Why Read if You Can Skim: Towards Enabling Faster Screen Reading

Faisal Ahmed Yevgen Borodin Yury Puzis I.V. Ramakrishnan

Stony Brook University, Computer Science Department, Stony Brook, NY 11790-4400
{faiahmed, borodin, ypuzis, ram}@cs.stonybrook.edu

ABSTRACT

Skimming broadly refers to different speed-reading methods that aim to enhance the rate of reading without unduly compromising on comprehension and retention of information. Skimming of content could be particularly useful for people with vision impairments, who frequently experience information overload when listening to reams of digital content online. Support for usable and useful skimming in modern screen readers remains very poor. This paper explores the user requirements for a usable non-visual skimming interface, informed by a large-scale human-subject experiment with blind individuals. Specifically, the study has: (1) helped identify the type of skimming that can be useful in screen reading main content in web pages; (2) led to the development of a usable interface for accessible online skimming; (3) demonstrated the utility of the accessible skimming interface in two realistic use scenarios; (4) identified automatic summarization techniques that could "closely" approximate skimming methods used by sighted people.

Categories and Subject Descriptors

H.5.2 Information Interfaces and Presentation]: User Interfaces; H.5.4 Information Interfaces and Presentation]: Hypertext/Hypermedia –navigation; H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors, Experimentation, Design

Keywords

Skimming, Summarization, Web Accessibility, Navigation, Speed Reading, Blind Users, Non-Visual, Web Browser, Audio Interface

1 INTRODUCTION

Since the invention of computers, researchers and practitioners have been trying to make computers more accessible to people with vision impairments. The introduction of screen readers has given blind people the ability to access web content. Nevertheless, a large gap remains between the ways blind and sighted people interact with web pages.

Sighted individuals can quickly process web content with their eyes; in contrast, blind people have to use screen readers [19, 25, 31] – assistive technology software that can narrate web page content. A screen reader typically reads *all* of the content, while allowing users to navigate in it. Although screen readers enable blind people to access the Web, screen-reader users often cannot determine whether the content in web pages is worth listening to, unless they hear at least some of it. As a result, blind users often suffer from information overload.

Sighted people who are used to reading large volumes of information often employ *skimming* [22, 24], a speed-reading technique that allows them to get the general idea of the main content in web pages and sift through information very quickly, without reading the entire content and, most importantly, without sacrificing basic understanding (a detailed overview of skimming techniques is provided in the Background Section). While screen-reader users can employ tools such as Apple's Reader or Instapaper to find main content in web pages, they cannot process information in the same way and as fast as sighted people can [4]. And this is where assistive technology should come to rescue.

JAWS screen reader [19] offers a feature called "skimming" [20], which allows users to read the 1st line or the 1st sentence of each paragraph. It also lets users specify regular expressions and then listen to the content that matches those expressions. Although with this feature, users can get some idea of the content, they miss most of the information and often have to reread the same content over again. Twenty experienced JAWS users we interviewed either did not know about the JAWS skimming feature or considered it inconvenient and/or "useless." Furthermore, in our user study with the same users, we observed that they use their own simple ad hoc "skimming" techniques, which we overview in the Ad-Hoc Techniques Section. Therefore, there is a clear need for an assistive technology that can facilitate effective skimming for screen-reader users. Before attempting to design an algorithm to support skimming, in this paper, we aim to identify the qualities of a usable skimming interface and evaluate it with end-users.

To design a skimming interface that would be useful for screen-reader users, we reviewed the skimming techniques employed by sighted people, identified the specifications of several possible skimming interfaces, and evaluated them in realistic use scenarios. In interviews with screen-reader users, we determined that, for faster reading, they wanted to have access to a shortened summary that would, however, preserve most of the original content. To obtain such summaries, we conducted an experiment with 12 sighted people who shortened 6 web articles and generated several variations of gold-standard summaries for skimming. We then developed a screen-reading interface that allowed blind people to switch seamlessly between an article and its summary. Finally, we conducted an extensive user study with 20 blind subjects to evaluate several types of skimming in listening-and-comprehension and in search scenarios that we overview later.

W4A2012 - Technical, April 16-17, 2012, Lyon, France. Co-Located with the 21st International World Wide Web Conference. Copyright 2012 ACM 978-1-4503-1019-2...$5.00.

The results of our experiments helped us compare several variations of skimming summaries, confirmed the general effectiveness and usability of our skimming interface for main content in web pages, and received highly positive feedback from the end-users who participated in the study. Statistical significance tests provided strong support of our two primary hypotheses: A: *"skimming is significantly faster than regular screen reading"* and B: "*user's comprehension does not fall off with skimming compared to regular screen reading.*"

To summarize, the 4 major contributions of this paper are:

1. Identifying the type of skimming that is perceived as "useful" and "efficient" by screen-reader users;
2. Proposing a usable interface for skimming;
3. Demonstrating high utility and effectiveness of the skimming interface in realistic use scenarios;
4. Providing insight into how automated summarization can be developed to support the skimming interface.

Although we realize that the proposed accessible skimming interface could be used with any digital content, in this paper we focus primarily on skimming main content in online articles.

2 BACKGROUND

In this section, we review the skimming techniques employed by sighted people, assess the skimming approaches available to screen-reader users, relate skimming and summarization, and, finally, establish the need for a more usable skimming approach.

2.1 Skimming (a.k.a. Scanning)

People typically read for enjoyment, to search for information, to complete a task, or to explore / review text [22]. Whatever the specific purpose of reading may be, in our information-driven society, people are faced with large volumes of information that they need to process. Cognitive overload is a fundamental problem that arises while reading volumes of text. The problem gets worse especially when the information context changes frequently across the content [6], which is especially true when browsing websites. Hence, the ability to sift through information can be very useful for reducing the cognitive load.

To overcome the cognitive overload, sighted people employ a number of different speed-reading techniques (a.k.a. skimming or scanning). We will treat "skimming" and "scanning" as synonyms. Usually, the technique is chosen depending on the purpose of reading. Ron Kurtus has suggested a couple of methods that students can use to read faster, which include reading first sentence, skimming several times, etc. [21].

