Interaction for Reading Comprehension on Mobile Devices

Rafael Veras*, Erik Paluka*, Meng-Wei Chang*, Vivian Tsang†, Fraser Shein†, Christopher Collins*

*University of Ontario Institute of Technology
Oshawa, Canada
{rafael.verasguimaraes, erik.paluka, meng-wei.chang, christopher.collins}@uoit.ca

†Quillsoft Ltd.
Toronto, Canada
{vtsang, fshein}@quillsoft.ca

ABSTRACT

This paper introduces a touch-based reading interface for tablets designed to support vocabulary acquisition, text comprehension, and reduction of reading anxiety. Touch interaction is leveraged to allow direct replacement of words with synonyms, easy access to word definitions and seamless dialogue with a personalized model of the reader's vocabulary. We discuss how fluid interaction and direct manipulation coupled with natural language processing can help address the reading needs of audiences such as school-age children and English as Second Language learners.

Author Keywords

e-reading; mobile; touch interaction

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

INTRODUCTION

Reading in the mobile HCI landscape is largely dominated by applications designed for proficient readers, featuring a variety of supports for active reading tasks, in particular, annotation. The needs of a large audience, composed of children, English as Second Language (ESL) learners, and struggling readers have not been well-addressed. Such needs consist mainly of support for vocabulary acquisition and text comprehension.

School-age children are faced each year with a massive number of new English words. Printed school English for grades three through nine contains about 88,500 word families, with over 100,000 distinct meanings, and the average number of words learnt yearly by children until the end of high school, when they know on average 40,000 words, is around 3,000 [13].

Research has shown that much of the vocabulary learning occurs incidentally while reading, with strategies that include using contextual cues, morphological information, cognate knowledge, and dictionaries [3]. The consensus is that the best way to improve vocabulary is by reading difficult texts, but texts with a high density of unknown words can quickly become unintelligible and frustrating [4, 10]. Moreover, there is evidence in the literature showing that ESL learners experience reading anxiety, the anxiety levels correlate with the reader's perception of the difficulty of the text, and the reading performance of learners decreases in conjunction with increased reading anxiety [16, 17].

In this paper, we present a novel reading interface for tablet devices aimed at improving the experience of struggling readers. Our design focuses on fluid interaction with the goal of empowering readers and relieving the anxiety that comes with reading unfamiliar texts. We take touch interaction beyond marking and highlighting to direct manipulation of the text's content. The main features of our prototype are a visual representation of text difficulty, interactive replacement of words with synonyms, a word definition lens, and a personalized model of the reader's vocabulary.

RELATED WORK

Several publications present multitouch tools for active reading, with some interaction design similarities to this work. LiquidText is a desktop application that leverages multitouch interaction to create a fluid reading experience where the main design goal is flexibility [18]. It supports annotation, highlighting and novel ways to manipulate the representation of content. Similarly, there are tools that explore pen and touch interaction for active reading on tabletops [12] and tablets [9]. Although these tools aim to develop more fluid ways to interact with text, they focus on active reading tasks, mainly annotation; therefore, they lack support for text comprehension and vocabulary acquisition.

Other publications propose user interfaces or methods to make texts easier to understand. For example, Zeng-Treitler et al. replace difficult words in electronic health records (EHR) with easier synonyms to improve comprehensibility for customers [21]. Unlike our prototype, their system is non-interactive and its scope is limited to improving sentences for EHR systems. Yu and Miller propose a method for transformation of content on web pages centred on sentence separation and spacing to enhance readability and comprehension for non-native readers [20]. Eom et al. offer sense-specific word definitions for texts on the web through a rudimentary HTML interface; their experiments demonstrated that the system can help ESL learners enhance vocabulary and understanding of texts [7].

MobileHCI 2014, September 23–26, 2014, Toronto, ON, Canada.

ACM 978-1-4503-3004-6/14/09 ...$15.00.
http://dx.doi.org/10.1145/2628363.2628387

Microsoft Word 2013 has functionalities that can be used to improve comprehension of texts, although this feature is targeted at writing, instead of reading. Readers can access dictionary definitions and synonyms of any selected word. Through the "touch mode" in hybrid laptops, these functions can be accessed by tapping on the screen, creating an experience somewhat similar to the one presented in this paper. The fundamental difference is that Word is still a desktop application based on the WIMP paradigm. The dictionary and thesaurus are add-ons that perform in isolation from the overall context of the word of interest. For example, the dictionary presents multiple definitions, some of them not corresponding to the word's part-of-speech (POS) in the text.

