RQ2," because they were identified and future researchers and developers will know what does, and does not work.

In conclusion, the "answer" to RQ1 can be found in the results of all five research methods or activities (and the design operations), in the form of many problems that were identified. The problems identified are applicable and can be extended to various facets of mobile technology design, which can improve mobile use, for the print disability group, and for all mobile users.

The potential solutions sought after in RQ2 have been identified and presented as well, even if some of them ultimately only became another problem for answering RQ1. The prototype offers a good model for the NLB to base a future version of Lydhør on, or to use to make the current Lydhør more accessible and usable.

The "answers" to RQ2, again spread throughout the results (and design) sections, can also be utilized by researchers and developers to make improvements for the print disability group.

16 Potential future research

16.1 Search result improvements

Testing the presentation of search results, for print disabled users, particularly with the way book results are presented by e-reader apps on mobile phones, certainly offers more research opportunities. A model to test, that may be viable, was presented in the discussion section of this thesis. The NLB or similar organizations, or researchers, can perform testing to answer the "unanswered" questions outlined there.

16.2 Context sensitive tutorials

The unreal engine is popular software, offered by Epic Games,[42] that allows everyone, including people with limited programming skills to create sophisticated games. Training users from various backgrounds, how to use their sophisticated software, is, needless to say, a monumental task. Yet, many people often successfully create outstanding games with the unreal engine.

The secret to success is probably due to the fact that Epic uses a "context sensitive" tutorial system, that offers advice about specific tasks and features, when it notices users utilizing them.

This concept could of adding context sensitive guidance, could be explored by future researchers, particularly those researching universal design (IT) topics. Such tutorials, might have a strong potential to increase software accessibility, particularly software displayed on tiny mobile phone screens. The implications become more important, the more sophisticated a particular piece of mobile software is. But the interactions of these proposed tutorials, require further exploration, specifically their interactions with assistive technologies.

16.3 Screen readers on mobile devices

The habits of screen reader users, in the context of mobile use, was an opportunity that was noticed during this research. Interestingly, there seemed to be a pattern followed by screen reader users, that was similar to the way reading, in many (not all) languages occurs. They started exploration with their fingers at the top left of the screen, and moved down and to the right.

This could prove to have significant implications about universal design/mobile accessibility topics, particularly with research about visual impairments.

[42] https://www.epicgames.com/

17 References

Barros, A. C. d., Leitão, R., & Ribeiro, J. (2014). Design and Evaluation of a Mobile User Interface for Older Adults: Navigation, Interaction and Visual Design Recommendations. *Procedia Computer Science, 27*, 369-378. doi:http://dx.doi.org/10.1016/j.procs.2014.02.041

Benedito, J., Guerruero, T., Nicolau, H., & Goncalvez, D. (2010). *The key role of touch in non-visual mobile interaction*. Paper presented at the Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, Lisbon, Portugal.

Berget, G., Mulvey, F., & Sandnes, F. E. (2016). Is visual content in textual search interfaces beneficial to dyslexic users? *International Journal of Human-Computer Studies, 92–93*, 17-29. doi:http://dx.doi.org/10.1016/j.ijhcs.2016.04.006

Berget, G., & Sandnes, F. E. (2015a). The Effect of Dyslexia on Searching Visual and Textual Content: Are Icons Really Useful? In M. Antona & C. Stephanidis (Eds.), *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being: 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part III* (pp. 616-625). Cham: Springer International Publishing.

Berget, G., & Sandnes, F. E. (2015b). On the Understandability of Public Domain Icons: Effects of Gender and Age. *ODA HIOA*.

Berget, G., & Sandnes, F. E. (2015). Searching databases without query-building aids: implications for dyslexic users. *Information Research, 20*(4), 18-34.

Berget, G., & Sandnes, F. E. (2016). Do autocomplete functions reduce the impact of dyslexia on information-searching behavior? The case of Google. *Journal of the Association for Information Science and Technology, 67*(10), 2320-2328. doi:10.1002/asi.23572

Bernard, R., Sabariego, C., & Cieza, A. (2016). Barriers and facilitation measures related to people with mental disorders when using the web: A systematic review. *Journal of Medical Internet Research, 18*(6). doi:10.2196/jmir.5442

Beyene, W. (2016). Resource Discovery and Universal Access: Understanding Enablers and Barriers from the User Perspective. *knowledgearc*. doi:http://hdl.handle.net/10642/4604

Blechner, A. J. (2015). Improving Usability of Legal Research Databases for Users with Print Disabilities. *Legal Reference Services Quarterly, 34*(2), 138-175. doi:10.1080/0270319X.2015.1048647

Brooks, F. P. J. (1987). No Silver Bullet Essence and Accidents of Software Engineering. *Computer, 20*(4), 10-19. doi:10.1109/MC.1987.1663532

Buxton, B. (2007). *Sketching User Experiences*. San Franciso, CA: Elseveir.