The following are the most frequent skimming techniques employed by sighted people. One of the simpler techniques is to look through the illustrations, titles, subtitles, and other visually prominent content. However, this only gives an idea of what the content is about. To gain a deeper understanding of the content one can read the first and/or the last sentence in each paragraph. This is the technique that JAWS [19] screen reader is trying to imitate with its skimming feature [20], as we describe in the next section. However, to gain understanding of the content, one has to scan the entire content; more advanced sighted readers can scan through the content catching separate phrases instead of individual words. So without reading every word, advanced readers can get most of the information slowing down or accelerating depending on their information needs. This can often be accomplished by diagonal reading, when the reader moves the eyes diagonally over the content picking out the most important phrases [29].

2.2 Skimming in Accessibility

During our study, a number of blind subjects said they wished they could skim web pages. Unlike sighted users, blind people have to use screen readers to access web content. Screen readers have come a long way from simple screen reading to intelligent content analysis that can segment web pages, categorize web objects by type, and offer the layered interface to navigate them [5, 7, 30]. However, screen readers still do not offer adequate support of skimming. So far, other than JAWS, no screen reader claims to support skimming. Also, Apple's VoiceOver [30] has several features that are relevant to skimming.

JAWS screen reader [19] has a skimming feature [20] which allows users to read the 1st sentence or a line of each paragraph. It also lets users create their own rules (regular expressions) for skimming. The rules tell JAWS to read certain patterns in the article, but most rules, other than those relating to text formatting, cannot be applied to different content. In our interviews, screen-reader users said that JAWS skimming "was not useful," that JAWS skimming "was not enough to look into the content of an article," and that it was similar to the paragraph navigation feature, but used for a different purpose and with a different name.

VoiceOver (VO) screen reader [30] has a user friendly accessible interface for interacting with the screen content. VO has a number of features that enable users to read faster, however, none of them can be called skimming. VO allows users to drag a finger over the screen and listen to the content under the finger. In [1] Ahmed et al. mentioned granularity issues in VO's "dragging interface" and pointed out that there was no way to ensure that content was read sequentially. While this feature could, in theory, be used for skimming, currently VO simply reads the content under the finger without trying to determine how important that content is. VoiceOver, just like Hearsay [5], provides a layered view on web content allowing users to access elements of a particular type (e.g., headings, links, images, etc.). While this feature could give an overview of the content, it cannot reproduce the experience of glancing over an article and picking the most salient phrases.

Based on our interviews with screen-reader users, none of the assistive technology tools has adequate support of skimming. As a result, screen-reader users try to use regular screen-reader navigation shortcuts to simulate skimming in web pages, as was confirmed during our user study. Unfortunately, as we show in our experiments, such ad hoc skimming (Section 4) does not prove to be particularly effective.

2.3 Summarization

Summarization is a process of condensing textual content into a shorter version, to help the audience get the gist in a short period of time. Summarization is supposed to highlight important points of the original long content (e.g., article, paper, book, etc.) to give the user the gist of the content, but summarization frequently does not preserve the mood or style of the original text. Automated summarization techniques aim to summarize text with the help of computer algorithms [28].

Several attempts have been made to use summarization in assistive technology tools. Harper et al. introduced Summate [15] – a summarization tool for blind individuals. Summate is a FireFox-based tool that summarizes web pages and presents the summary in an alert box. Another accessibility tool, AcceSS [27] attempted to simplify and summarize web contents for web users. AcceSS was doing simplification by removing the clutter and retaining the important sections to give the user a preview of the page. Oxford Brookes University have built a Web navigation tool called BrookesTalk [34] that uses information retrieval techniques to summarize a Web page for quick orientation. While these tools gave screen-reader users the ability to preview content, they have not given the blind users the ability to skim and switch between the summary and the original content easily.

Unfortunately, summarization alone does not enable skimming. A summary often does not preserve all of the salient content, because, otherwise, the summary will be too long. Reading a summary, a blind person can get the gist of the content, but cannot control how much information appears in the summary and has no easy way to get more information. Our subjects reported that, after reading regular summaries, they would often have to read the entire content again to look for missing information.

In our research, we are trying to recreate the skimming experience that blind people have not been able to enjoy. We believe that, at its core, skimming is an interface has to rely on summarization.

Broadly, there are two main types of summarization techniques: summarization by extraction and summarization by abstraction. In short, extraction deals with picking out the most salient sentences or phrases from the original text [2, 12, 13], and abstraction is compressing sentences [3, 9], which involves paraphrasing the existing content, for example by automatically filling out a template [18]. Most of the existing summarization techniques [10, 11, 14, 18, 32, 36] use one or both of these two techniques.

Because visual skimming involves "extracting" salient content with one's eyes, the extractive summarization appears to be the most suitable technique to support computerized skimming. Evaluation and comparison of the existing approaches to extractive summarization was not in the scope of this research, instead, we focused on confirming that non-visual skimming was useful for screen-reader users. Because we did not want shortcomings of the existing summarization technique to affect the results, we used human-generated summaries to identify the qualities of the ideal automated skimming approach.

2.4 The Need for Accessible Skimming

Our review of literature, as well as personal experience have shown that sighted people employ numerous speed-reading techniques; sighted readers can scan web content with their eyes and pick out salient sentences and phrases that help the readers get the gist of the content. Assistive technology does not provide adequate support of skimming that would allow blind people to skim with efficiency comparable to that of sighted people.

Summarization by itself does not satisfy the need for skimming; for screen-reader users, reading a summary is more like reading a separate, albeit shorter, narrative. We are not aware of any interfaces that can connect the summary and the original text – the feature that would allow the screen-reader user to switch seamlessly between the two.

Based on our interviews with screen-reader users, extractive summarization techniques that preserve the original content appear to be the best match for enabling the skimming interface. The design of an appropriate summarization approach is the subject of future research that will be informed by the experiments described in this paper.

3 DESIGN & EVALUATION

We have conducted a user study in order to verify the utility of skimming and evaluate our skimming interface. In the first step we prepared the materials for the user study, by asking 12 sighted subjects to create *gold-standard* summaries of several web articles, i.e. manually summarize the articles. In the second step, 20 blind subjects were asked to use our skimming interface to do two realistic tasks (listening and comprehension and searching for information) with and without the skimming interface.

3.1 Generating Gold-Standard Summaries

The quality of automated summarization is gradually approaching that of humans [28, 35]. However, to enable and evaluate the proposed skimming interface without committing to a particular automated summarization technique, we employed human-generated summarization. We will use the terms "gold-standard summary" and "gold summary" to refer to the human-generated summaries throughout the paper. By using sighted subjects to manually produce gold summaries of articles, we were able to guarantee that the results of our experiment would not be affected by the shortcomings of any particular summarization method.

