Analysis of Navigability of Web Applications for Improving Blind Usability

HIRONOBU TAKAGI, SHIN SAITO, KENTAROU FUKUDA,
and CHIEKO ASAKAWA
Tokyo Research Laboratory, IBM Research

13

Various accessibility activities are improving blind access to the increasingly indispensable WWW. These approaches use various metrics to measure the Web's accessibility. "Ease of navigation" (navigability) is one of the crucial factors for blind usability, especially for complicated webpages used in portals and online shopping sites. However, it is difficult for automatic checking tools to evaluate the navigation capabilities even for a single webpage. Navigability issues for complete Web applications are still far beyond their capabilities.

This study aims at obtaining quantitative results about the current accessibility status of real world Web applications, and analyzes real users' behavior on such websites. In Study 1, an automatic analysis method for webpage navigability is introduced, and then a broad survey using this method for 30 international online shopping sites is described. The next study (Study 2) focuses on a fine-grained analysis of real users' behavior on some of these online shopping sites. We modified a voice browser to record each user's actions and the information presented to that user. We conducted user testing on existing sites with this tool. We also developed an analysis and visualization method for the recorded information. The results showed us that users strongly depend on scanning navigation instead of logical navigation. A landmark-oriented navigation model was proposed based on the results. Finally, we discuss future possibilities for improving navigability, including proposals for voice browsers.

Categories and Subject Descriptors: H.3.5 [**Information Storage and Retrieval**]: Online Information Services—*Web-based services*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evolution / methodology, graphical user interfaces (GUI), screen design (e.g., text, graphics, color), standardization, style guides, theory and methods*; H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia—*Navigation, user issues*; K.4.2 [**Computers and Society**]: Social Issues—*Assistive technologies for persons with disabilities*

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: Accessibility, Web accessibility, usability testing, voice browsers, online shopping

ACM Reference Format:

Takagi, H., Saito, S., Fukuda, K., and Asakawa, C. 2007. Analysis of navigability of Web applications for improving blind usability. ACM Trans. Comput.-Hum. Interact. 14, 3, Article 13 (September 2007), 37 pages. DOI = 10.1145/1279700.1279703 http://doi.acm.org/10.1145/1279700.1279703

Authors' address: Tokyo Research Laboratory, IBM Japan, 1623-14 Shimo-Tsuruma, Yamato-shi, Kanagawa-ken, 242-8502, Japan; email: {takagih, shinsa, kentarou, chie}@jp.ibm.com.

1. INTRODUCTION

The Web has become a crucial part of the social infrastructure, not only for sighted "visual" users, but also for blind "nonvisual" users. The population of legally blind people is 179,000 in Japan [Cabinet Office Japan 2004] and 1.3 million in the U.S. [NCHS 1995], and the number is gradually increasing. Needless to say, online services should be usable for this nonvisual population as an important aspect of the modern social infrastructure. Countless online services are delivered through the net, such as news services, public services, and online banking services. Online shopping sites are one category. These sites have brought invaluable benefits to blind people by providing ways to get a broad range of products from their living rooms without leaving home. One important benefit is the scripted detailed explanations of the items offered for sale. It is difficult for blind people to get sufficient explanations in (brick and mortar) shops without human help. Pages on online shopping sites have a wide range of attributes for each item, such as the prices, nutritional data for foods, where perishables were produced, or specifications for home appliances. These explanations allow blind people to effectively choose appropriate items and be wiser shoppers.

However, a massive wave of visually rich online shopping sites is increasingly preventing blind users from purchasing goods on these sites. Severe competition between sites has resulted in various types of promotional tools, such as best-seller rankings, previously viewed items' lists, comments from previous purchasers, seasonal sales, promotional affiliates, personalized recommendations, and so on. These tools are integrated into each page on a site according to visually oriented business logic, which is focused around "screen real estate". Each small part of the screen is seen as having a clear business (real-estate) value based on human visual cognition.

This business-centric page design has been causing severe accessibility issues, especially for ease-of-navigation (navigability) when using voice browsers. Sighted users can use eye movements to scan a newly opened page, and they grasp the structure of the page at a glance. In contrast, voice access users access a page one line at a time by listening to a voice, and use simple navigation commands for scanning a newly opened page. They use commands such as next or previous line, next or previous link, or skip 10 lines. The content is rendered as a one-dimensional string of fragments of the original content in their mental model. The workload for moving back and forth on the string is extremely tedious and time consuming.

We found that the definition of navigation actions in CWW (Cognitive Walkthrough for the Web) theory [Blackmon et al. 2002] gives a good description of the difficulties blind users are facing. CWW itself is a general-purpose usability inspection method for websites, and it aims at discovering usability issues for visual users. In this method, the navigation actions are divided into the following two steps. We inserted the bracketed words into the original description to emphasize the original visual focus of CWW. *"Step one is an* [visual] *attention process that* [visually] *parses a web page into* [visual] *subregions and* [visually] *attends to the* [visual] *subregion of the page that is semantically most similar to* *the user goal. Step two is an action selection process that selects and acts on a widget from the* [visually] *attended-to subregion, the widget semantically most similar to the user goal."*

From this definition, we can see that visual Web navigation heavily depends on visual cognition. Now, let us think of the process of nonvisual navigation. Do blind users parse a webpage into subregions when they access a page? It is not clear but we would say "No" from our experience with IBM Home Page Reader [Asakawa and Itoh 1998]. Users mostly concentrate on a search task seeking their target content in the pages, and they are trying to create a different type of mental model of the page compared to the sighted users' mental models. Do users nonvisually attend to the subregion of the page that is semantically most similar to the user goal? Probably "Yes", in that they often try to figure out the boundaries between the target content and other content. Sometimes, they run over the boundary or fail to find the subregion, but it still seems we can answer "Yes" to this question.

