How People use Presentation to Search for a Link: Expanding the Understanding of Accessibility on the Web

Caroline Jay[†], Robert Stevens[†], Mashhuda Glencross[†] and Alan Chalmers[‡]

[†] School of Computer Science
University of Manchester

[‡] Department of Computer Science
The University of Bristol
E-mail: caroline.jay@manchester.ac.uk

ABSTRACT

It is well known that many Web pages are difficult to use by both visually disabled people and those who use small screen devices. In both cases there exists a problem of viewing a great deal of information with presentation capabilities reduced from the intended formatted large-screen colour display. It is pertinent, however, to ask how the presentation of Web pages on a standard display makes them easier for sighted people to use. To begin to answer this question, we report on an exploratory eye-tracking study that investigates how sighted readers use the presentation of the BBC News Web page to search for a link. We compare the standard page presentation with a "text-only" version and observe both qualitatively and quantitatively that the removal of the intended presentation alters "reading" behaviours. The demonstration that the presentation of information assists task completion suggests that it should be re-introduced to non-visual presentations if the Web is to become more accessible. Finally, we propose that models derived from studies that reveal how presentation is used to aid task completion can form the basis for annotation and transcoding of Web pages to present pages in a more usable non-visual form.

Categories and Subject Descriptors

K.4.2 [**Computers and Society**]: Social IssuesAssistive Technologies for Persons with Disabilities; H.1.2 [**Information Systems**]: User/Machine SystemsSoftware Psychology; H.5.2 [**Information Interfaces and Presentation**]: User InterfacesEvaluation/methodology; H.5.2 [**Information Interfaces and Presentation**]: User InterfacesScreen Design; I.7.5 [**Document and Text Processing**]: Document CaptureDocument Analysis

1. INTRODUCTION

In this paper we explore the problems encountered by people with profound visual disabilities when browsing Web pages in a manner that replicates interaction with small screen devices. We ask two questions: firstly, what is it about audio screenreader presentation of a Web page that really makes it difficult to use; secondly, what is it about the standard presentation that is missing and needs to be replaced? The obvious and unattainable answer is that the external memory provided by paper or a screen, together with the speed and accurate control of information flow afforded by the human visual system is required [17]. All assistive technology for visually disabled people attempts, in some way, to provide some kind of replacements for these fundamental facilities.

The difficulties with access to Web content by visually disabled people is well documented [6, 5], however what is less well documented is the exact nature of the problems encountered by those users. Profoundly Visually disabled people usually use a screenreader[1] (such as Jaws [1, 8]), which speaks screen content under the direction of the user, in order to "read" what is on the screen. When a Web page is loaded into a browser, by default the page contents are typically spoken from top-left to bottom-right. Naturally, such a spoken presentation is difficult to use—memory cannot handle such quantity and review is not really possible. Of course, the screenreader allows finer control over what is spoken than this cumbersome mechanism. A user can move around a page using cursor keys. This allows movement at the level of lines, characters, words, paragraphs, *etc*. It would seem that all parts of a page can be accessed; so, what is the problem?

The screenreader access described above is slow and cumbersome, but still presented as difficult. What makes such access difficult? Much effort in Web accessibility has focused on access to content. This includes access to graphics; tables; figures; *etc*. Much has been achieved, but access to Web pages continues to be a problem; why is this?

In a text-only presentation (such as that shown in Figure 1), the majority of the formatting and layout is removed [19]. The user is given what amounts to a traditional linear document, but with little differentiation between text serving different purposes, such as headings, plain text and emphasised text that usually supports the processing of the printed information [20]. Through eye-tracking studies, we have examined how readers use a Web page's presentation to accomplish their task. We have compared this with how the presentation of the text-only version of the same page is used

[1]In this article we restrict ourselves to those users who are profoundly visually disabled. Assistive technology for users with impaired vision cover a broad range of tools including magnification, colour or contrast enhancement, *etc*.

W4A at WWW2006 23-26 May 2006, Edinburgh, UK

to accomplish the same task. The latter simulates, to an extent,both the presentation on small screen and the experience of a visually disabled person using a Web page through reduced visual effects and reduced information content on the screen.