In the commercial mobile landscape, a large number of educational applications are meant to improve reading comprehension and vocabulary. However, the great majority is composed of games for kindergarten-aged children aimed at teaching words in isolation (reproducing the flashcard experience) [11], distinguishing facts from opinions [5], identifying the five W's of reading [6], or reading out loud curated stories [15]. Our application creates a mobile reading experience for older children, aimed at vocabulary acquisition and easing reading anxiety. It can be used with any text chosen by the learner or suggested by a teacher.

In summary, there is a gap left unaddressed between out-of-context, kindergarten word learning and independent reading. In other words, there is yet a system that provides the appropriate scaffolding leading to sustained, independent and critical reading, for which some of the cited systems provided support. Our proposed system provides one such scaffold. To the best of our knowledge, no publication or commercial system has brought together solutions for vocabulary acquisition and text comprehension in a highly interactive, general purpose, mobile reading system.

PROTOTYPE

We developed a multitouch web reading interface tailored to tablets. The prototype allows the creation of a personal profile and accepts as input plain text documents. In this section we introduce the main features of the prototype.

Breaking Down the Wall of Grey

When faced with a page of dense text, colloquially called a "wall of grey," readers can become intimidated and discouraged before even beginning to read. One of the authors of this work relates an anecdote in which his child was intimidated by a reading challenge and refused to begin to read a new book. The parent discovered that by pointing out how many words his child already knew on the page, the reading anxiety was reduced. This observation inspired our first design feature, the Personal Vocabulary View (PVV) which "breaks down" the wall of grey text by showing the reader how much of the text they already know how to read, based on a record of their personal vocabulary. The technique also isolates the proportion and location of unknown words, which leads the reader to use the other tools, described later, to expand their vocabulary. An interaction mechanism allows readers to inform the application of additional words they know. With a personalized, incremental model of the reader's vocabulary knowledge, our Personal Vocabulary View conveys a personalized visual representation of the text difficulty.

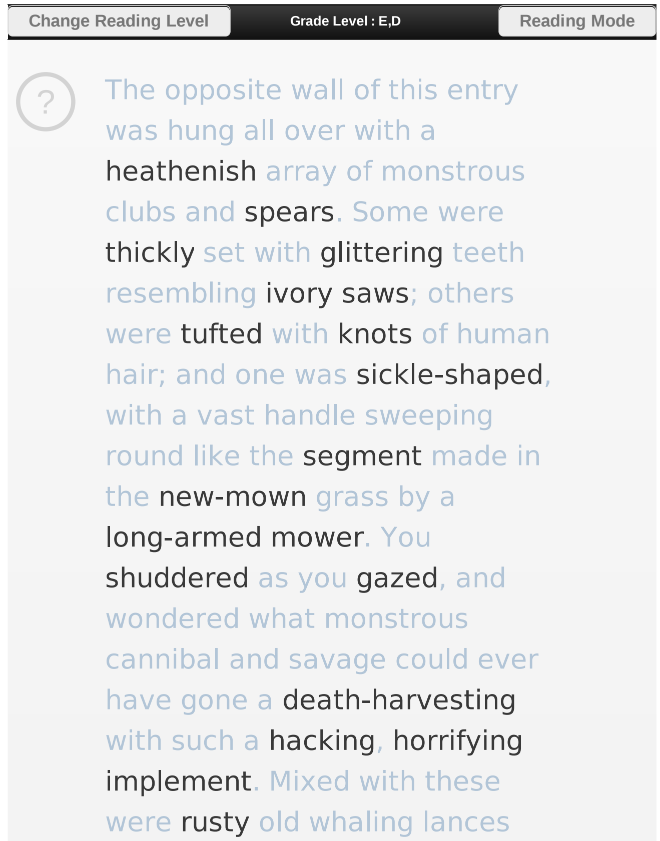

Figure 1: The Personal Vocabulary View highlighting unknown words according to a personalized model of the reader's vocabulary. Excerpt of Moby Dick by Herman Melville.

The PVV (Figure 1) fades out the words that are known by the reader, while the unknown words are highlighted in black. Every word in the text can be toggled with a single tap gesture. Tapping an unknown word informs the tool's internal model that the word (with the current part-of-speech) is known by the reader, preventing it from appearing as unknown in the future. If the word appears in other context playing a different syntactical role, it will be considered unknown. Document scrolling is constrained to dragging in the margins of the text to prevent gesture ambiguity.