The concern of this research is that we do not have any evidence for the answers to such questions. In particular, we do not have any empirical evidence, and we do not even have any method to investigate the moment-to-moment cognition of blind users for nonvisual Web access processes.

In this article, we will describe our new approach towards obtaining quantitative results for these questions, and investigate the fine-grained behaviors of blind users on the Web while using online shopping sites. We initially chose the following questions for this study:

(1) How inaccessible are current online shopping sites?
(2) How is the regulation for Web accessibility (see Section 3.4) contributing to increase accessible online-shopping sites?
(3) How different are the workloads between the sighted and the blind when shopping using online shopping?
(4) How are blind users accessing the less accessible sites? What leads to failure?
(5) What types of navigation commands are they using? What types of scanning strategies are they deploying?
(6) How much of the listening time is lost for unnecessary content?
(7) Do automatically evaluated "navigable" sites ("logical contents", see Section 2.1) help users realistically? Can heading tags and skip links improve navigability effectively?
(8) How does the use of information scents linked to nodes affect the users' navigation? (See Section 2.1 for "information scent".)

Though this work is of crucial importance to the blind people, these issues will be shared by all users with limitations, such as people whose visual acuity is declining (the elderly) or even people who are temporarily blinded (mobile phone users outdoors on sunny days). We believe the larger significance of this work is what it teaches about designing more usable Web interfaces for everyone.

In Study 1 (Section 3), an automatic analysis method for webpage navigability is introduced, and then a broad survey using this method for 30 international online shopping sites is described. Study 2 (Section 4) focuses on a fine-grained analysis of real users' behavior on some of these online shopping sites. We modified a voice browser to record each user's actions and the information presented to that user. We conducted user testing on existing sites with this tool. We also developed an analysis and visualization method for the recorded information. Finally, we will discuss future possibilities for improving navigability, which includes proposals for site owners, for checking tools, and for voice browsers.

2. ISSUES OF NONVISUAL NAVIGATION AND RELATED WORKS

2.1 Differences between Visual and Nonvisual Navigation

General-purpose Web usability models, such as CWW [Blackmon et al. 2002, 2005], premise that users use vision for their access. They premise that the parse of a webpage into subregions can be done at a glance, and therefore the time for parsing can be ignored while examining Web usability. However, the difference of bandwidth between the senses of vision and hearing (or the sense of touch for Braille reading) makes the parsing task difficult for blind users. Sighted users use their eye-movements for "scanning" a page and can parse a whole page by using everything in the visual field, even including peripheral vision. Blind users are scanning "one line at a time" through a text-to-speech engine (or a one-line Braille output device) by using navigation commands in voice browsers. Goble et al. [2000] explained this scanning (mobility) by using "traveling" metaphor, since a user needs to travel among objects in a page by using navigation commands. It is obvious that full-page parsing is not practical and only partial parsing of a page can by done with partial scanning.

Therefore, the definition of navigation is different for the sighted and the blind. For sighted users, navigation means hypertext navigation by selecting hyperlinks. For blind users, navigation means hypertext navigation and **scanning navigation**,which corresponds to eye-movements for the sighted. Figure 1 shows a proposed mental model of a webpage and examples of scanning navigation. Webpages are recognized as a one-dimensional string of fragments of the original content. One fragment denotes a unit of a navigation command, such as "one line" for line-based command scanning and "one link" for link-based command scanning. In IBM Home Page Reader version 3.04 (HPR), an "item" is the most frequently used unit for navigation. An item is defined as *a small chunk of text that visually causes a hard break new line, such as (1) a paragraph, (2) a heading, (3) a list item in an unordered (bulleted), ordered (numbered), or definition list, (4) a simple table cell, or (5) a control (such as a text box, radio button, or clickable button, etc.)* in the manual.

One of our goals for this study is to discover the various scanning methods, and measure their frequencies. Figure 1(a) shows scanning without navigation to find the title of an article in a news article page. This corresponds to a user just pressing the "play" command at the top of the page and waiting while listening to the content until the target is read aloud. Figure 1(b) shows an