The presentation of a page generally facilitates a sighted user in achieving their goal, be it searching, browsing, following links *etc*. It is well known from the principles of graphic design and information presentation that how content is laid out can influence how a task is performed and what interpretation is made of data [20]. If we are to support visually disabled and small-screen users to use Web pages it is important to characterise the situation beyond "difficult". We would like to have insight into how Web presentation is used by sighted users. It is not the presentational qualities themselves that need to be re-introduced, but what users do with those formatting and layout features. We attempt to do this by examining eye-tracking data during the exercise of basic Web based tasks on standard and text-only versions of a page—which features are looked at and for how long in each presentational style?

Once we know the difficulties a lack of presentation causes, it will be easier to know what to replace or re-introduce into audio presentations. Our eye-tracking studies have produced a vast quantity of data, of which we examine only a portion here. From these studies, we can see how a human uses the page's presentation during navigation and orientation. From our observations of how readers use presentational information to orientate and navigate in order to accomplish tasks, we propose that eye-tracking studies can be used to provide models of page usage that can be applied to different styles of presentation in order to transcode those pages. It is known that such transformations can, to some extent, replace some of the facilities of a standard Web page presentation [2]. Our investigations should provide a basis for driving such transformations.

In Section 2 we describe the context in which this work has been performed. We then describe our experiment: its design, procedure and equipment (see Section 3). Our results are presented in Section 4. A discussion of the results is given in Section 5.

2. RELATED WORK

Profoundly visually impaired users usually access the Web using screen readers [8] or specialised browsers [1]. There are many sources that state the difficulty that such users have using the Web *via* tools such as screenreaders. The Disability Rights Commission (DRC) report of 2004 [6] concludes that most UK Web sites (81%) fail to satisfy even basic accessibility requirements. A similar situation exists in the USA [5]. Coyne and Nielsen [5] concluded that "the Web is about three times more difficult to use for people who are visually impaired than it is for sighted users". The question still remains, however, as to exactly what makes it so difficult.

One part of the answer is in simply extracting information encoded within the HTML of a Web page. Thatcher *et al.* [19] give practical advice on constructing accessible Web sites by elucidating guidelines [4] and using HTML to best advantage. Text-only alternatives to standardly formatted Web pages are offered as "accessible" options. However, [19, Chapter 1] states "the issue of text only versions crosses into the idea of separate versus inclusive design". Why should visually disabled people use different pages to sighted users? As well as problems of maintenance *etc*., there is a resistance to the development of Ghettos. [15] observe that, "designers look at sites that are meant to serve as models of accessibility and are appalled by the aesthetics. For most designers, accessibility equates with boring, uninteresting designs. The state of accessibility on the Web today represents a failure of the imagination".

Even when pages are built according to guidelines that are meant to increase accessibility, there seem to be "difficulties". Takagi *et al.* [18] classified the problems into three categories:

1. Adherence to guidelines, not usability issues;
2. Over-reliance on syntactic checking of Web pages;
3. No attention on the time component in the operations provided to aid access.

These observations usefully describe some of the sources of problems encountered by visually disabled users without really describing the nature of those problems.

The problem is summed up by Hanson [7], who says, "specifications for accessibility of Web pages do not necessarily guarantee a usable or satisfying Web experience for persons with disabilities. It is not uncommon to have pages that meet standards but are still difficult to use by persons who have difficulties."

Harper *et al.* [3] and Yesilada *et al.* [22] have used the metaphor of travel to raise the notion of using a Web page above that of dealing with mere 'sensory translation'. Just as people use *travel* objects in the environment (signs, landmarks and other cues) to help them orientate and navigate, so a Web user can use *travel objects* on a Web page to aid mobility—the ease of travel in a Web page. The layout and presentation of a Web page provides these travel objects. A well designed page eases travel and a badly designed page hinders travel and decreases mobility. A text-only page will tend to lack more of these travel objects and thus the reader is again hindered in their task.

Screen readers, unlike sighted users, cannot see the implicit structural and mobility knowledge encoded within the visual presentation of Web pages [21]. "It is impossible for blind users to distinguish visually fragmented groupings only from the sequence of tags read to them [2]. " So, we begin to see that basic sensory translation of what is encoded within HTML still lacks what is necessary to support usable access to the Web. We can identify the Web correlates of what in printed material enables and supports effective information processing [20], as something lacking in current non-visual renderings of Web based material.