3.1.1 *Participants*

To generate the gold summaries, we recruited 12 sighted participants – all college-educated and fluent in English.

3.1.2 *Setup*

We collected 6 news articles on different topics with 5-6 paragraphs in each of the articles on average. Each of the 12 sighted participants was asked to summarize 2 articles, thus, generating 4 gold-standard summaries for each article.

Since the purpose of skimming was to pick out the most salient information without rephrasing the original content, we stipulated that the subjects use extractive summarization. Since the point of skimming was to save time to the end users by presenting less information, we set the upper bound on the length of the summary to 1/3 of the original content. However, we did not want to dictate how exactly to do summarization, so each participant was provided only with the following "loose" guidelines:

- Each sentence is to be summarized separately;
- The summary may only include words and phrases found in the summarized sentence;
- Words in the summary should be in the same order they appear in the original text;
- The length of the summary should be no more than one third of the length of the original article;
- Summaries should be as informative as possible.

3.1.3 *Analysis*

Sighted human subjects generated 4 gold-standard summaries for each of the 6 articles. Because of the loose guidelines, subjects generated summaries of variable length choosing different words. Specifically, some summaries were composed of mostly nouns, others also had prepositions, and yet others preserved some verbs, adjectives, and adverbs. Figure 1 shows that in the original text nouns (31%), adjectives (9%), adverbs (4%), verbs (15%) preposition (13%) were the most common parts of speech. After the gold-standard summarization, the ratios changed: nouns (54%), adjectives (11%), adverbs (11%), verbs (12%), and prepositions (7%), and these were still the most commonly chosen parts of speech [Figure 2]. This indicates that these parts of speech carried the most meaning in the original text. The "other" category included interjections, pronouns, determiners, predeterminers, etc. In almost all summaries, the punctuation was preserved. We believe this is because punctuation partitions sentences and clauses and helps preserve some of the original semantics.

The analysis of the parts of speech in the gold-standard summaries served as the key to designing the experiments for blind subjects. The results of this experiment prompted us to choose 3 types of summarization to be used in our subsequent experiments. To obtain the three types of summaries, we first merged the 4 gold-standard summaries of each article into a single summary using the following approach: if at least two subjects chose a certain word – we included it in the final summary. We preserved the original word order and punctuation. This summary is shown as summary C in [Table 1]. From summary C, we, then, created two more types of summaries: one with only nouns (summary A) and the other with nouns and prepositions (summary B). In Summary A. we kept only nouns and numerals that also appeared in Summary C. And to produce Summary B, we extended summary A with prepositions. Therefore, summary C contains human summary as it is, which has a mix of parts of speech (noun, verb, adj., prep, adv., others) which were chosen by human. In [Table 1], we provide samples of the 3 summaries (passages A, B, and C) for a paragraph (D) taken from a New York Times article [26]. Note the increasing level of detail from one summary to the other.

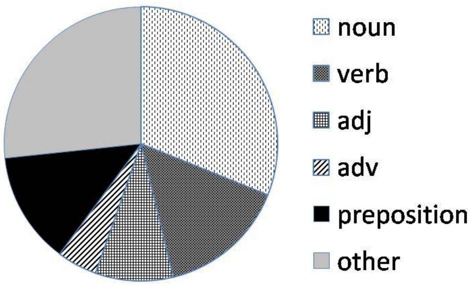

Figure 1. Parts-of-speech distribution in original text

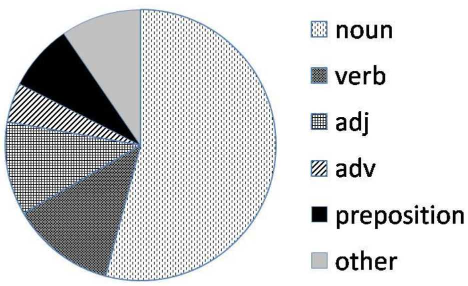

Figure 2. Parts-of-speech distribution in gold-standard summaries

3.2 User Study

To evaluate accessible skimming, we conducted an extensive user study with 20 blind subjects who used our skimming in listening-and-comprehension, as well as in search scenarios; these two common scenarios, where skimming could be most useful, were selected in informal discussions with blind screen-reader users. We then conducted a user study that demonstrated that accessible skimming improved user efficiency in both scenarios. Furthermore, this study helped obtain both quantitative and qualitative data that further validate the usability and effectiveness of the accessible skimming interface. In this section, we describe the experimental set up and discuss the results of the user study.

Table 1 – Gold-Standard Summaries

A: Gold summary with nouns only:
Twitter, 10 person startup San Francisco, Obvious. Mixture networking microblogging. idea, people omnipresence. Use Iran election.
B: Gold Summary with nouns and prepositions only:
Twitter, 10 person startup San Francisco, Obvious. Mixture of networking microblogging. on idea, people omnipresence. Use in Iran election .
C: Combined gold summary
Twitter, 10 person startup San Francisco, called Obvious. Mixture of social networking microblogging. based on idea, people enjoy virtual omnipresence. Use in Iran disputed election.
D: Original paragraph
Twitter, which was created by a 10 person startup in San Francisco was called Obvious. It is a heady mixture of messaging, social networking, 'microblogging' and something called 'presence.' It's shorthand for the idea that people should enjoy an 'always on' virtual omnipresence. Twitter's rapid growth made it the object of intense interest. The object of fair amount of ridicule, as it was derided as high tech trivia or the latest in time-wasting devices. But its use in Iran in the wake of the disputed presidential election of June 2009 brought it new respect. It was used to organize protests and disseminate information in the face of a news media crackdown.

3.2.1 *Participants*

Disability Resource Center of Anonymous State University helped us to recruit 20 blind subjects who participated in the user study. Gender representation was approximately equal. The ages of participants were evenly distributed and varied from early twenties to late sixties. Four of the subjects considered themselves expert computer users, ten of them were very comfortable and six were comfortable with computers. All subjects were well-versed in the use of screen readers, with JAWS as their primary screen reader. Computer usage was more than 20 hours per week for ten subjects, 11-20 hours per week for seven subjects, 6-10 hours per week for one subject, and 1-5 hours a week for the remaining two subjects.