One of the challenges related to this approach is inferring the initial vocabulary of the reader. Biemiller, using partly empirical methods, roughly estimated the vocabulary of pre-elementary and elementary students in a list of 11,000 root word meanings categorized as: useful to teach during the primary grades (K–2), easy by second grade, useful to teach during the upper elementary grades, and difficult for elementary children at grade 6 or earlier [1]. We use that list, along with a list of stop words, to populate the tool's vocabulary model of the reader at profile creation time, based on the reader's

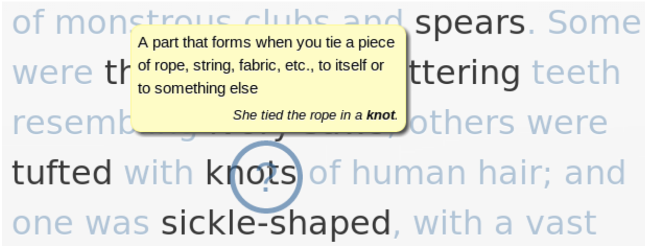

Figure 2: The word definition lens being used on the word *knots*.

selection of school level: a) lower than grade 2, b) lower than grade 6 or c) higher or at grade 6. Note that this level system is most appropriate for native readers, as it refers to school grades. We consider alternative methods for inititalizing the personal vocabulary list in the future work section.

Interactive Replacement of Words with Synonyms

The data-driven learning paradigm advocates that systems should provide evidence (as opposed to answers) and rely on users' intelligence to reach the correct conclusions. For example, researchers have shown that corpus-based usage examples, based on concordances, can be just as effective as dictionary definitions in conveying the meaning of words for L2 learners [8]. In favour of a minimal switch in the focus of attention, we take the inverse approach: instead of showing different contexts where the target word appears, we allow readers to browse synonyms suitable to the context. Evidence for the efficacy of synonyms for text comprehension can be found in Wittrock et al., where the authors demonstrate that stories whose rare words were replaced by more frequent synonyms were significantly better comprehended by children in sixth grade [19]. We hypothesize the replacement of words with (more familiar) synonyms can make the understanding of rare words easier and faster than dictionary definitions. Moreover, through exploring an array of nearly equivalent words in a real usage context, readers have more opportunities to improve their vocabulary.

In our prototype, readers can manipulate the words of a document directly using their fingers. Building upon a roller metaphor, as in a combination lock, the application allows readers to "roll" over several choices of synonyms for any word in the document, as illustrated in Figure 3. The rolling gesture is registered by touching and holding a word for at least 600ms, causing the word to subtly shake, as if it were "detached" from the document. Then readers can drag their fingers up and down to roll over synonyms. Lifting the finger terminates the gesture and replaces the original word with the synonym chosen. To relieve the constraint of having fingers down during the whole gesture (there may be many synonyms), we delay the termination by 2 seconds after the finger is lifted; any touch on the word within the delay is interpreted as a continuation of rolling, allowing one to perform multiple swipes, as opposed to a continuous dragging motion. The substitute words remain underlined to facilitate their identification. Compared with dictionary definitions, which usually appear in pop-ups and carry longer text with their own unfamiliar words [14], we believe this design is more fluid and less distracting.

Note that we intentionally did not include a visual affordance indicating which words have synonyms available. As the view is both an exploratory interface as well as a reading application, we felt this would be distracting. Also, through engaging with the text in an exploratory way, we hypothesize that learners will develop a deeper connection to the vocabulary.

Ranking Synonyms

The order in which the list of synonyms is made available to the reader is important to minimize the time spent browsing, and it can implicitly inform the user of the most appropriate synonyms, since the semantic equivalence of words varies across contexts. Also, the knowledge about the reader's vocabulary captured by the PVV can be leveraged to prioritize more familiar words, as the utmost purpose of the tool is to make the text more comprehensible.

Thus, the ranking of synonyms is based on two attributes, in order of importance: familiarity and fitness to context. Upon registration of the rolling gesture, a list S of synonyms of the target word with matching part-of-speech is retrieved from the Merriam-Webster's Dictionary Public API and sorted in two stages according to the ranking criteria. The sorting determines the order the words appear while the reader is performing the rolling gesture.

The *familiarity score* of each $s \in S$, where s is a tuple of the form $\{word, POS\}$, is determined straightforwardly based on its membership in the set of the reader's known words (0 or 1). The *fitness score* is computed according to the frequency of s in a large N-gram corpus, within the context of the next word w_1. Therefore, the fitness score $F(s, w_1)$ is defined as $C(s, w_1)/C(*, w_1)$, where the numerator corresponds to the frequency that s occurs before w_1, and the denominator corresponds to the frequency that w_1 occurs after any word. As a source for the frequencies, we employ the bigram collection of the Contemporary Corpus of American English, containing over 48 million rows with POS information.