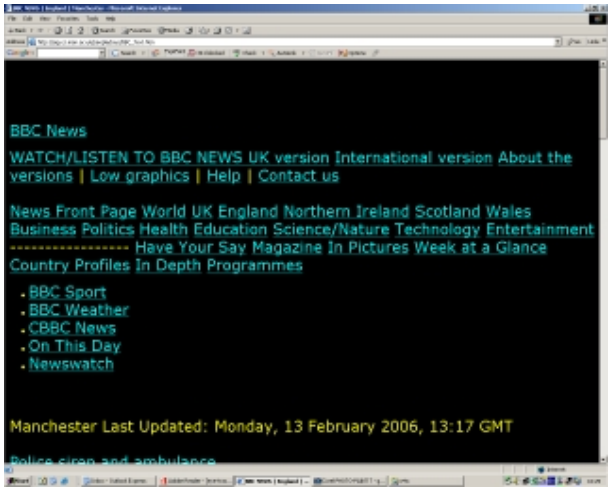
We have used eye-tracking in this study because we wish to find out how sighted users exploit the visual presentation of a Web page. Once we have an insight into what it is that visually disabled people can no longer do without this information then it is possible to begin to plan to replace those facilities. Whilst eye tracking has been used to investigate cognitive processes for over 100 years [14], monitoring people's gaze during Web use is a relatively new discipline. Recording the pattern of fixations on a Web page is a powerful tool, enabling us to determine those areas that are most salient (attract the most fixations), and those that receive little attention.

Its most obvious applications are in improving the standard design and layout of Web pages, and evaluating their usability [16]. Studies have also examined the saliency of items on a page under varying conditions (*e.g.*, [9, 12, 10]), how eye movements vary according to information scent [13] and how looking for a menu is influenced by page complexity and prior expectations [11]. To our knowledge, no one has investigated how gaze patterns differ when searching on either the standard or text-only versions of the same page. Here, we do just that, by tracking participants' eye movements as they search for specific links on the standard and text-only versions of the BBC News Web site.

3. METHODOLOGY

Eye-trackers provide an effective means of recording where a user's gaze falls on a particular part of a computer screen, and hence

(a) Standard formatted BBC Webpage



(b) Text-only version of BBC Web page

Figure 1: The BBC Web page used in this study. 1(a) is the standard formatted versions and 1(b) is the text-only version of the same page. Cached versions can be found at `http://aig.cs.man.ac.uk/people/jayc/bbc_study.htm`.

a good indication of the location of his or her attention. The number of fixations an area receives indicates the amount of attention it attracts; the average duration of these fixations corresponds to its complexity (more complex stimuli require longer fixations to process) [14].

In the current study, participants' eye movements were tracked while they were searching for a link on either the standard or text-only versions of the BBC News Website (see Figure 1).[2] The time to locate the link provided a measurement of task difficulty. The number of fixations indicated the amount of attention required to find the link, and the average duration of fixations indicated the relative complexity of the information presented on each page.

3.1 Equipment

The equipment used for the study is displayed in Figure 2. The experiment was run on a SONY VAIO VGN-FS315S laptop. Stimuli were displayed on a SONY VAIO SDM-HS75P monitor and a Tobii x50 Eye Tracker, positioned at the base of the monitor, tracked the participant's gaze. The Tobii ClearView Analysis software was used to record and analyse eye-movement and event data.

3.2 Design

The study used a between-subjects design in which participants searched for links on the BBC News Manchester Web page (see Figure 1). This site was chosen as it is widely known in the UK, but the actual content on any one day will change. BBC pages are also automatically provided with a text-only page. Sighted users, despite having the facilities of the visual system outlined in Section 1, will experience much of the effect of an audio screenreader rendering through experiencing a text-only version. The text-only version will also reflect the effects of a small-screen rendering, irrespective of the BBC's provision of a version for small-screen devices. This mimicry arises from the text-only version having little of the visual formatting of the standard presentation and the consequent reduction of information content. Of course, the correspondence

Figure 2: The equipment used

[2] `http://aig.cs.man.ac.uk/people/jayc/bbc_study.htm`

between sighted users reading a text-only page and the reality of profoundly visualy disabled people's use of an audio screenreader is only partial at best. It does, however, provide some flavour of the interaction.

Half of the participants searched for a link to the BBC Manchester Website on the standard version of the page, and a link to a story about the Chinese community on the text-only version of the page; the other half searched for the link to the BBC Manchester Website on the text-only page, and the link to the story about the Chinese community on the standard page. The links were chosen because they were next to each other on the page, and positioned on the right in the standard version and approximately two-thirds of the way down in the text-only version, so participants would not see them immediately on entering the page. The presentation order of the two site versions alternated between users. At the end of the experiment, users were asked to indicate whether they used the BBC website more than once a week, and whether they found it easier to perform the search on the standard or text-only version of the page.