3.2.2 *Setup: Listening and Comprehension*

When experienced sighted readers want to get a gist of the text, they can quickly scan through the text picking out salient information. To test how well our skimming interface can facilitate speed-reading, we designed an experiment testing how much information our subjects could retain while skimming, compared to regular screen reading as a baseline.

In a within-subjects experiment that had a 4 by 4 cell design, we had each subject listen to 4 different articles using 3 types of summaries A, B, and C, and the full text D [Table 1]. We varied the order of the tasks for counterbalancing and to avoid bias.

We considered including a condition that would test the 1st-sentence-of-each-paragraph summary (similar to JAWS skimming), but, after trying it in a pilot study, we concluded that such a summary contained too little information and could not be used effectively in either the listening-and-comprehension or the searching scenarios.

Before the experiment, the subjects were given a sample listening-and-comprehension task to practice with. They were not allowed to use navigation shortcuts in this task. Each task took approximately 10 minutes.

Every task was followed by a set of 10 questions that tested the retention of information. The first question was open-ended and the rest were multiple-choice. The questions we formulated covered the major points of each article without regard for which content actually appeared in the corresponding gold summary and covered the parts of speech that appeared in the gold summaries: 1 question was on the article topic (i.e. "What is the article about?" – "Twitter"), 4 questions on the nouns (e.g., "What was the name of the tweeter start up" – "Obvious"), 3 questions on the verbs (e.g., "What was twitter used for in Iran?" – "organize protests"), 1 on the numeric values (e.g., "How many people organized Twitter?" – "10"), and 1 on the adjectives/adverbs (e.g., "What kind of interest did Twitter generate?" – "Intense"). We approximately followed the distribution of the parts of speech in the original articles and were consistent in this break-down across all articles.

3.2.3 *Setup: Ad-Hoc Searching Scenario*

When experienced sighted or blind readers want to find some information in a lengthy text, but cannot search for the exact keywords, they quickly skim the text until they find a relevant passage and then start reading normally. It is also notable that many screen-reader users do not even use the search feature. To assess whether accessible skimming can help in this scenario, we designed an experiment that tested whether the subjects could find the answer to a specific question quickly using accessible skimming vs. their own ad-hoc skimming techniques as the baseline.

In a within-subjects experiment that had a 2 by 2 cell design, we had each subject skim through 2 different articles using the combined gold summary C and the full text D. Throughout the rest of the paper, we will use the term *skimming* to mean skimming using gold summary and we will use *reading* to mean reading of text using regular shortcuts. The conditions were counterbalanced across subjects to avoid bias. Before each task, the subjects were asked a question, the answer to which they had to find in the article. The questions were formulated used different wording than the actual article, so that the subjects could not search for the exact keywords, e.g., "What is the name of the country in which the results of the voting caused riots?" – "Iran." The subjects were instructed not to try keyword searching.

The answer to the question was contained in the 4th paragraph of each article to increase the time required to find the answer. Having the answer in the beginning of the articles would have caused the users to find the answer equally fast with and without skimming, defeating the purpose of the experiment. Further, we only used one type of summary in this experiment to avoid the learning effect that would have been observed had our subjects figured out that the answer was toward the end of the article. Varying the location of the answer would have increased the variation of the completion times.

In this task, we measured the time it took subjects to reach the sentence that contained the answer to the question, as well as the time needed to understand and give an answer to the question. We also measured the number of keystrokes the subjects pressed to find the answer. Finally, we observed and noted the ad-hoc skimming strategies employed by the subjects while reading the full text. The subjects were given a sample task before the experiment and 5 minutes to practice with the skimming interface. They were allowed to use any navigation shortcuts which they were used to. Each task took approximately 10 minutes.

3.2.4 *Questionnaires*

Following each task in either of the scenarios, the participants were asked to rate the perceived difficulty of the task on a 5-point Likert scale (1=Very Easy to 5=Very Hard) [Figure 4]. At the end of the experiments, we read statements about the skimming experience to the subjects and asked them to rate the statements on a 5-point Likert scale (1=Strongly Disagree to 5=Strongly Agree) – Table 2 summarizes the ratings.

Table 2 – Average 5-Point Likert scale values (St. Dev.) (Scale 1=Strongly Disagree to 5=Strongly Agree)

General Statements	Avg (St Dev)
I wish I could look through articles faster, than I can with a screen reader	3.80 (1.15)
I often experience difficulties looking for desired information within an article	3.40 (0.99)
Skimming made reading through articles faster	4.20 (1.00)
I was able to understand articles equally well when skimming and when reading full article	2.70 (1.34)
I want to use the skimming feature in the future	4.30 (1.03)

3.2.5 *Interface for Accessible Skimming*

To evaluate the accessible skimming interface, we used our Hearsay [5] platform. We used IVONA TTS [17] voice *"Eric"* with speech rate of 180 words per minute. The screen-reader interface enabled subjects to use all the typical navigation shortcuts used by the mainstream screen-readers such as JAWS [19], e.g., paragraph / sentence / word / character navigation, pause / resume, etc. The articles were plain text containing no links or images.

To enable skimming in the searching scenario, we added a new shortcut that allowed users to switch seamlessly between the summary and the original article preserving the current reading position. When switching from the regular screen-reading to the skimming mode, the reading position was reset to the closest preceding word that was in both the original text and the summary. As a result, the user gets an impression that the system continues to read the same content, but is now skipping parts of it.

By switching back and forth, the user can, essentially, control the amount of detail. This feature helped us emulate the behavior of experienced sighted readers who usually can scan quickly and then, at any point, slow down to read the text regularly.

In the searching-with-skimming condition, the subjects could use all navigational shortcuts. The search shortcut was disabled. The subjects were instructed not to try keyword searching because some would and some would not depending on their experience with screen readers. However, to make the scenario realistic, the questions were formulated using the wording different from that of the actual article, so that one could not easily search for the exact keywords.

3.3 Hypotheses and Results

To test the two general hypotheses proposed in the introduction, we formulated a series of specific hypotheses that we accepted or rejected based on the statistical significance testing. In this section we provide the details and the results on all tested hypotheses. Also we present the results of the analysis of the post-completion questionnaire.

We considered statistical significance level of alpha = 0.01, while conducting the t-test tests on the data. The results may have been significant even with a smaller rejection region, but we did not want to increase the probability of the Type II error. The following are the specific hypotheses tested in the experiments and the results.

3.3.1 *Hypotheses and Results: Comprehension*

H1: The questions on general comprehension of the articles can be answered equally well for summary A, B, C, and D.