Cáliz, D., Alamán, X., Martínez, L., Cáliz, R., Terán, C., & Peñafiel, V. (2016) Examining the usability of touch screen gestures for elderly people. *Vol. 10069 LNCS. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 419-429).

Casare, A. R., Silva, C. G. d., Martins, P. S., & Moraes, R. L. O. (2016). *Usability heuristics and accessibility guidelines: a comparison of heuristic evaluation and WCAG*. Paper presented at the Proceedings of the 31st Annual ACM Symposium on Applied Computing, Pisa, Italy.

Charoenchaimonkon, E., & Janecek, P. (2015). Characterizing non-visual target acquisition tasks with the aid of a tactile display: investigating factors beyond the classical Fitts' theorem. *Universal Access in the Information Society, 14*(4), 459-475. doi:10.1007/s10209-014-0352-5

Charters, E. (2003). The Use of Think-aloud Methods in Qualitative Research

An Introduction to Think-aloud Methods. *Brock, Vol. 12*(No. 2), 68-82.

Chen, C. J., & Keong, M. W. Y. (2016). Affording inclusive dyslexia-friendly online text reading. *Universal Access in the Information Society*, 1-15. doi:10.1007/s10209-016-0501-0

Chen, W. (2013). Gesture-Based Applications for Elderly People. In M. Kurosu (Ed.), *Human-Computer Interaction. Interaction Modalities and Techniques: 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part IV* (pp. 186-195). Berlin, Heidelberg: Springer Berlin Heidelberg.

Chourasia, A., Sesto, M., Kuehn, K., & Wiegmann, D. (2017). Effects of activation method and mental effort on touchscreen task performance for users with and without upper extremity motor control disabilities. *Universal Access in the Information Society, 16*(2), 469-481. doi:10.1007/s10209-016-0471-2

Craven, J., & Booth, H. (2006). Putting awareness into practice: practical steps for conducting usability tests. *Library Review, 55*(3), 179-194. doi:doi:10.1108/00242530610655984

Davies, D. K., Stock, S. E., King, L. R., Brown, R. B., Wehmeyer, M. L., & Shogren, K. A. . (2015). An interface to support independent use of facebook by people with intellectual disability. *Intellectual and Developmental Disabilities, 53*(1). doi:https://login.ezproxy.hioa.no/login?URL=?url=http://search.proquest.com.ezproxy.hioa.no/docview/1654943416?accountid=26439

Dawe, M. (2007). *Understanding mobile phone requirements for young adults with cognitive disabilities*. Paper presented at the Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, Tempe, Arizona, USA.

Díaz-Bossini, J.-M., & Moreno, L. (2014). Accessibility to Mobile Interfaces for Older People. *Procedia Computer Science, 27*, 57-66. doi:http://dx.doi.org/10.1016/j.procs.2014.02.008

Dim, N. K., & Ren, X. (2014). Designing Motion Gesture Interfaces in Mobile Phones for Blind People. *Journal of Computer Science and Technology, 29*(5), 812-824. doi:10.1007/s11390-014-1470-5

Eli Toftøy-Andersen, J. G. W. (2011). *Praktisk Bruker-Testing (Practical User Testing)*.

Fuglerud, K. S. (2014). *Inclusive design of ICT: The challenge of diversity.* University of Oslo.

Gould, J. D., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Commun. ACM, 28*(3), 300-311. doi:10.1145/3166.3170

Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., & Cajander, Å. (2003). Key principles for user-centred systems design. *Behaviour & Information Technology, 22*(6), 397-409. doi:10.1080/01449290310001624329

Guo, Q. (2016). An icon recognition study on different simplicity

levels.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH. *MIS Quarterly, 28*(1), 75-105.

Horton, W. K. (1994). *The ICON Book: Visual Symbols for Computer Systems and Documentation*: John Wiley \\& Sons, Inc.