3.3 Participants and Procedure

Eighteen participants between the ages of 17 and 50 with normal or corrected vision took part. The study was carried out as part of visit days by prospective students to the School of Computer Science.[3] Both prospective students and their parents were asked to volunteer to take part in the study. The procedure was explained and the study performed, which took about five minutes. Participants were shown the recording of their eye-movements after the experiment.

Each participant sat 50 cm from the monitor and went through the calibration process. The participant was told to look for either the BBC Manchester or Chinese community link. The participant entered the appropriate version of the page and started searching, indicating that he or she had located the link by hovering over it with the mouse. The participant then searched for the other link on the alternative version of the page.

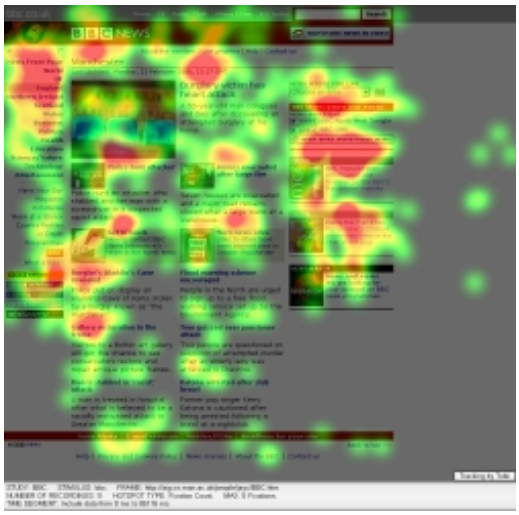
4. RESULTS

The eye tracker recorded the position of the participant's gaze on the Web page throughout the experiment. We used the ClearView analysis software to calculate the number of fixations that occurred (a fixation is where the eyes rest on part of the stimulus for more than 100 msec, in order to process information), their position, their order, and their average duration. Figures 3 and 4 show the areas of the page that received the most fixations in the standard and text-only versions of the page respectively (red = 9 or more fixations, green = 4 or more, grey = 0). Participants tended to fixate on the salient areas when searching the standard page: headlines, some images, and prominent words in the text. The areas with the most fixations also appeared to vary according to the link participants were searching for - when looking for the Manchester link, they looked more at the menus, as this may seem a logical place from which to navigate to a BBC regional Website. On the text-only page, participants fixated in a uniform manner on nearly all of the text on the left-hand side of the page, indicating that it was allocated equal importance as they read down from the top until they found the link.

Qualitative analysis of the gaze replay and fixation order data confirms that participants simply read down the page in a serial fashion when looking for the link on the text-only page. On the standard page, however, participants' eyes dart around as they attempt to locate what appears visually to be the next most likely location for the link. The gaze plots in Figure 5 illustrate this process for participants searching for each link on the standard page and those in Figure 6 show the same two participants searching for the other link on the text-only page. When looking at the standard page, participants will make large saccades to move from one salient area to the next; on the text page however, saccades are much smaller, as the participants read through the text from top to bottom.

A two-way ANOVA (page $\times$ link) showed that participants made significantly more fixations on the text-only site, regardless of the link they were searching for, than they did for the standard site ($F_{1,32} = 2.183$, $p<0.005$). This suggests that the text-only version of the page required more of their attention to process (see Fig-