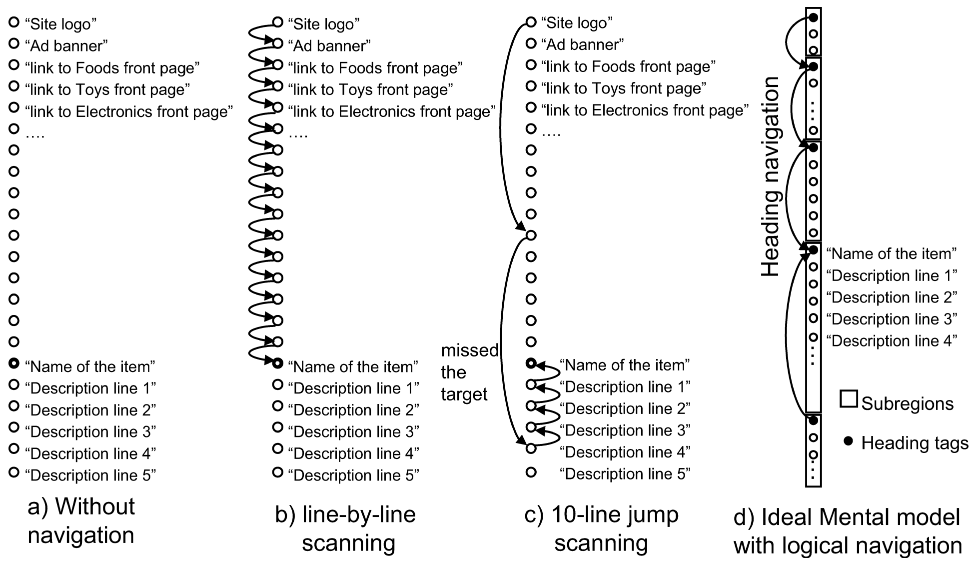


Fig. 1. Mental model of a Webpage and models of scanning navigation.

example of line-by-line scanning navigation to find the title line. In HPR, this corresponds to repeated use of the next-item command (the right arrow key) starting from the top of the page. This scan takes time, but the risk of missing the target is low. The risk is not zero, since each line can be terminated by a next line command before the Text-to-Speech engine reads out it. Figure 1(c) shows an example of gambling by scanning in 10-line jumps. All voice browsers have these gambling commands, such as HPR's "10th following item (page-down key)" command. This scanning has a high probability of missing the target line. Figure 1(d) shows an ideal mental model, which accesses the page in logical way. The target page should have appropriately inserted heading tags, so the (idealized) user can navigate among subregions by using heading navigation commands.

In this article, the term "accessible content" is used for the theoretically navigable content based on this logical navigation model. This model corresponds to the navigation model defined by the HTML specification and the W3C (World Wide Web Consortium) Web Accessibility Initiative (WAI) guidelines [Chisholm et al. 1999, 2000]. In Section 2.3, it is described that the difficulty of defining appropriateness of navigability in literal guidelines.

In Study 2 (Section 4) in this article, we will discuss the effects of "information scents" by applying the "Information foraging theory" [Card et al. 2001; Pirolli 2006] to explain blind users' behavior. This is a theory to understand human behavior on information search tasks, and it emphasizes the similarity to animal food foraging strategies. The most important concept of this theory is the notion of an "information scent", which denotes how human searchers rely on various visual cues in the content to get to their answer, such as link descriptors, images, contextual clues from page arrangements, and so on. For nonvisual access, users might detect a "distinctive scent" for each line during

Table I. Frequently Used Navigational Methods

	IBM Home Page Reader (HPR) v3.04	FreedomScientific Jaws® v7.0	Notes
Heading navigation	<H> and <Shift + H>		
Link navigation	<Tab> and <Shift + Tab>		HPR reads link text in a female voice and Jaws precedes link text with the word ‘link’.
Content (item) navigation	<Left> and <Right>	<Up> and <Down>	
Block navigation	<Ctrl + Enter>	<N> and <Shift + N>	The definition of “blocks” is differed for each voice browser.
Intra-page search	<Ctrl + F>		
Play	<Space>	<Insert + Down>	
Ten-lines jump	<Page Up> and <Page Down>		

scanning navigation, and based on that scent, decide which command to use next. Using Figure 1(c) as an example, the user might jump past the target line, but then notice that the scent is different from the previous content, so the user would then know to go back to the title line. One difference between the senses of smell and sight is the granularity of the information. In Card [2001], a particular scent was assigned to each page, but the unit should be each line or item for nonvisual navigation because of the size of the navigation units. One of our experimental goals was to study the effects of the information scent of each line.

2.2 Assistive Technology Support for Navigability

Current voice browsers and screen readers provide a variety of navigation commands. Table I shows several frequently used navigational methods with two commercially available voice browsers, IBM Home Page Reader (HPR) v3.04 and FreedomScientific Inc. Jaws® v7.0. Each browser provides both basic navigation commands (heading navigation, link navigation, content navigation, etc.) and unique navigation commands (block navigation, 10-line jump, etc.). However, the usability of these commands has not been sufficiently evaluated. Theofanos and others have reported that blind users don’t use most of these commands [Theofanos and Redish 2003]. Their results showed that average users only used a small set of the functions of their screen reader, since the mental load of using a browser, making a mental model of the website, and working with assistive devices or programs was already very high compared to the mental load for sighted users.

In our new study (Study 2), we conducted usability testing sessions and recoded various data related to users’ operations, such as users’ navigation operations, and the text strings sent to the Text-to-Speech engine. Salampasis et al. [2005] reported that they conducted usability testing sessions for their enhanced browser for the blind by collecting keystroke logs, but they did not report on the utility of their browser in practical environments, nor provide any detailed results. Barnicle [2000] reported the results of usability testing for GUI applications and a Web browser. They recorded audio and video, but no key operations. We could not find any research that analyzed the details of users' operations and voice browser behaviors.

2.3 Navigability Considerations in Guidelines and Automatic Checking Tools

The concept of navigability is considered in Web accessibility guidelines. The W3C's Web Content Accessibility Guidelines (WCAG) is the standard guideline for Web accessibility [Chisholm et al. 1999]. In this guideline (and in the HTML specifications), the Heading tags (H1, H2, H3, ..., H6) are defined to support easy navigation. This was described as an ideal logical navigation model in Section 2.1. CSS [Jacobs and Brewer 1999] allows developers and designers to control the visual appearance of pages while preserving heading tag structures.