Huang, P.-H., & Chiu, M.-C. (2016). Integrating user centered design, universal design and goal, operation, method and selection rules to improve the usability of DAISY player for persons with visual impairments. *Applied Ergonomics, 52*, 29-42. doi:http://dx.doi.org/10.1016/j.apergo.2015.06.008

Ibrar Hussain, L. C., Hamid Turab Mirza,, & Gencai Chen, S.-U. H. (2014). Right mix of speech and non-speech: hybrid auditory feedback

in mobility assistance of the visually impaired. *Universal Access in the Information Society, 14*(4), 527-536. doi:10.1007/s10209-014-0350-7

Izzo, M. V., & Bauer, W. M. (2015). Universal design for learning: enhancing achievement and employment of STEM students with disabilities. *Universal Access in the Information Society, 14*(1), 17-27. doi:10.1007/s10209-013-0332-1

Kiger, J. I. (1984). The depth/breadth trade-off in the design of menu-driven user interfaces. *International Journal of Man-Machine Studies, 20*(2), 201-213. doi:http://dx.doi.org/10.1016/S0020-7373(84)80018-8

Kyrkjebø, A. (2016, 22/12/2015) *Interview/Interviewer: T. Godwin*.
L. Maatta Stephanie, J. B. L. (2014). An evaluation of the functionality and accessibility of ereaders for individuals with print disabilities. *The Electronic Library, 32*(4), pp.493-507. doi:10.1108/EL-01-2013-0012
Lazar, J. F., Jinjuan Heidi; Hochheiser, Harry. (2010). *Research Methods in Human-Computer Interaction* Wiley.
Leporini, B., Buzzi, M. C., & Buzzi, M. (2012). *Interacting with mobile devices via VoiceOver: usability and accessibility issues*. Paper presented at the Proceedings of the 24th Australian Computer-Human Interaction Conference, Melbourne, Australia.
Lowdermilk, T. (2013). *User Centered Design*: O'Reilly
Lundh, A. H. (2015). The Use of Digital Talking Books by People with Print Disabilities: A Literature Review. *Library Hi Tech 33*(1), 54-64. doi:http://dx.doi.org/10.1108/LHT-07-2014-0074
Maatta, S. L., & Bonnici, L. J. (2014). An evaluation of the functionality and accessibility of e-readers for individuals with print disabilities. *The Electronic Library, 32*(4), 493-507. doi:doi:10.1108/EL-01-2013-0012
McCarthy, J. E., & Swierenga, S. J. (2010). What we know about dyslexia and Web accessibility: a research review. *Universal Access in the Information Society, 9*(2), 147-152. doi:10.1007/s10209-009-0160-5
McNiff, J. (2013). *Action Research*. Florence, UNITED STATES: Taylor and Francis.
Mi, N., Cavuoto, L. A., Benson, K., Smith-Jackson, T., & Nussbaum, M. A. (2014). A heuristic checklist for an accessible smartphone interface design. *Universal Access in the Information Society, 13*(4), 351-365. doi:10.1007/s10209-013-0321-4
Miao, M., Pham, H. A., Friebe, J., & Weber, G. (2016). Contrasting usability evaluation methods with blind users. *Universal Access in the Information Society, 15*(1), 63-76. doi:10.1007/s10209-014-0378-8
Nandigam, D., Symonds, J., Kayes, N., & McPherson, K. (2010). *Mobile phone user interface design for patients with traumatic brain injury*. Paper presented at the Proceedings of the 11th International Conference of the NZ Chapter of the ACM Special Interest Group on Human-Computer Interaction, Auckland, New Zealand.
Newell, A. F., Gregor, P., Morgan, M., Pullin, G., & Macaulay, C. (2011). User-Sensitive Inclusive Design. *Universal Access in the Information Society, 10*(3), 235-243. doi:10.1007/s10209-010-0203-y
Nielsen, J. (1993). *Usability Engineering*: Wiley.
Nielsen, J. (1994). *Enhancing the explanatory power of usability heuristics*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, Massachusetts, USA.
Nielsen, J. (2000). Why You Only Need to Test with 5 Users. Retrieved from https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/
Nielsen, J., & Landauer, T. K. (1993). *A mathematical model of the finding of usability problems*. Paper presented at the Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands.
Nielsen, J., & Molich, R. (1990). *Heuristic evaluation of user interfaces*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, USA.
Oehl, M., Dahlmanns, L., & Sutter, C. (2013). Small Input Devices Used by the Elderly – How Sensorimotor Transformation and Task Complexity Affect Interaction. In C. Stephanidis & M. Antona (Eds.), *Universal Access in Human-Computer Interaction. User and Context Diversity: 7th International Conference, UAHCI 2013, Held as Part of HCI International 2013, Las Vegas,*