After sorting, S is further trimmed by removing the words for which $F = 0$. This ensures that synonyms that never occur in a certain context are cut off, preventing sentences such as "a strong cup of tea" being erroneously turned into "a powerful cup of tea". Finally, we change the inflection of the synonym, if necessary, to match the same conjugation of the original word, in case of a verb, or the same number, in case of a noun.

Word Definition Lens

The replacement of words with synonyms does not always lead to comprehension, as sometimes the replacement is just as unfamiliar as the original word; in addition, the algorithm described in the previous section might return an empty list, in case no appropriate replacement is found. In order to cover these cases, we designed a minimalistic, gesture-based word definition lens. Many of the available tools that offer word definitions, such as the built-in iPad dictionary, present the user with an overwhelming amount of information, usually containing definitions for several senses of the word, along with phonetic symbols and usage examples. The goal behind our word definition lens was to make it simpler and more interactive.

The lens is represented by a question mark icon placed in the upper left corner of the screen (Figure 1). The lens can be dragged around freely; holding it over any word triggers the appearance of a box showing the first, main definition of the word along with a usage example from Webster's Learner's Dictionary (Figure 2). The word definitions presented correspond to the POS of the target word (distinguishing, for example "sit in a chair" from "chair a meeting"). Swiping the definition box brings up additional definitions of the word, if they exist. Any touch out of the definition box cancels it. The lens returns to its home position when released, allowing reading to resume quickly. The word definition lens allows readers to investigate the meanings of many words quickly by just dragging the lens over them.

Implementation Details

Our prototype is implemented in a client-server architecture. The client-side is based on HTML5, CSS3, and JavaScript, powered by the D3 library. The server-side is based on Python, and uses NLTK and nodebox for the natural language processing. Communication is done via JSON documents.

When the reader inputs the text in the client, the text is sent to the server for processing and returns with POS tags. Every interaction in the client requires the target word to be sent along with surrounding words and POS tags to the server. The personal vocabulary of each reader is held on the server side. Latency depends on requests to the third-party API (Webster's) where it gathers synonyms and definitions from. The server maintains the N-gram corpus in a fully indexed database, so that lookup is done without compromise to interaction. POS tagging is done following the Brill algorithm [2]. All interactions with the text (tapping words, activating synonym replacement, and requests for definitions) are tracked in order to build a data resource about reading behavior and vocabulary development on an individual and community basis. Leveraging this data for educational analytics, and investigating the privacy issues related to its collection are left for future work.

FUTURE WORK

Future work includes a study to evaluate our tool using instruments for measuring reading performance and anxiety. There are also many opportunities to expand the capabilities of our prototype. The current estimation of the reader's initial vocabulary does not support groups other than English first language children; estimates of ESL learners' vocabulary, for example, are necessary. Ideas to initialize the model include creating a self-adjusting model of learner vocabulary based on community statistics. At first deployment of our software, the vocabulary initialization would be based upon rough estimates and models taken from the grade-level vocabulary dataset. Then, leveraging the collective knowledge about vocabulary encoded in the usage statistics from the system, and

Figure 3: Direct replacement of words with synonyms. The word "heathenish" is replaced with the more familiar word "wild" through a rolling gesture, captured here in three snapshots. The grey circle represents the touch point.

meta-data about the learner, we could create an increasingly accurate estimate of initial learner vocabulary for new users. Another potential avenue for vocabulary estimation could be through analyzing some writing or reading samples provided by the learner. When considering how this prototype could fit into a larger learning ecology, the data collected by the interactions of readers with the software, if properly protected and anonymized, could be useful for educators to mine in curriculum planning.

The suggestions of synonyms and dictionary definitions are currently based on part-of-speech and bigram frequency of use statistics. Future work will investigate how grammar (parsing) and word sense disambiguation (semantics) can improve the quality of the suggestions. For example, rather than looking only at bigrams of the candidate word and next word to filter synonym candidates, we could examine counts of the candidate word and syntactically related verbs or nouns using the parse tree.