However, the appropriateness and practical effectiveness of heading tags is barely defined as a part of the guideline items. The WCAG mentions appropriateness as regards the nesting order of the heading tags, where it notes that heading tags "should not skip levels (e.g., H1 directly to H3)" [Chisholm et al. 2000]. In the U.S. Section 508 of the Rehabilitation Act,[1] Item (o) is the main reference to navigability; "*(o) a method shall be provided that permits users to skip repetitive navigation links.*" This item is limited to encouraging the insertion of skip links for a set of navigation links.

Automatic checkers were developed to support these guidelines, so each checker has some specific functions, such as warning about a "skip of heading levels". However, the appropriateness of the navigability is essentially hard to assess automatically, considering the appropriateness of the chapter to subtopics to sections structure as described in WCAG 1.0. Visualization is a new way of ensuring navigability [Takagi et al. 2004]. The method is introduced in Section 3.3 of this article.

3. STUDY 1: EVALUATION OF ONLINE SHOPPING SITES' NAVIGABILITY

3.1 Overview

Study 1 aimed at discovering the current status of accessibility for online shopping sites. We first discuss characteristics of and difficulties with online shopping sites, and then introduce our automatic navigability analysis method. Then evaluation results for 30 online shopping sites around the world are presented.

3.2 Characteristics of Online Shopping Sites

In recent years, we have had meetings with the site owners and developers of many online shopping sites. They frequently noted that it is difficult for them to

[1]See `http://www.section508.gov/`.

make their sites accessible, even if they have a strong interest in accessibility. We classified the reasons into the following three areas.

(1) *Variety of Small Promotional "Tools" on a Page*. Fierce competition among sites has resulted in various types of promotional tools, such as best seller rankings, personalized recommendations, comments from buyers, price-sensitive search, buyers' rating-sensitive searches, and so on. These tools are combined into one page and the tools are routinely optimized for visual users to easily recognize their functions. It is sometimes difficult for nonvisual users to understand these tools.

    Taking Amazon.com™ as a widely known example, we can see various types of promotional tools within an item page, such as a link to other books by the same author, SIP recommendations (using distinctive phrases in the book for the Search Inside™ Program), CAPs (capitalized phrases: people, places, events, or important topics mentioned frequently in the original book), a link to customer reviews and ratings, links to other editions and used copies, and so on.

    These tools are visually well organized. For example, SIPs and CAPs use a smaller font size, the sale price of this item is red and bold, and the other edition information is surrounded by thin sky blue line. However, without vision all of these visual cues are lost. In this case, the nonvisual target content would probably be the price, some of the details, and the add-to-cart buttons. This essential content is buried among the various promotional tools.

(2) *Strong Domination of Visually Oriented Business Logic, in the Form of "Screen Real Estate" Thinking*. Each small part of the screen has a clear business value (as screen real estate) based on human visual cognition. Therefore, various tools are integrated into one page and positioned according to their business value in order to maximize revenue. This design strategy clearly increases the complexity of the pages. In the case of Amazon.com™, the SIPs and CAPs are actually explained at the bottom of the page, but the lower part of the page has a lower value as real estate, and therefore only a small bit of labeling information appears just below the title of the book in the high value area. If a blind visitor constructs a linear model (as in Figure 1) it is not at all obvious how to find such details as more information about this book, but a sighted visitor will easily be guided to links that are intended to sell other books.

(3) *Rapid Changes (Update or Replacement or Removal) of Items*. Items are continuously added to such sites. Therefore, new pages with new layouts and new promotions are continuously developed and published. If the pages are static and the layouts are not stable, it is not possible to renovate all of the pages on such a site to make it accessible. The replacement speed of pages is too fast for after-the-fact accessibility efforts to catch up. In one case, we found a site where the accessibility had been improved at one time, but the accessibility was degraded after a year in spite of their large investment.

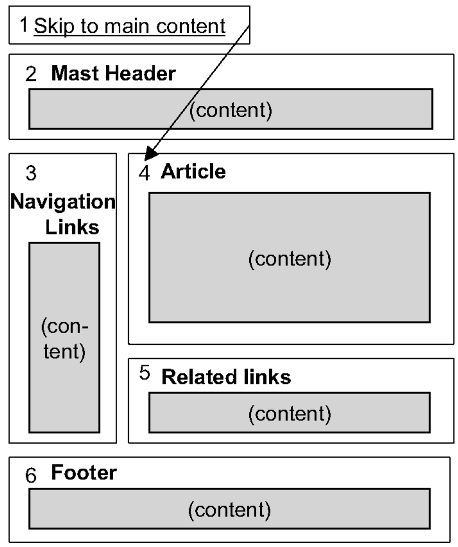


Fig. 2. Example of heading tags and intra-page links.

3.3 Nonvisual Navigability Analysis and Visualization

We have developed an automatic analysis and visualization method for nonvisual navigation, called "Nonvisual Usability Visualization Method" [Takagi et al. 2004]. We used this method for measuring the navigability of the pages on online shopping sites in this study.

This is a method to analyze the navigability for each part of a page based on graph structures defined by the existence of heading tags and skip-links (intra-page links), which then calculates the reaching time to each part of a page when using voice browsers. In this application, the "reaching time" is a metric to express the navigability for page authors with little understanding of accessibility. As discussed in Section 2.3, navigability is basically ignored by existing checking tools and rarely reported to page authors. This method aims at providing a method to evaluate navigability and other issues at a glance.