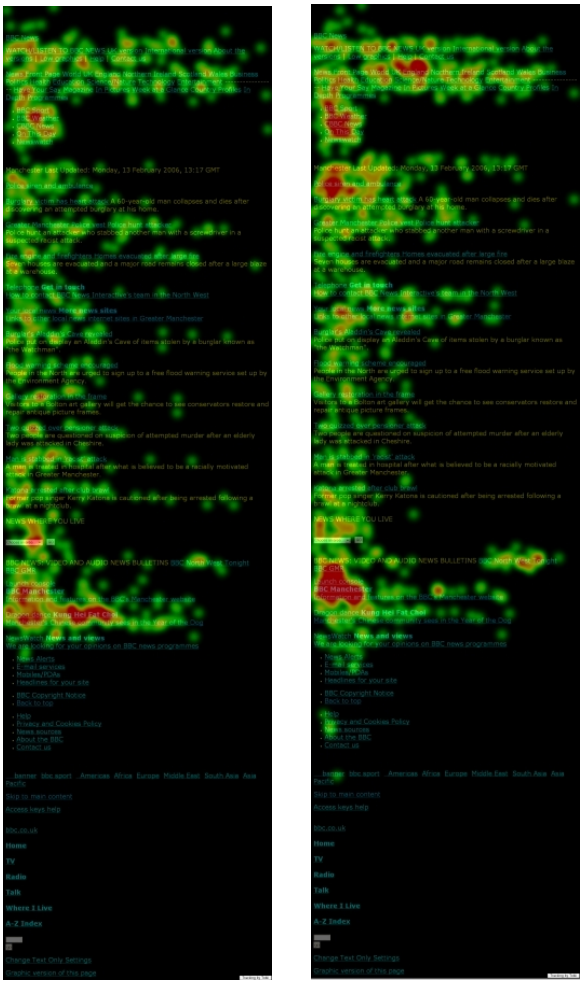
(a) China Search



(b) Manchester Search

Figure 3: Hotspot analysis of standard presentation of the BBC Manchester Web page.

[3] http://www.cs.manchester.ac.uk



(a) China Search (b) Manchester Search

Figure 4: Hotspot analysis of text-only presentation of the BBC Manchester Web page.

(a) Gazeplot for Manchester link

(b) Gazeplot for China link

Figure 5: Gazeplots for participants searching for the Manchester and China links on the standard BBC Manchester Web page.

ure 7). The duration of a fixation, however, lasted significantly longer on the standard page ($F_{1,32}$ = 0.208, p<0.005), indicating that a greater amount of cognitive processing was occurring in a fixation on the standard page than the text-only page (see Figure 9). It may be that the information obtained in a single fixation on the standard page was more complex, or that some of the extra processing time was due to participants orientating themselves and planning their navigation to the next part of the page (fixations during visual search are known to be longer than fixations during silent reading [14]).

The majority of participants found it easier to search for the link on the standard version of the page (seefigure 8).It is important to acknowledge that familiarity with the page may have had a strong influence – all but two of the participants used the BBC website more than once a week, and of those who did not, one rated the text-only page as easier and the other did not express a preference. However, the possibility that participants found it easier to search the information in the standard page because of the way it was laid out cannot be dismissed.

Completion times (shown in figure 10) did not vary significantly as a result of page type. Closer analysis shows that this result occurred due to the serious difficulty two participants had locating the Manchester link (taking more than 50 seconds to find it, in contrast to the 13 seconds it took the other participants). This may have arisen due to the prior expectations of the participants. Both spent a long time looking at menus, convinced that the link should be located on one, rather than positioned on the right of the page. A T-test considering the China link searches alone shows that participants locate it significantly more quickly on the standard page (t_{16} = 3.696, p<0.005). It is also worth considering the fact that the two pages varied not only in layout, but also in colour. Sighted users, used to seeing white text on a black background, may have have had more difficulty processing yellow text on a black background.

(a) Gazeplot for Manchester link

(b) Gazeplot for China link

Figure 6: Gaze plots for participants searching for the Manchester and China links on the textual BBC Manchester Web page.

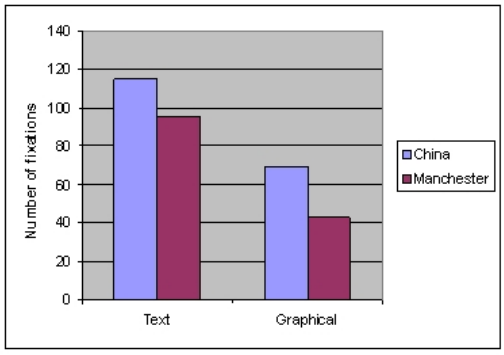


Figure 7: Mean number of fixations on each type of page.

5. DISCUSSION

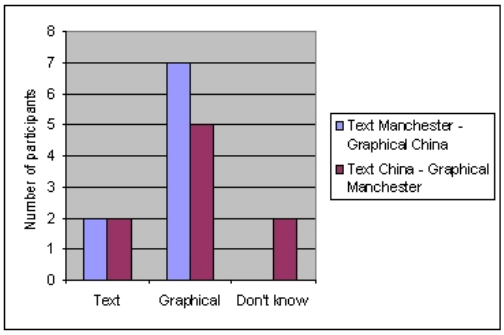


Figure 8: Number of participants who found it easier to locate the link on the standard and text-only pages.