Results: Subjects were able to answer correctly the question "What is this article about?" with accuracy of 80% (St. Dev.=40%) for summary A, 90% (St.Dev.=30%) for summary B, 100% (St. Dev.=0) for summary C, 100% (St. Dev.=0) for summary D.

Discussion: One of our important goals was to check if skimming adversely affected the general understanding of a text. This hypothesis clearly shows after listening to summary D, the subjects had no problem answering the general comprehension question. This was not the case with summaries A and B that contained no verbs, adjectives, or adverbs.

H2: The use of verbs, adverbs, and adjectives (VAAs) in skimming, in addition to nouns, enables greater comprehension overall compared to skimming without VAAs.

Results: Subjects were able to give correct answers to 56.50% (St. Dev. = 13.86%) and 62.0% (St. Dev. 16.73%) of questions after reading summary A and B, which do not contain any VAAs. In contrast, subjects were able to answer correctly 77.50% (St. Dev.=11.64%) and 80.50% (St. Dev.=14.31%) of questions after listening to summary C and D respectively which contain VAAs. Using the one-tailed paired t-test, we found the differences in the question answering accuracy to be statistically significant ($t=6.352$, $df=19$, $p<0.0001$), accepting H2.

Discussion: The parts-of-speech distributions in the original text and the gold-standard summary presented in Figure 1 and Figure 2 show the essential parts of speech that need to be preserved in the summary. This hypothesis, although unsurprising, suggests that VAAs can be very important in a summary.

H3: Comprehension of any action described in text, increases if the summary includes Verbs, Adjectives, and Adverbs (VAAs).

Results: Our questions about actions / verbs were designed to check subjects' comprehension level of the actions described in the articles, they did not test the ability to guess / remember the exact verb. Subjects were able to answer questions about actions correctly in 50% (St. Dev. = 35%) and 45% (St. Dev.=27%) of cases for summaries A and B respectively. For summaries C and D, subjects were able to answer questions correctly in 78.33% (St. Dev. = 24.83%) and 76.66% (St. Dev.=26.71%) cases on average respectively. So average correctness in answering questions after reading summaries C and D are greater than after reading summaries A and B. This was an obvious result, since summaries A and B did not contain VAAs and we still asked questions about VAAs. However, we have also compared the accuracy of answers for questions about only nouns, which were present in all summaries. Subjects answered questions on nouns correctly: 63.75% (St. Dev.=26.25) for summary A, 72.50% (St. Dev.=17.98%) for summary B, 81.25% (St. Dev.=15.96%) for summary C, and 83.75% (St. Dev.=18.62%) for summary D. The differences between A and B, B and C, C and D were not found to be significant, but the difference between A and C was: ($t=-2.570$, $df=19$, $p=0.009$). This result suggests that punctuation and VAAs together have improved the answers of noun questions. Thus we can only partially accept H3.

Discussion: It is notable, that the difference between A and B, as well between C and D, was not statistically significant, which means that subjects' ability to answer questions about actions was comparable for the gold-standard summary and the original text.

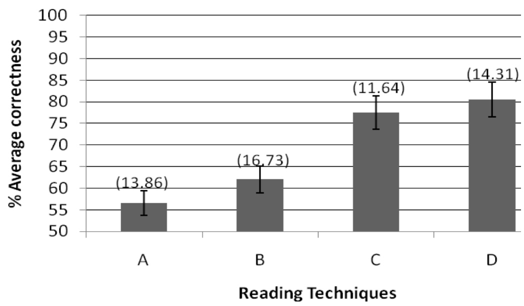

Figure 3 Average correctness (St. Dev.) of question answering with different summaries and full text D

H4: After reading each of the summaries A, B, C, and D subjects can answer increasingly more questions correctly.

Results: Subjects were able to give answers with the average accuracy of 56.50% (St. Dev. = 13.86%), 62.00% (St. Dev. = 16.73%), 77.50% (St. Dev. = 11.64%), and 80.50% (St. Dev. = 14.31%) after listening to summaries A, B, C, and D respectively (Figure 3), even though the actual order of reading these summaries was randomized.

A one-way ANOVA was used to test accuracy differences in answering four types of summaries A, B, C, and D. The accuracy with which subjects gave answers differed significantly across the four types of summaries, $F(3, 76) = 13.44$, $p < 0.0001$. The post-hoc test for linear trend shows that there is a significant trend, that answering accuracies increases as you read summary type A, B, C, and D respectively, Slope 0.04375, $p < 0.0001$.

Discussion: Overall, we observed an increase in the average correctness, but we could not accept H4 for all pairs. Adding prepositions to the nouns (summary B) did not result in significantly better accuracy than using only nouns (summary A). While the presence of prepositions gave the subjects a better idea about the relationship between nouns, it was not enough to help them answer questions more accurately. It is notable, that the biggest difference in comprehension was with the addition of verbs, adjectives, and adverbs. Even more interesting is that *there was no significant difference in question answering accuracy between gold-standard summary and the original text. This implies that summary C carried almost as much information as the full text.*

H5: Subjects perceive the questions to be increasingly easier for summaries A, B, C, and D respectively.

Results: Using the Likert scale (1=very easy to 5=very hard) the subjects rated the difficulty of the listening-and-comprehension task on average as 3.98 (St. Dev. = 0.95) for summary A, 3.35 (St. Dev. =1.03) for summary B, 2.40 (St. Dev. =0.88) for summary C, 1.80 (St. Dev. = 0.89) for the full text, which shows a clear decrease in the difficulty level (Figure 4).

A one-way ANOVA was used to test differences of the difficulties perceived in answering four types of summaries A, B, C, and D. The difficulty perceived by the subjects while answering questions, differed significantly across four types of summaries, $F(3, 76) = 21.05$, $p < 0.0001$. The post-hoc test for linear trend shows that there is a significant trend, that answering difficulty decreases as you read summary type A, B, C, and D respectively, Slope -0.3738, $p < 0.0001$.