NV, USA, July 21-26, 2013, Proceedings, Part II (pp. 181-190). Berlin, Heidelberg: Springer Berlin Heidelberg.
Panchanathan, S., & McDaniel, T. (2015). Person-centered accessible technologies and computing solutions through interdisciplinary and integrated perspectives from disability research. *Universal Access in the Information Society, 14*(3), 415-426. doi:10.1007/s10209-014-0369-9
Passini, S., Strazzari, F., & Borghi, A. (2008). Icon-function relationship in toolbar icons. *Displays, 29*(5), 521-525. doi:http://dx.doi.org/10.1016/j.displa.2008.07.001
Persson, H., Åhman, H., Yngling, A. A., & Gulliksen, J. (2015). Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. *Universal Access in the Information Society, 14*(4), 505-526. doi:10.1007/s10209-014-0358-z
Rello, L. (2015). *Dyslexia and web accessibility: synergies and challenges*. Paper presented at the Proceedings of the 12th Web for All Conference, Florence, Italy.
Rocha, T., Carvalho, D., Bessa, M., Reis, S., & Magalhães, L. (2016). Usability evaluation of navigation tasks by people with intellectual disabilities: a Google and SAPO comparative study regarding different interaction modalities. *Universal Access in the Information Society*, 1-12. doi:10.1007/s10209-016-0489-5
Rodrigues, É., Carreira, M., & Gonçalves, D. (2014). Developing a Multimodal Interface for the Elderly. *Procedia Computer Science, 27*, 359-368. doi:http://dx.doi.org/10.1016/j.procs.2014.02.040
Rømen, D., & Svanæs, D. (2012). Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society, 11*(4), 375-385. doi:10.1007/s10209-011-0259-3
Schwaber, K. (1997). SCRUM Development Process. In J. Sutherland, C. Casanave, J. Miller, P. Patel, & G. Hollowell (Eds.), *Business Object Design and Implementation: OOPSLA '95 Workshop Proceedings 16 October 1995, Austin, Texas* (pp. 117-134). London: Springer London.
Shaywitz, S. E., & Shaywitz, B. A. (2007). The neurobiology of reading and dyslexia. *ASHA Leader, 12*(12), 20-21.
Shimomura, Y., Hvannberg, E. T., & Hafsteinsson, H. (2010). Accessibility of audio and tactile interfaces for young blind people performing everyday tasks. *Universal Access in the Information Society, 9*(4), 297-310. doi:10.1007/s10209-009-0183-y
Small, J., Schallau, P., Brown, K., & Appleyard, R. (2005). *Web accessibility for people with cognitive disabilities*. Paper presented at the CHI '05 Extended Abstracts on Human Factors in Computing Systems, Portland, OR, USA.
Stevens, R. D., Edwards, A. D. N., & Harling, P. A. (1997). Access to Mathematics for Visually Disabled Students Through Multimodal Interaction. *Human-Computer Interaction, 12*(1/2), 47.
Teixeira, A., Ferreira, F., Almeida, N., Silva, S., Rosa, A. F., Pereira, J. C., & Vieira, D. (2016). Design and development of Medication Assistant: older adults centred design to go beyond simple medication reminders. *Universal Access in the Information Society*, 1-16. doi:10.1007/s10209-016-0487-7
Trewin, S., Swart, C., & Pettick, D. (2013). *Physical accessibility of touchscreen smartphones*. Paper presented at the Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, Bellevue, Washington.
Van Biljon, J., & Renaud, K. (2016). *Validating mobile phone design guidelines: Focusing on the elderly in a developing country.* Paper presented at the ACM International Conference Proceeding Series.
Vision-Australia. (2016). CCA Download Page. Retrieved from http://www.visionaustralia.org/digital-access-cca
W3C. (2008). Web Content Accessibility Guidelines (WCAG) 2.0. Retrieved from https://www.w3.org/TR/WCAG20/
W3C. (2015). Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile. Retrieved from https://www.w3.org/TR/mobile-accessibility-mapping/

Zaphiris, P., Kurniawan, S., & Ghiawadwala, M. (2006). A systematic approach to the development of research-based web design guidelines for older people. *Universal Access in the Information Society, 6*(1), 59. doi:10.1007/s10209-006-0054-8

18 Appendices

The following appendices are included in the appendices.zip file:

1. Ratings review screenshots
2. Survey questions
3. Survey results
4. User testing consent form
5. User testing “speech” that was read to participants before they started testing
6. User testing task transcript
7. Prototype code

Note that some of the individual folders containing these items provide a .doc file with applicable information when necessary. The notes from the interview and from user testing are not included, because they were taken in shorthand. The denotations were cryptic, and had to be reviewed directly after completion of an activity, to ensure that they would be fully understood in the future. They offer no insight into the process, because they are illegible, and appear similar to ancient hieroglyphics, many of which have yet to be fully understood.