The analysis is based on two preconditions: (1) the virtual user uses heading navigation keys and the available intra-page skip links effectively. This condition allows us to estimate the reaching time as the shortest path to each element. (2) The virtual user accesses all of the text information contained in each accessed line. This condition corresponds to a user who always uses the "play" button for checking the content. In other word, this is a very conservative user, who is afraid of skipping over content by using the navigation keys, since the user always misses some of the content when using such navigation commands. These two preconditions allow us to estimate navigability as the simple metric of "reaching time".

Figure 2 shows an example of heading tags and intra-page links. There are six areas, with five heading tags and one intra-page link. The calculation of reaching time can be regarded as a shortest path problem for a weighted directed graph. Figure 3 shows a graph representation for Figure 2. The nodes

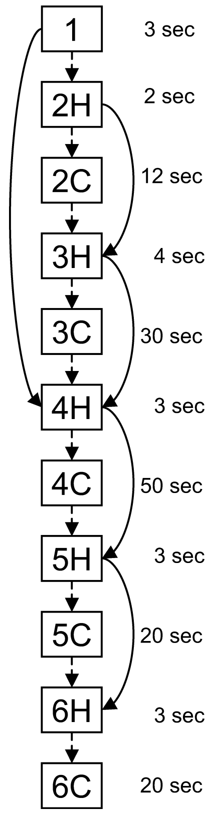
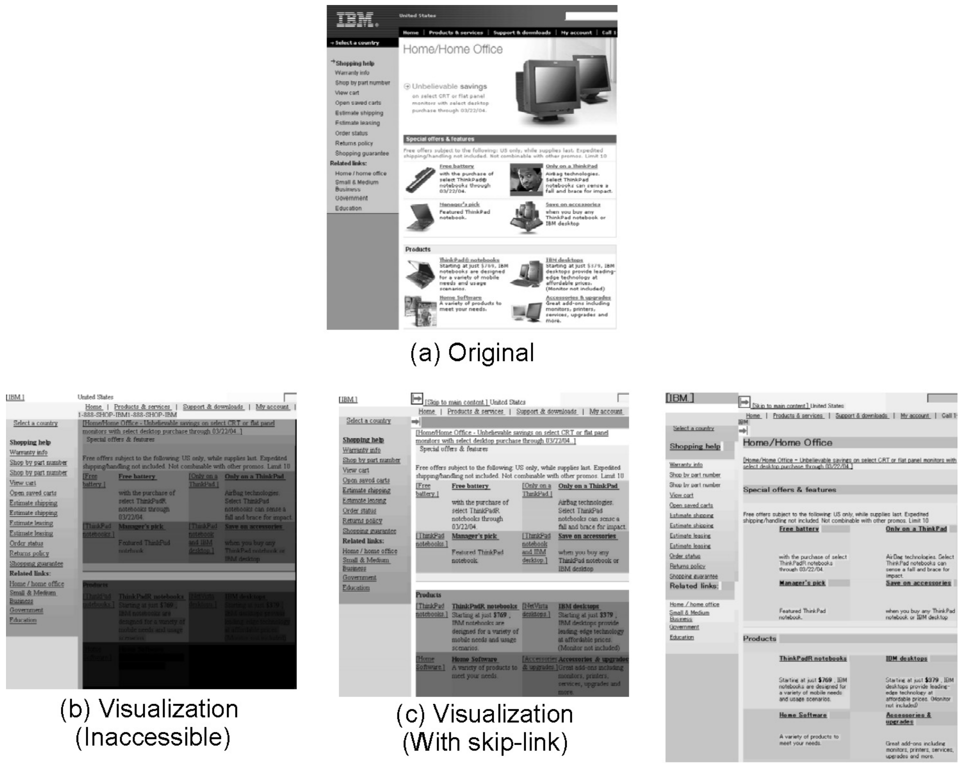


Fig. 3. Graph representation for Fig. 2.

are the indicated areas, the numbers are the area identification numbers, the dotted lines show the tag order, and the solid lines are the related headings and links. Each node has a weight representing the time required to read through the corresponding area. The duration is estimated from the number of words and the length of each word in the area. For example, the shortest path to get to Area 6 is via 4H, 5H, and then 6. This will take a minimum of 9 seconds $(3 + 3 + 3)$.

Figure 4 shows an example of visualization. The gradations of the background color show the times to reach each element in a page with a standard voice browser, white for 0 seconds and black for 120 seconds or longer, beginning from the top of the page. You can see the consideration given to navigation even from these small thumbnail images. Figure 4(b), 4(c), and 4(d) have just the same appearance (Figure 4(a)) when rendered by standard browsers (such as Internet Explorer), but they are significantly different from the blind usability point of view when rendered by standard voice browsers. Figure 4(b) shows an inaccessible page, since there are no heading tags or any intra-page links, such as the skip-link and the page index links. The main content area (here including a title and a large image that was converted to an ALT text in the blind views) is shaded as black, showing it will take a long time to arrive at this area.



(a) Original

(b) Visualization (Inaccessible)

(c) Visualization (With skip-link)

(d) Visualization (With headings)

Fig. 4. Examples of aDesigner's reaching-time visualization. (excerpt from Takagi et al. [2004]).