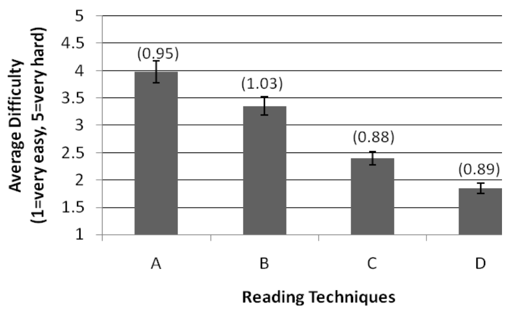

Figure 4 Average perceived difficulty (St. Dev.) of question answering with summaries and full text D

Discussion: The results show that the subjects thought the questions for gold-standard summary were almost as easy as those for the original text. It is notable that, although subjects thought the questions about the full text were a little easier than those for the gold-standard summary C, we did not find the difference to be statistically significant.

3.3.2 *Hypotheses and Results: Searching Scenario*

H6: Finding information using gold summaries is faster than in ad hoc searching.

Results: While finding information using our skimming interface we recorded two timings. First, the time user took to navigate to the answer, which means the time when the screen reader TTS (text-to-speech) engine reads out the answer. Second one is the time when user gives the answer. In Figure 5 we referred to them as "time for TTS" and "time for subject" respectively. While searching for information, subjects took on average 76.35 seconds (St. Dev. = 24.89) to reach the answer using skimming and 148.45 seconds (St. Dev. = 45.16) using ad-hoc searching (Figure 5). There was a significant difference in the speed of skimming ($t=8.011$, $df=19$, $p<0.0001$).The time it took the subjects to actually answer the questions was on average 103.70 seconds (St. Dev. 37.30 seconds) and 169.40 seconds (St. Dev. = 49.54 seconds) for skimming and reading respectively. The one-tailed paired t-test ($t=6.448$, $df=19$, $p<0.0001$) found this difference to be statistically significant; thus, we accept H6.

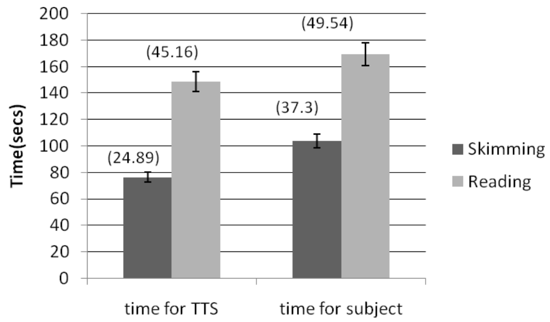

Figure 5 – Average time to reach and understand the answer to a question in skimming vs. normal reading with (St. Dev.)

Discussion: Subjects were able to navigate to the answer location 1.9 times faster using skimming compared to using ad-hoc searching. However, to actually answer the questions, subjects were 1.6 times faster while using skimming compared to ad hoc searching. This can be a substantial time saving to screen-reader users.

H7: There is a difference between the time to reach the answer and the time to actually answer the question.

Results: While using skimming, subjects spent on average 76.35 seconds (St. Dev. = 24.89) to reach the answer and 103.70 seconds (St. Dev. = 37.30) to actually answer to the questions. This difference is significant verified by a one-tailed paired t-test ($t=-6.092$, $df=19$, $p<0.0001$). While using ad-hoc searching, subjects spent on average 148.45 seconds (St. Dev. = 45.16) to reach the answer and 169.40 seconds (St. Dev. = 49.54) to actually answer the questions. One-tailed paired t-test showed this difference is significant ($t=-6.092$, $df=19$, $p<0.0005$). Thus, we accept H7.

Discussion: The result show that it takes subjects time to realize that they have found the answer when they are either skimming or using ad-hoc searching. We expect the difference to be dropping for longer text. But there is still a substantial gain in the speed. We have also computed the differences between the time to answer and time to reach the answer, and then compared these differences between skimming and ad-hoc searching – *it was not significant.*

We found that skimming was almost twice as fast (1.9) as ad hoc searching for reaching the answer. But because, while skimming, it took subjects longer to realize that they have found the answer, skimming was only 1.6 times faster compared to ad-hoc searching. We believe the reason may have been the lack of practice with this new kind of interface. Some of the user comments (in the Testimonials Section) also mentions that it requires practice.

H8: Skimming was easier than ad hoc searching.

Results: We found that, on average, people rated the difficulty of the tasks with skimming as 1.25 (St. Dev. = 0.55) and the difficulty of tasks with reading as 2.30 (St. Dev. = 0.98). Using the one-tailed paired t-test, we found skimming appeared to be significantly easier than reading ($t=-4.702$, $df=19$, $p<0.0001$), rejecting the null hypothesis.

Discussion: This was a surprising finding, especially considering that the subjects have had very little experience with skimming. Recall, that in the reading-and-comprehension scenario, the subjects rated questions for the full text as easier, even though that difference was not found to be statistically significant.

H9: The number of keystrokes pressed while skimming seemed to be always greater than the number of keystrokes used in normal mode. But the one-tailed paired t-test showed there was no difference between these two variables ($t=1.475$, $df=19$, $p=0.922$).

We also found no relationship between: H10) the completion time and the number of hours the subjects used computers per week, H11) the completion time and subjects' comfort level with computers, H12) the amount of time spent at the computer weekly and the perceived difficulty of the tasks (for either searching or listening and comprehension), and H13) the comfort level with computers and the perceived difficulty of the tasks (for searching or listening and comprehension).

3.3.3 *Questionnaires – Analysis and Discussion*

The post completion questions and answers about the skimming experience are presented in Table 2. The analysis showed that the standard deviation for all questions was within one point, demonstrating high consistency.

From the statements made by subjects, we can conclude that skimming has a lot of potential to improve reading experience for screen-reader users. Although the subjects thought that their comprehension was somewhat impeded in the skimming mode, almost all subjects agreed that skimming made the reading faster and they wanted to use skimming in the future. We did not see significant difference in the question-answering performance between the gold-standard summary and the regular text. Two of the subjects mentioned that the tasks may have been easier if they could get more time to practice before the experiment.

4 AD-HOC SKIMMING TECHNIQUES

In our experiments we observed the subjects using their own ad-hoc techniques for skimming using regular navigational shortcuts. Some of the subjects used paragraph navigation to quickly move among paragraphs of the article on the web page. A lot of subjects were reading sentences half way. Some subjects did not use any shortcuts and just listened to the entire articles.