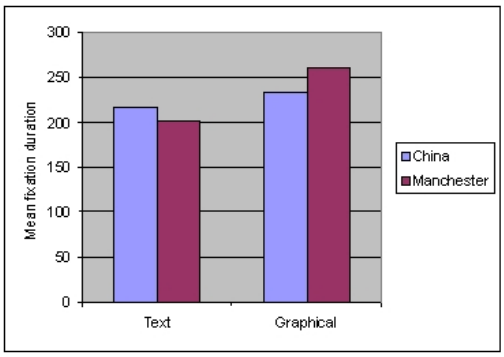


Figure 9: Mean duration of a fixation on each type of page.

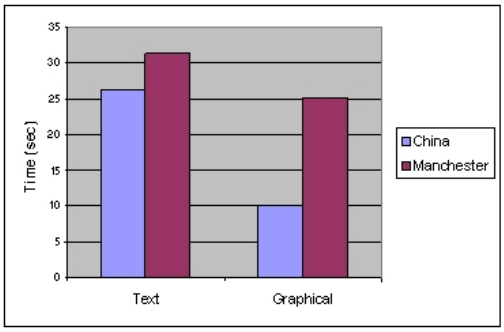


Figure 10: Mean time in seconds to locate the link.

An Audio rendering lacks the very facilities that make the visual system the predominant human information gathering system, that is, the ability to control information flow from the outside with speed and accuracy, because there is an external memory holding the information [17]. In the case of Web pages, there is the additional factor of even being able to access what is being presented—the focus of much of the work reviewed in Section 2. Current audio renderings lack the formatting and presentation available in the visual mode: Spacing; typefaces; font-size; colour; *etc.* Non-speech and audio cues can potentially add some of this information, but this has not so far been greatly exploited in the rendering of Web pages.

In this exploratory experiment we have attempted to further characterise the problem of Web accessibility. In this study qualitatively we see a stark difference in eye-movement behaviour in the

two conditions. In the standard, formatted presentation (preferred by the majority of participants), saccades are greater and fixations longer. The link menu is being used, along with the distinct presentational fragmentation of the page into areas. In contrast, the text-only page provoked eye-movements more akin to those seen in reading of ordinary printed text. The page is being "read", rather than being used to navigate towards the link goal. The speculation is that a sighted user can use the formatting of the page to achieve the task and that the longer gaze duration is a consequence of decision making on orientation and navigation. The text-only page is in effect similar to that which a blind user encounters in an audio presentation. The eye-movements seen in this condition are similar to those seen in cursor movements during reading. The lack of opportunity to "dart" around the page to orientate and navigate are removed.

Text-only versions of Web pages have been advocated as options for presentation on small-screen devices. The observations of this investigatory study would suggest that this is not sufficient for a usable interaction with a page on a mobile device. Similarly, for visually disabled users, a text-only version of a Web page does make the information available, but using that information is seen to be difficult without the facilities that a formatted presentation affords a sighted Web user.

So the question is how to replace the facilities afforded by presentation to both visually disabled users and those sighted people who use small screen devices? Layout provides borders for portions of information. Ideally, the objects on a page guide a reader through the information such that he or she can accomplish their task [20]. This is the basis of the use of the travel metaphor [3] in increasing Web accessibility. Here, the *travel objects* that facilitate movement through a Web page are identified and semantically marked up [22]. This markup is then used to help transcode a page into either an order that facilitates use or a fragmentation that facilitates use.

Eye-tracking studies have confirmed our hypothesis about the utility of a standard and a text-only presentation. One way of bringing this information back into a presentation is *via* transcoding. Does this eye-tracking data provide any mechanism by which transcoding can be driven? Obviously, markup and transcoding cannot occur by hand alone. Our aim is to see whether we can create models of how presentation is used for navigation and orientation about a page such that we can look at a page's presentation and infer how the presentation will be used. Such models will be able to drive the transcoding of a page and begin to re-introduce into an audio rendering the same support for tasks that a sighted user has in presentational cues for task completion.

Acknowledgements: Caroline Jay is funded by EPSRC grant EP/D036518/1.

6. REFERENCES

[1] C. Asakawa and T. Itoh. User interface of a home page reader. In *ASSETS'98*, pages 149–156. ACM Press, 1998.