Figure 4(c) shows an improved page with a skip-link. The main content area is shaded with a light gray color, showing how blind users can now access the main content easily. In this way, by referring to the differences in these colors, Web designers can easily recognize how short or long is the time blind users need to reach the main content. The darker colors indicate longer times, while lighter colors indicate shorter times. The insertion of a "skip-to-main-content" link, however, only helps in accessing the main content area. In Figure 4(d), there are several headings followed by brighter colored areas. Standard voice browsers provide functions to jump directly to each heading tag in a page, so this allows blind users to navigate through each heading quickly and it helps them to grasp the page overview effectively.

3.4 Sites

We selected 10 major online shopping sites for each country, the U.S., Japan, and the U.K. In total, we evaluated 30 sites. For each country, four of the sites (A, B, C, and D) were localized versions of global sites, and the other six sites were domestic sites, which were selected to represent six categories: an online mall, an online version of a department store, a specialty site for toys, one for books and CDs, one for electronic goods, and one for stationary goods. We selected the U.K. sites in order to measure the effectiveness of the new U.K. Disability Discrimination Act (U.K. DDA). This regulation requires that all public online services should follow the Web Content Accessibility Guidelines 1.0 [Chisholm et al. 1999]. Therefore, we hoped that the current efforts to comply with the U.K. DDA would be clearly shown in the results.

3.5 Measured Criteria

We measured three criteria, (1) the reaching times to result lines in the search result pages, (2) the ratios of pages with heading tags, and (3) the ratios of pages with skip-links. The search result pages were selected as comprehensive evaluation targets for all of the shopping sites, since all such sites have search functions for their merchandise. For search result pages, it is obvious that the result lines are the main content. The reaching time was measured by Nonvisual Navigability Analysis method (Section 3.3), which is implemented in the aDesigner's visualization function.[2] The experimenter opened the top page of a target site and searched for one of the selected keywords, and then measured the reaching time by checking the visualization result. Three keywords were selected for each site category (e.g., tea, a handbag, and a suit for department store sites). The ratios of pages with heading tags and skip links were measured by using the Site-wide aDesigner, which has a function to automatically crawl the pages on a site and to generate various statistical data for the target site. For this study, pages within one click of the top page were examined. For the skip-link ratio, only complete and valid skip-links were counted and broken skip links (e.g., a skip-link without a target) were not counted.

3.6 Results

3.6.1 *Reaching Time to Result Lines in Search Result Pages.* Table II shows the results of the reaching time analysis. There are no results for site J (in the US), since it uses an internal frame technique for the results list and aDesigner cannot analyze it. We classified the sites into three categories according to the reaching times, 0–20 seconds, 20–40 seconds, and more than 40 seconds. This is only an arbitrary reference categorization to examine the results based on our experiences with users.

In Table II, only four sites qualified under the shorter threshold (20 seconds). Three of these sites were U.K. sites, one was in Japan, and there were none in our U.S. sample. Three of these sites marked the item names in the result lists as heading tags, so voice browser users can easily navigate to them by using the heading navigation functions. The other site has a skip link to its result list, but the destination point is not exactly on the result list and users need to skip some content to get to the target. The site was still under 20 seconds, but it would be even shorter if the destination position were corrected.

Four more sites were classified in the group from 20–40 seconds, one from Japan, two from the U.K., and one from the U.S. They all used heading tags

[2] See `http://www.alphaworks.ibm.com/tech/aDesigner/`.

Table II. Results of Reaching Time Analysis for Product Search Result Pages

		Item1 [sec]	Item2 [sec]	Item3 [sec]	**Average [sec]**	Distribution
US	A	37	37	21	**31.7**	5.3
	B	70	69	68	**69.0**	0.5
	C	100	112	103	**105.0**	3.5
	D	157	111	106	**124.7**	16.2
	E	102	102	102	**102.0**	0.0
	F	313	233	255	**267.0**	23.0
	G	54	56	52	**54.0**	1.0
	H	58	49	54	**53.7**	2.3
	I	84	74	75	**77.7**	3.2
	J					
UK	A	34	31	32	**32.3**	0.8
	B	49	47	47	**47.7**	0.7
	C	29	27	27	**27.7**	0.7
	D	106	92	82	**93.3**	6.3
	K	42	42	42	**42.0**	0.0
	L	24	22	15	**20.3**	2.7
	M	5	5	5	**5.0**	0.0
	N	76	76	76	**76.0**	0.0
	O	5	5	5	**5.0**	0.0
	P	139	144	140	**141.0**	1.5
JP	A	215	212	237	**221.3**	7.8
	B	129	141	137	**135.7**	3.3
	C	11	11	11	**11.0**	0.0
	D	143	108	97	**116.0**	13.5
	Q	300	304	300	**301.3**	1.3
	R	35	26	24	**28.3**	3.3
	S	168	193	168	**176.3**	8.3
	T	127	122	130	**126.3**	2.2
	U	547	545	553	**548.3**	2.3
	V	119	117	120	**118.7**	0.8

or skip-links in their pages, which is why the reaching time did not exceed the 40-second criterion. Some sites added heading tags only for the header (not the headings) and the footer. This means the main content does not have a heading tag. There are two likely explanations for this status, one being that their accessibility work is ongoing (and currently only halfway finished), and the other being that they have done some makeshift repairs in response to complaints from automatic checking tools. It is difficult to determine which inference is correct for these cases, but it is clear that their sites should be improved further.

