An Accessible, Large Print, Listening and Talking E-book to Support Families Reading Together

Abbas Attarwala

University of Toronto

TAGlab, Computer Science Dept.

Toronto, ON, Canada

attarwal@cs.toronto.edu

Cosmin Munteanu

University of Toronto and

National Research Council Canada

Fredericton, NB, Canada

Cosmin.Munteanu@nrc-cnrc.gc.ca

Ron Baecker

University of Toronto

TAGlab, Computer Science Dept.

Toronto, ON, Canada

ron@taglab.ca

Abstract

Reading is an activity that is not only informative or pleasurable, but can have significant social benefits. Especially in a family setting, it is part of the interaction between children and their parents, it helps create a bond between children and their grandparents, and even bring adults and their older parents closer. However, with families increasingly living or spending time in different locations or managing busy schedules that afford very little time together, the social opportunities enabled by reading are often lost. Furthermore, reading can be a challenge for older adults or for those with impaired eyesight. To address these problems, we are proposing ALLT – an Accessible, Large-Print, Listening and Talking e-book. ALLT is a tablet-based e-reading application that enhances the capabilities of e-book readers through customizable and intelligent accessibility features. It provides support for asynchronous "reading together" by synchronizing the audio recording of one user with the text that is later read by another user. This addresses the needs of a variety of users, from visually impaired adults reading together with a loved one, to children being able to replay an interactive story previously read together with their grandparents. In this demo paper we present ALLT's features and detail how they support asynchronously reading together.
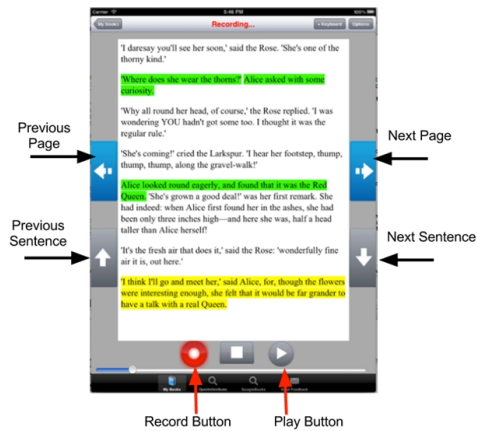
Figure 1: The current sentence is highlighted in yellow. Sentences that have audio recordings are highlighted in green.
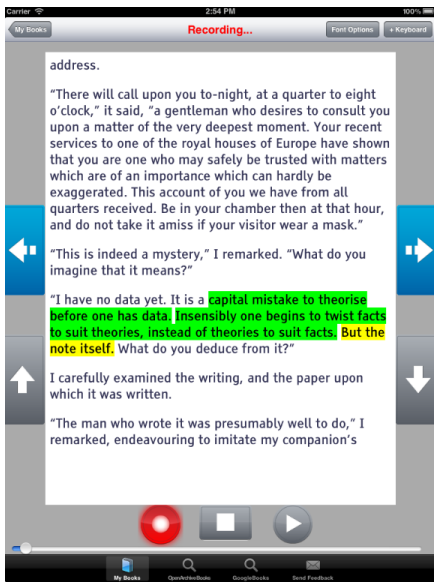

Figure 2: Automatically-synchronized recording in ALLT.

Keywords
Collaborative reading; aging; visual impairment; assistive technology.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces; I.7.4 [Document and Text Processing]: Electronic Publishing

Introduction
Current e-book reading applications and e-readers offer a variety of features that support accessibility or enhance the experience of reading. For example, many applications support text-to-speech – automatically having the text read to the user, or the ability to control the font size. However, such features merely address the accessibility shortcomings of printed books, and not necessarily improve users experience when reading. On a telephone survey that was conducted with 25 blind persons, regarding their use of and satisfaction with reading aids [4], the results indicated that it is tiresome to listen to synthesized speech output for prolonged periods. When looking at the social aspects of reading, one can observe that accessibility is part of reading together; for example, an adult reading to their older parent that suffers from age-related impaired eyesight. Audio books are a replacement for printed books that address the issue of poor vision; however, they are expensive to produce (or rely on volunteers such as the Librivox project [5]) and lack the social interaction aspect of reading together – surveys such as [2] found that a third of partially sighted (and over half of blind) individuals read in this way. This was the inspiration for ALLT ([1] and [7]), which goes beyond enabling access to a read-aloud book, and allows a collaborative reading experience in which people read together, with the recorded audio synchronized with the book content and later played back (in a familiar voice) by the visually-impaired user.