In the follow up interview, we asked open-ended questions to gain deeper understanding into ad hoc skimming techniques. One of the expert users who had 15 years of experience in using JAWS said: "I usually read the first one or two words of a sentence and then go to the next one. Obviously first 1-2 words cannot always give me the whole picture, but it works to some extent." This is an interesting comment that can lead to a simple, yet effective skimming approach. Seven of the subjects mentioned that they often speed up the speech rate for faster reading (in our experiment, the speech rate was controlled and was fixed at 180 words per minute). Most of the subjects would have tried to use JAWS search feature to look for certain information if they knew what exactly they were looking for. Two of the subjects, who frequently read research papers, used JAWS skimming. However, they said JAWS skimming is not comparable to our skimming.

5 SUMMARIZATION FOR SKIMMING

Literature review, examination of visual speed-reading technique, and the experiments with the gold-standard summaries helped us identify the main characteristics of the automated summarization technique that is best suited for supporting non-visual skimming.

Nouns carry the most information and should be the majority of the summary. Prepositions help establish relationship between the nouns. And verbs, adjectives, and adverbs add additional meaning explaining the qualities of and relationship among objects.

When people scan the text they take it as it is, extracting parts of it. Therefore, extractive summarization appears to be the most appropriate method for supporting non-visual skimming. A number of extractive summarization approaches have been proposed in the research literature.

Extractive summarization approaches can be mainly divided into two types: a) sentence extraction and b) keyword extraction. The majority of the existing extractive summarization approaches are doing sentence extraction [3, 8, 13, 15]. While sentence extraction can help get a general overview of text, this approach cannot enable skimming, in which the reader scans through all or most of the sentences picking out the most salient words and phrases. Sentence extraction leaves out most of the original content, which makes it unusable for scanning for information that cannot be found without reading the text, e.g., when keyword search is not possible due to unknown keywords. Sentence extraction assumes coherent text and will not work in web pages, where there is unrelated content, e.g., in front pages of news sites.

After examining the available keyword extraction techniques, we narrowed down the list to the following approaches. In [33], the summary is constructed by extracting word trigrams that frequently occur in the entire document. Marina Litvak [23] presents a keyword extraction approach based on the analysis of a whole document. In [16], the authors build separate graphs for each sentence and weigh each vertex based on the word and structural dependencies within a sentence, as well as word frequency in the entire document. An interesting relation-driven summarization algorithm is proposed in [18]. Any of these approaches could potentially be used to support skimming.

In conclusion, the desired approach should be able to summarize one sentence at a time without, extract meaningful words and word combinations from every sentence, and preserve the original order of the extracted text. The ideal summarization technique should use some syntactic and/or semantic analysis of text for extracting salient parts of the original content. Human generated summaries can serve as a benchmark for evaluating the selected automated summarization methods and the statistical analysis of them can provide clues to the design of automated summarization.

6 TESTIMONIALS

To complete the paper, we are offering these verbatim quotes of our subjects commenting on the skimming:

"If I don't know what I am looking for, this is definitely very handy. I can decide whether to read further or not."

"Search doesn't always work well because there may be multiple results, while skimming is more contextual."

"It's definitely a valid operational concept, which will be very useful."

"I think skimming would be really helpful for long articles, books. In the skimming mode, it was like I was going over my class notes!"

"Usual way of reading takes a lot longer than skimming."

"It's easier, more practical, to find the most important things. I usually have difficulties in reading longer class assignments."

"I usually read the article in its full to get the idea, but with skimming I did not have to read it entirely. It's definitely faster."

"Skimming makes work faster if you need a synopsis."

"It gives key information, gets rid of unnecessary information. Sometimes I increase speech rate to quickly read the article, which is no longer required if we use this method."

"I usually speed up the speech rate to read faster and use paragraph navigation. With skimming it made it easier by giving important words. But it needs more time to practice."

Although the majority of feedback was elated, not all subjects were happy with the skimming approach:

"Skimming breaks up information… Introduces disorganization... There were no reference points… I don't know where I am..."

"Skimming is faster, but important info is sometimes missing."

Nevertheless, the subjects who made those comments did well in the listening-and-comprehension experiment and were able to complete the searching task faster with skimming than in the regular reading mode. This suggests that accessible skimming, just like visual skimming, may require some practice and getting used to.

7 DISCUSSION & FUTURE WORK

The problem we proposed in this paper is, the information overload that people with visual impairment, face during web access. We suggested that the solution lies in skimming. Unlike the sighted people visually impaired ones, will use a specialized screen reader interface for skimming. Though controlled experiments we have proved the importance of this problem and sketched the outline of our solution.

The generic skimming interface we developed for our user studies enabled screen-reader users to skim through the textual content with full control, allowing them to go back and forth between the skimmed summary and the original content. And to enable this interface a specialized summarization is required in its core. By conducting the experiments we have identified the important qualities of the summaries (Section 5) to be used for skimming interface. These qualities of the summaries give sufficient insights on how to design and develop automated summaries. This direction of research will require more experiments on the time and space complexities of the automated summarization approaches and implementation factors with screen readers.

There were highly positive feedback from the study participants, which validates the usability of the interface. Furthermore, since subjects were able to perform several browsing tasks significantly faster with skimming, and skimming did not impede subjects' understanding of the content, skimming proves to be a very useful feature for any screen-reading software. All our hypotheses have been analyzed with appropriate statistical tests (i.e., t-test, ANOVA). In particular to verify the hypotheses that involved three or more groups, we used ANOVA instead to pair wise t-tests, to reduce the chance of Type I error. We also conducted a post-hoc test for linear trend to determine the direction.

Our user interface had only two options: read the full text and read its summarized version. The results suggest that variable-speed skimming may be viable. Changing the speed of skimming in the process of reading may help better mimic the behavior of sighted readers and give screen-reader users more control over the interface.

A comment – "I wish skimming could be applied to non article page as well, where you can give a general name to a group of page elements. And I wish it could strip a web page into important parts similar to skimming" – reveals another possible direction of research that would allow users to skim other types of content, for example a list of links. Other summarization approaches would be necessary to enable this type of skimming.

8 ACKNOWLEDGEMENTS

We would like to express our deepest gratitude to Terri Hedgpeth, Ed.D., the Director of the Disability Resource Center at the Arizona State University for helping us organize and conduct the user study with blind individuals. This work has been supported by NSF Awards IIS-0808678 and CNS-0751083.

9 REFERENCES

[1] Ahmed, F., M.A. Islam, Y. Borodin, and I.V. Ramakrishnan, *Assistive web browsing with touch interfaces*, in *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. 2010, ACM: Orlando, Florida, USA. p. 235-236.