Currently, a single tap on a word in the PVV is taken as a permanent indication that a word is known by the reader. This design decision needs to be tested with an empirical study. It may take multiple exposures to a new term before a reader is comfortable with it. We are investigating ways to adapt the number of exposures required before the word is marked as known. One idea is to require multiple exposures (and taps) on words that are ranked as significantly above the reading level of the learner, as determined by community statistics. In addition, it may be useful to employ a time function, requiring multiple exposures of a term in a specified time period before it is permanently marked as learned, or a morphology function, requiring exposure to multiple variants of a word in a variety of contexts.

The PVV currently only shows the difficulty of the words that are within the viewport. A more abstract representation that can portray the difficulty of the words across the entire text could make this view more powerful and should be considered.

CONCLUSION

With the design proposed in the current paper, we hope to have empowered struggling readers with tools to effectively overcome reading anxiety and comprehension difficulty using mobile technology. Recognizing that the main obstacle for reading is vocabulary difficulty, we designed a view where readers can easily visualize the unknown words in a text, an interaction mechanism by which readers can substitute difficult words with familiar synonyms and a minimalistic word definition lens. Furthermore, we believe the direct manipulation and the fluidity of our interaction play a significant role in connecting the reader to the text, which might also be a factor in reducing anxiety.

REFERENCES

1. Biemiller, A. *Words Worth Teaching: Closing the Vocabulary Gap*. McGraw-Hill SRA, 2009.
2. Brill, E. A simple rule-based part of speech tagger. In *Proc. of the Workshop on Speech and Natural Language*, HLT '91, Association for Computational Linguistics (1992), 112–116.
3. Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. J., and White, C. E. Closing the gap: Addressing the vocabulary needs of english-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly 39*, 2 (2004), 188–215.
4. Carver, R. P. Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Literacy Research 26*, 4 (1994), 413–437.
5. E. Skills Learning, LLC. MiniMod Fact or Opinion Lite. `https://itunes.apple.com/ca/app/minimod-fact-or-opinion-lite/id492491726`, 2013.
6. E. Skills Learning, LLC. MiniMod Reading for Details Lite. `https://itunes.apple.com/ca/app/minimod-reading-for-details/id409669762`, 2013.
7. Eom, S., Dickinson, M., and Sachs, R. Sense-specific lexical information for reading assistance. In *Proc. of the Seventh Workshop on Building Educational Applications Using NLP*, Association for Computational Linguistics (2012), 316–325.
8. Frankenberg-Garcia, A. The use of corpus examples for language comprehension and production. *ReCALL 26*, 02 (2014), 128–146.
9. Hinckley, K., Bi, X., Pahud, M., and Buxton, B. Informal information gathering techniques for active reading. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, ACM (2012), 1893–1896.
10. Hsueh-Chao, M. H., and Nation, P. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language 13*, 1 (2000), 403–30.
11. Innovative Mobile Apps. Sight Words List. `https://itunes.apple.com/ca/app/sight-words-list-learn-to/id445708245`, 2013.
12. Matulic, F., and Norrie, M. C. Supporting active reading on pen and touch-operated tabletops. In *Proc. of the International Working Conference on Advanced Visual Interfaces*, AVI '12, ACM (2012), 612–619.
13. Nagy, W., and Herman, P. Limitations of vocabulary instruction (tech. rep. no. 326). *Center for the Study of Reading. Urbana: University of Illinois* (1984).
14. Nesi, H., and Haill, R. A study of dictionary use by international students at a british university. *International Journal of Lexicography 15*, 4 (2002), 277–305.
15. Rock 'N Learn, Inc. Rock 'n Learn Phonics Easy Reader. `http://rocknlearn.com/html/EasyReadersiPad.htm`, 2011.
16. Saito, Y., Garza, T. J., and Horwitz, E. K. Foreign language reading anxiety. *The Modern Language Journal 83*, 2 (1999), 202–218.
17. Sellers, V. D. Anxiety and reading comprehension in spanish as a foreign language. *Foreign Language Annals 33*, 5 (2000), 512–520.
18. Tashman, C. S., and Edwards, W. K. Liquidtext: A flexible, multitouch environment to support active reading. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11 (2011), 3285–3294.
19. Wittrock, M. C., Marks, C., and Doctorow, M. Reading as a generative process. *Journal of Educational Psychology 67*, 4 (1975), 484.
20. Yu, C.-H., and Miller, R. C. Enhancing web page readability for non-native readers. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, ACM (2010), 2523–2532.
21. Zeng-Treitler, Q., Goryachev, S., Kim, H., Keselman, A., and Rosendale, D. Making texts in electronic health records comprehensible to consumers: A prototype translator. In *Proc. of the AMIA Annual Symposium*, vol. 2007, American Medical Informatics Association (2007), 846.