Reading Together in ALLT
Typically, reading together is a synchronous activity, where usually one reader follows the text and reads aloud to his or her companion. This lacks persistence, and the companion is thus not able to relieve the experience at a later time. Of course, the reader can explicitly record the interaction; however, this is not synchronized with the text. Some interactive books such as children's stories have the ability to record a parent reading the text, but in a limited and very direct manner (e.g. by activating an explicit record button for a short paragraph). Since ALLT is aimed at a wider-range of users, including adults reading together, it provides support for asynchronous reading of longer texts, such as novels, which are rarely read in their entirety in a single session.

The audio recording takes place while one user reads aloud to the other users. The recording can then be played back at a later time by the older user, with highlighted text being synchronized with the audio recording to assist those who are partially-sighted. A novel feature proposed in ALLT is a seamless recording, in which the audio recording and text are automatically synchronized through a combination of speech recognition and implicit user gestures. Alternatively, the main user can explicitly control the audio recording, as described below.

User-controlled Audio Recording
ALLT supports recording and playback as seen in Figure 1. The user initiates recording starting at the
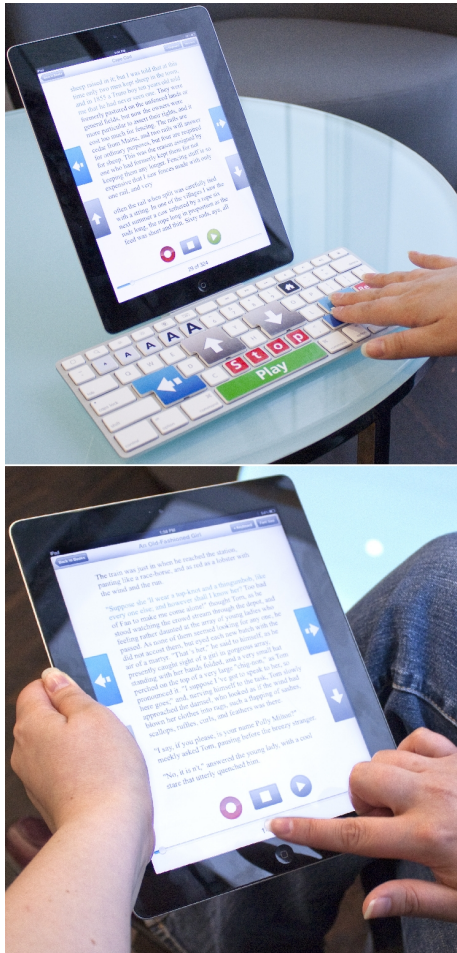
Figure 3: Support for multiple input modalities in ALLT i.e. keyboard (top) and touch based gesture (bottom).

current sentence by clicking on the record button. The Next Sentence and Previous Sentence buttons support sentence navigation through the book. Likewise page navigation is supported via Next Page and Previous Page buttons. A major design goal for the reading aloud environment is that it "stays out of the way" of the reader. This makes it easy for an untrained friend or family member of a partially-sighted individual to easily create an audio-enhanced book.

Users can play back recorded readings made either by themselves or their caregiver to re-live the reading session. To fill in recording gaps, ALLT uses a text-to-speech engine to enable playback of passages that were not previously read.

Seamless Audio Recording

When reading together users should be able to enjoy the experience as much as possible and not have to worry about controlling the recording. For this, ALLT records the audio without users needing to explicitly synchronize the recording with the text being read. The synchronization is carried out in the background through automatic speech recognition using the PocketSphinx system from CMU. The audio recording is transcribed and the resulting text is aligned at word-level with the original audio track.

As the forced alignment of automatically-transcribed text is subject to speech recognition errors, additional sources of information about when the user pauses or at what speed he/she reads can be useful in adjusting the time alignment between the audio channel and the text. For this, we propose a novel feature in ALLT (Figure 2) that implements the metaphor of following

the text with the finger scrolling under the words being read, as many readers do with printed books.

The finger scrolling metaphor is used to determine the users' reading speed and long pauses by following the finger's touches and motions across the screen as the reader follows the text. The current sentence is highlighted on the screen, and moves along as the users drags the finger on the screen. Throughout this process, the reading speed is measured subject to the user's finger movement. When the user lifts the finger off the screen, the recording, together with automatic text highlighting and behind-the-scene synchronization proceeds at the previously-computed averaged speed. The user can touch the screen again and move the finger faster or slower in order to adjust the synchronization speed. Various heuristics are also implemented to prevent false measurements when the users reaches the end of the line and moves the finger back to the other side, or when he/she accidentally moves the finger in a vertical or diagonal direction.