The remaining 21 sites were over the minimum of 40 seconds threshold. Most of them did not pay any attention to navigability for voice access users on their sites. This result is not surprising and fits well with our expectations. In the U.K., five sites exceeded the threshold, but those sites still had relatively short reaching times compared with U.S. and Japanese sites. The national average for the U.K. of 49 seconds is close to the threshold and this result should be regarded as evidence of the effectiveness of the British law. For the U.S., only one site passed, and the national average was much longer than for the U.K., 98 seconds. For Japan, only two sites passed and the national average was 178 seconds.

Table III. Ratios of Pages with Skip Links and Heading Tags

			Skip Link		Heading Tags	
		Total # of page	# of pages with skip links	Ratio	# of pages with heading tags	Ratio
JP	A	142	0	0.0%	53	37.3%
	B	207	3	1.4%	35	16.9%
	C	12	0	0.0%	8	66.7%
	D	158	0	0.0%	4	2.5%
	Q	183	1	0.5%	6	3.3%
	R	89	85	95.5%	72	80.9%
	S	77	0	0.0%	0	0.0%
	T	136	0	0.0%	25	18.4%
	U	300	0	0.0%	0	0.0%
	V	93	0	0.0%	0	0.0%
	Average	139.7		9.8%		22.6%
UK	A	59	0	0.0%	59	100.0%
	B	64	0	0.0%	59	92.2%
	C	39	39	100.0%	39	100.0%
	D	196	0	0.0%	1	0.5%
	K	37	0	0.0%	30	81.1%
	L	80	0	0.0%	80	100.0%
	M	63	61	96.8%	47	74.6%
	N	42	0	0.0%	41	97.6%
	O	133	0	0.0%	64	48.1%
	P	60	0	0.0%	3	5.0%
	Average	77.3		19.7%		69.9%
US	A	190	0	0.0%	162	85.3%
	B	149	3	2.0%	144	96.6%
	C	343	0	0.0%	2	0.6%
	D	257	0	0.0%	7	2.7%
	E	56	0	0.0%	4	7.1%
	F	36	0	0.0%	0	0.0%
	G	34	0	0.0%	20	58.8%
	H	141	140	99.3%	0	0.0%
	I	85	0	0.0%	0	0.0%
	J	228	0	0.0%	4	1.8%
	Average	151.9		10.1%		25.3%
Total Average		123.0		13.2%		39.3%

3.6.2 *Ratio of Pages with Heading Tags and Skip Links.* Table III shows the results of the ratios of pages with heading tags and skip-links. Each ratio was calculated from the number of checked pages and detected pages with heading tags or skip-links. The number of checked pages on each site was 123 pages on average.

It was better than we expected to find that all pages had heading tags for three sites (A, C, and L) and more than 90% for two sites (B and N), all in the U.K. The quality should be discussed, but this result is evidence of the ongoing and massive effort in the U.K. In the U.S., one site had 96.6% and another had 85.3% with heading tags. These two sites are global suppliers, so they might share their U.K. experience with their U.S. site designers and developers. In Japan, one site had 80.9% with heading tags. This site is very active for accessibility enablement, so this result shows the ongoing efforts.

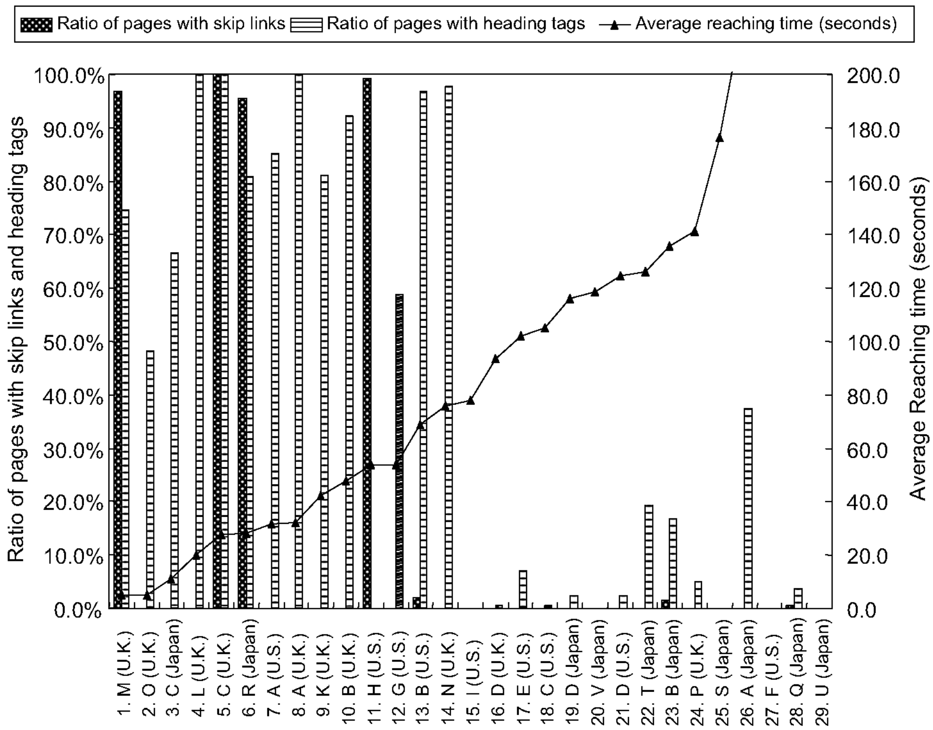


Fig. 5. Rankings for average reaching times and comparisons with skip links/heading tags.