[2] C. Asakawa and H. Takagi. Annotation-based transcoding for nonvisual web access. In *ASSETS'00*, pages 172–179. ACM Press, 2000.

[3] Goble C.A., S. Harper, and R.D. Stevens. The Travails of Visually Impaired Web Travellers. In *ACM 2000 Hypertext (Proceedings of the Eleventh ACM Conference on Hypertext and Hypermedia)*, pages 1–10. ACM Press, 2000.

[4] W. Chisholm, G. Vanderheiden, and I. Jacobs. Web content accessibility guidelines 1.0. World Wide W eb Consortium (W3C), 1999.

[5] K. P. Coyne and J. Nielsen. Beyond ALT text making the web easy to use for users with disabilities. Nielson Norman Group, 2001.

[6] Disability Rights Commission. The web access and inclusion for disabled people. Technical report, Disability Rights Commission (DRC), UK, 2004.

[7] V.L. Hanson. The user experience designs and adaptations. In S. Harper, Y. Yesilada, and C. Goble, editors, *Proceedings of the international cross-disciplinary workshop on Web accessibility*, pages 1–11, 2004.

[8] Henter-Joyce, Inc. *Jaws*.

[9] T. Joachims L. Granka and G. Gay. Eye-tracking analysis of user behaviour in www search. In *SIGR'04*, pages 478–479, Sheffield, UK, 2004. ACM Press.

[10] A. Tholke M. Schiessl, S. Duda and R. Fischer. Eye tracking and its application in usability and media research. In *MMI-Interactiv '03*, pages 41–50, 2003.

[11] J. McCarthy, A. Sasse, and J. Riegelsberger. Could i have a menu please? an eye tracking study of design conventions. In *Proceedings of HCI2003*, pages 401–414, 2003.

[12] B. Pan, H. A. Hembrooke, G. K. Gay, L. A. Granka, M. K. Feusner, and J. K. Newman. The determinants of web page viewing behavior: An eye-tracking study. In *Eye Tracking Research & Applications Symposium on Eye-Tracking Research & Applications*, pages 147–154, San Antonio, Texas, 2004. ACM Press.

[13] P. L. Pirolli, S. K. Card, and M. Van Der Wege. Visual information foraging in a focus+context visualization. In *Proceedings of ACM Conference on Human Factors and Computing Systems (CHI 2001)*, pages 506–513, Seattle, 2001. ACM Press.

[14] K. Rayner. Eye movements in reading and information processing: 20 years of research. In *Psychological Bulletin*, pages 372–422, 1998.

[15] B. Regan. Accessibility and design a failure of the imagination. In S. Harper, Y. Yesilada, and C. Goble, editors, *Proceedings of the international cross-disciplinary workshop on Web accessibility*, pages 29–37, 2004.

[16] M. Russell. Using eye-tracking data to understand first impressions of a website. In *Usability News 7.2*, 2005.

[17] R. D. Stevens, A. D. N. Edwards, and P. A. Harling. Access to mathematics for visually disabled students through multi-modal interaction. *Human-Computer Interaction (Special issue on Multimodal Interfaces)*, pages 47–92, 1997.

[18] H. Takagi, C. Asakawa, K. Fukuda, and J. Maeda. Accessibility designer visualizing usability for the blind. *ASSETS'04*, pages 177–184, 2004.

[19] J. Thatcher, P. Bohman, M. Burks, S.L. Henry, B. Regan, S. Swierenga, M.D. Urban, and C.D. Waddell. *Constructing Accessible Web Sites*. Glasshaus, 2002.

[20] P. Wright. The psychology of layout: Consequences of the visual structure of documents. In R. Power and D. Scott, editors, *American Association for Artificial Intelligence Technical Report FS-99-04*, pages 1–9. AAAI Press, 1999.

[21] Y. Yesilada, S. Harper, C. Goble, and R. Stevens. Screen readers cannot see (ontology based semantic annotation for visually impaired web travellers). In N. Koch, P. Fraternali, and M. Wirsing, editors, *Web Engineering - 4th International Conference, ICWE 2004 Proceedings (LNCS 3140)*, pages 445–458. Springer, 2004.

[22] Yeliz Yesilada, Robert Stevens, and Carole Goble. A

foundation for tool based mobility support for visually impaired web users. In *Proceedings of the Twelfth International World Wide Web Conference*, 2003.