Accessible e-reading in ALLT
Accessibility in ALLT

ALLT supports multiple input modalities. Alongside standard touch screen gestures, the ALLT book can be controlled using a wired or wireless keyboard with large target buttons as seen in Figure 3. We conducted a pilot study with a 40-year old highly educated participant who has multiple sclerosis and lives with active family members and caregivers who can read to her. Due to tremors caused by her affliction, she was unable to hold a traditional book or tablet device. She used ALLT with a stand and the accessible keyboard to read for pleasure and to read together with family members. Since she had not been reading for several years, she
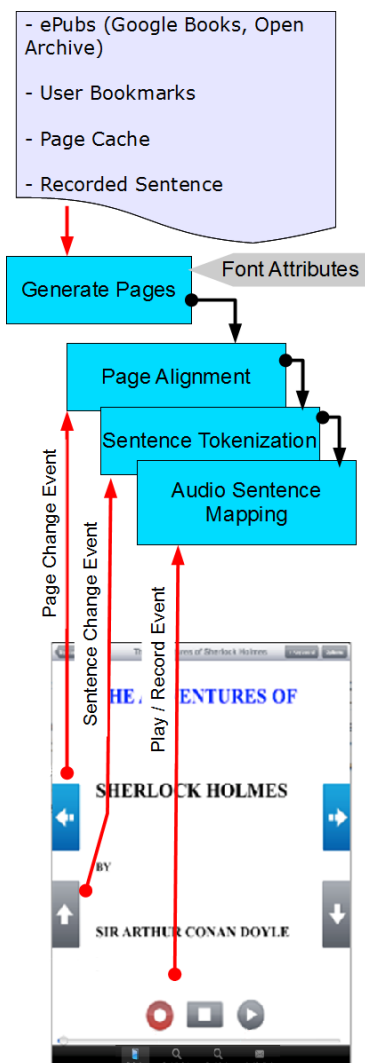
- ePubs (Google Books, Open Archive)
- User Bookmarks
- Page Cache
- Recorded Sentence

Font Attributes

Generate Pages

Page Alignment

Sentence Tokenization

Audio Sentence Mapping

Page Change Event

Sentence Change Event

Play / Record Event

SHERLOCK HOLMES

THE ADVENTURES OF

BY

SIR ARTHUR CONAN DOYLE

Figure 4: Architecture of ALLT.

also found it helpful to read aloud herself and listen to a recording of her own reading.

Large Print in ALLT
ALLT supports different type sizes, and allows the user to select various attributes of text: typeface, text color and background color, and color contrast. These can be adjusted for users with different visual acuities.

Implementation of the Accessible Features
A high-level architecture diagram of ALLT is represented in Figure 4. ALLT is developed on the Model-View-Controller design pattern. The controller is responsible for generating and aligning pages, performing sentence tokenization, and mapping user-recorded audio (if any) with sentences. The page generation receives as input: font attributes selected by the user and the e-book (epub format). Font attributes consists of user selected typeface, font size, font color, and contrast. The mapping of recorded audio to selected sentence is dependent on the record event. Likewise, the playback of recorded audio or text-to-speech for the selected sentence is a function of the play event. If text has been recorded, this is read aloud; otherwise text-to speech is used.

Future Work
To better understand how people read together, and to inform future versions of ALLT, we conducted a study of pairs of people who currently read together with current e-book technologies [6]. We noticed that participants value audio playback of e-books, both when listening alone or together. Reading aloud with friends and family can be a way to pass the time and enjoy the company of others. Human voices add

emotion and depth to the playback, having both utilitarian and entertainment values.

It is important to support reading together both synchronously and asynchronously. Drawing from the design implications listed in our previous study [6], an exploratory study is currently being planned which will explore how ALLT can improve on partnered reading experiences in collocated settings, both when reading for pleasure and for practical purposes (e.g. reading information brochures, with the help of a caregiver or social worker). We also plan to investigate more formally some of the novel features such as the finger scrolling metaphor as well as the accessibility features.

References
[1] A.Attarwala, R.Baecker, and C.Munteanu. Accessible, Large-Print, Listening & Talking E-book (ALLT). In CIKM BooksOnline, 2012.

[2] G. Douglas, C. Corcoran, and S. Pavey. Network 1000. Opinions and circumstances of visually impaired people in Great Britain: Report based on over 1000 interviews. University of Birmingham, 2006. ISBN: 0704426048 / 9780704426047.

[3] M. F. Theofanos and J. Redish. Bridging the gap: between accessibility and usability. J. interactions, 10(6):36–51, 2003.

[4] R. Steele RD et al. Reading aid technology for blind persons: Responses to a questionnaire of experienced users. J. Assistive Technology :23-30, 1989.

[5] Librivox. http://librivox.org/

[6] M. Massimi, et al. Reading Together as a Leisure Activity: Implications for e-Reading. In INTERACT 2013.

[7] W.X. Snelgrove and R.M. Baecker.2010. A system for the collaborative reading of digital books with the partially sighted. In CIKM BooksOnline, 2010.