[2] Amitay, E. and C. Paris., *Automatically summarising web sites: is there a way around it?*, in *CIKM '00: Proceedings of the ninth international conference on Information and knowledge management*. 2000, ACM Press. p. 173–179.

[3] Berger, A.L. and V.O. Mittal., *Ocelot: a system for summarizing web pages*, in *SIGIR '00: Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. 2000, ACM Press. p. 144–151.

[4] Bigham, J.P., A.C. Cavender, J.T. Brudvik, J.O. Wobbrock, and R.E. Ladner, *WebinSitu: A Comparative Analysis of Blind and Sighted Browsing Behavior*, in *Assets '07*. 2007, ACM New York, NY, USA ©2007: Tempe, Arizona, USA.

[5] Borodin, Y., F. Ahmed, M.A. Islam, Y. Puzis, V. Melnyk, S. Feng, I.V. Ramakrishnan, and G. Dausch, *Hearsay: a new generation context-driven multi-modal assistive web browser*, in *Proceedings of the 19th international conference on World wide web*. 2010, ACM: Raleigh, North Carolina, USA.

[6] Buyukkokten, O., H. Garcia-Molina, and A. Paepcke, *Seeing the Whole in Parts: Text Summarization for Web Browsing on Handheld Devices*, in *Proceedings of the 10th International World Wide Web Conference*. 2001.

[7] Chieko, A. and C. Lewis, *Home page reader: IBM's talking web browser*, in *Closing the Gap Conference Proceedings*. 1998.

[8] Chuang, W.T. and J. Yang, *Extracting sentence segments for text summarization: a machine learning approach*, in *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*. 2000, ACM: Athens, Greece. p. 152-159.

[9] DeJong, G., *Skimming stories in real time: An experiment in integrated understanding*, in *Research Report 158*. 1979, Department of Computer Science, Yale Univesity.

[10] Delort, J.Y., B. Bouchon-Meunier, and M. Rifqi, *Enhanced web document summarization using hyperlinks*, in *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia*. 2003, ACM: Nottingham, UK. p. 208–215.

[11] Farhoomand, A.F. and D.H. Drury, *Managerial information overload*, in *Commun*. 2002, ACM. p. 127–131.

[12] Fischer, G. and C. Stevens., *Information access in complex, poorly structured information spaces*, in *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*. 1991, ACM Press. p. 63–70.

[13] Goldstein, J., M. Kantrowitz, V.O. Mittal, and J.G. Carbonell, *Summarizing Text Documents: Sentence Selection and Evaluation Metrics*, in *In Research and Development in Information Retrieval*. 1999. p. 121-128.

[14] Gong, Y. and X. Liu, *Generic text summarization using relevance measure and latent semantic analysis*, in *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. 2001, ACM Press. p. 9–25.

[15] Harper, S. and N. Patel, *Gist summaries for visually impaired surfers*, in *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility*. 2005, ACM: Baltimore, MD, USA.

[16] Hori, C., S. Furui, R. Malkin, H. Yu, and A. Waibel, *A statistical approach to automatic speech summarization.* EURASIP J. Appl. Signal Process., 2003. 2003: p. 128-139.

[17] IVONA. *IVONA multi-lingual speech synthesis system*. 2005; Available from: http://www.ivona.com/.

[18] Jacobs, P.S., *To parse or not to parse: relation-driven text skimming*, in *Proceedings of the 13th conference on Computational linguistics - Volume 2*. 1990, Association for Computational Linguistics: Helsinki, Finland. p. 194-198.

[19] JAWS, *Screen reader from Freedom Scientific*. 2011. http://www.freedomscientific.com/products/fs/jaws-product-page.asp.

[20] JAWS_Skimming. *Skim Reading*. 2011; Available from: http://www.freedomscientific.com/Training/Surfs-up/Skim_Reading.htm.

[21] Kurtus, R. *Reading Faster*. 2001; Available from: http://www.school-for-champions.com/grades/reading.htm.

[22] Lamb, A. and L. Johnson. *Skimming and Scanning*. 1999; Available from: http://42explore.com/skim.htm.

[23] Litvak, M. and M. Last, *Graph-based keyword extraction for single-document summarization*, in *Proceedings of the Workshop on Multi-source Multilingual Information Extraction and Summarization*. 2008, Association for Computational Linguistics: Manchester, United Kingdom. p. 17-24.

[24] MindTools. *Speed Reading*. Available from: http://www.mindtools.com/speedrd.html.

[25] NVDA, *Non-Visual Desktop Access*. 2011. http://www.nvda-project.org/.

[26] NYTIMES. *Twitter*. 2010; Available from: http://topics.nytimes.com/top/news/business/companies/twitter/index.html.

[27] Parmanto, B., R. Ferrydiansyah, A. Saptono, L. Song, I.W. Sugiantara, and S. Hackett, *AcceSS: accessibility through simplification & summarization*, in *Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A)*. 2005, ACM: Chiba, Japan. p. 18-25.

[28] Radev, D.R., E. Hovy, and K. McKeown, *Introduction to the special issue on summarization.* Computational Linguistics, 2002. 28(4): p. 399-408.

[29] Serenson. *Speed Reading Methods*. Available from: http://www.write-better-english.com/speed-reading-methods.aspx.

[30] VoiceOver, *Screen reader from Apple*. 2010. http://www.apple.com/accessibility/voiceover.

[31] Window-Eyes, *Screen Reader GW Micro*. 2010. http://www.gwmicro.com/Window-Eyes.

[32] Young, S. and P. Hayes, *Automatic classification and summarization of banking telexes* in *The Second Conference ok Artificial Intelligence Applications*. 1985, IEEE Press. p. 402-208.

[33] Zajicek, M., C. Powell, and C. Reeves, *A Web navigation tool for the blind*, in *Proceedings of the third international ACM conference on Assistive technologies*. 1998, ACM: Marina del Rey, California, United States. p. 204-206.

[34] Zajicek, M., C. Powell, and C. Reeves, *Web search and orientation with BrookesTalk*, in *Technology and Persons with Disabilities Conference (CSUN)*. 1999.

[35] Zechner, K., *Fast Generation of Abstracts from General Domain Text Corpora by Extracting Relevant Sentences*, in *COLING '96 Proceedings of the 16th conference on Computational linguistics - Volume 2*. 1996.

[36] Zhang, Y., *World Wide Web Site Summarization*, in *Technical report, Faculty of Computer Science*. 2002, Dalhousie University.