Four sites had skip-links. Skip-links can be used to compensate for a lack of heading tags. One site (H) is using skip links for this purpose. There are no heading tags on the site, but 99% of the pages had skip-links. However, the average reaching time on Site H (Table II) was 53.7 seconds, so unfortunately it seems the skip-links are not placed with maximum effectiveness in the search result pages. Other sites had both skip links and heading tags on their pages.

3.7 Discussion

3.7.1 *Correlation between Reaching Times and Ratios of Heading Tags/ Skip Links.* In order to show the correlation between reaching times and the ratios of heading tags and skip links, we combined these results into one chart and sorted the results by increasing reaching times (Figure 5).

It is obvious that the highly ranked sites had heading tags, skip links, or both. Most of the highly ranked sites with ranks from 1 to 14 had more than 80% of their pages with heading tags or skip links, showing that this tendency correlates with low reaching times. Sites O, C_{Japan}, and G are exceptions. In their search result pages they added tags effectively, instead of inserting tags into all pages. Sites $B_{U.S.}$ and N show sharp contrasts with these three sites (O, C_{Japan}, and G). Sites $B_{U.S.}$ and N had good ratios for heading tags, but reaching times were long. For Site $B_{U.S.}$, the results page has many heading tags and most of them are appropriate. However, the "list of sellers (shops who sell the items)" was inserted before the result line and after heading tag, so the reaching time was longer than it needed to be. Such problems are caused by their business decision to feature the sellers along with the search results. Site N shows another severe misunderstanding of navigability. They only added heading tags to the site-navigation list on the left side of each page. Therefore, users will encounter too many heading tags but not find the main content (the search results) by using heading navigation. These inappropriate uses of heading tags and skip links are scarcely visible to sighted people, but they are made clear by using our navigability analysis method.

3.7.2 *Differences among Localized Sites of Global Companies and the Effects of Local Regulations.* In the results, there were significant differences between the U.K. sites and the other countries' sites. The reaching time was significantly different between the U.K. sites and the U.S. sites (t-test, 1%, $p < 0.01$), and between the U.K. sites and the Japanese sites (t-test, 1%, $p < 0.01$). We can see this tendency not only for domestic sites, but also for global sites. Sites A, B, C, and D were localized sites owned by global online shopping sites. It is interesting that each localized site has a different level of navigability compared with other localized versions of that site. For the A sites, the U.S. and U.K. sites had similar reaching time values (32.3 and 31.7 seconds), but there was a much longer value (221.3 seconds) for the Japanese site. We believe that the U.K. site was renovated to comply with the U.K. DDA and then the experience was shared with the U.S. site designers and developers. However, the Japanese designers and developers were isolated for some reason. The type B sites show the same traits. The type C cites show another pattern, with Japan and the U.K. sites having short reaching time values (11.0 and 27.7 seconds), but much longer (105.0 seconds) for the U.S. site. In this case, it also seems the U.K. site might play an important role. The type D sites are an exception, since there are no localized sites with shorter reaching times and only a few pages had heading tags or skip links. Worldwide, this company is clearly paying less attention to nonvisual access.

From these results, we feel that the U.K. sites are becoming testbeds for building voice-accessible sites. Sites in the U.K. still have many issues and the level of accessibility varies among sites, but the regulation is driving their sites towards being accessible online shopping sites.

4. STUDY 2: EXPERIMENT ON VOICE-ACCESS BEHAVIOR IN THE SHOPPING PROCESS

4.1 Overview

The results in the previous section showed us the poor status of navigability of online shopping sites. In contrast to the previous study, this second study aims at analyzing detailed data on real users' behaviors during the shopping process on real online shopping sites, and at obtaining quantitative evidence to answer the questions numbered from 3 to 7, as listed in Section 1.

This study consists of two sub-studies. Experiment 1 (Section 4.3) aimed at examining the entire process of realistic online shopping, and comparing

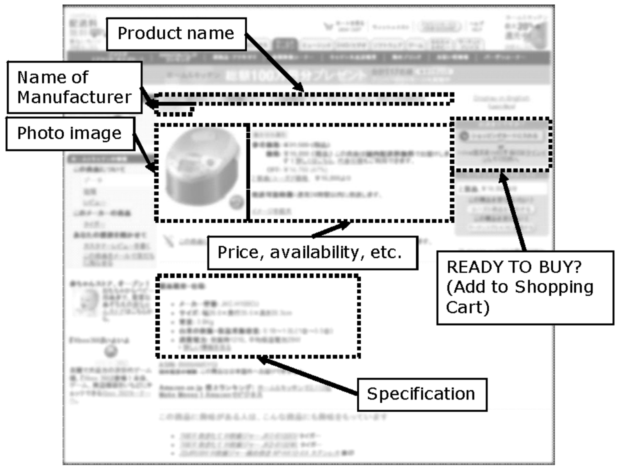
Table IV. Types of Automatically Logged Data

Event Type	Description	Example
CreateController	New window created.	CreateController:
DestroyController	Window was closed.	DestroyController:
EnterMenu	Enter menu bar by ALT key, etc.	EnterMenu:
FindCancel	Cancel Find dialog.	FindCancel:
FindOK	Select OK on Find dialog. Parameters are search keyword, direction, and case sensitivity.	FindOk: "Sonata of Winter" FindNext IgnoreCase
FindOpen	Find dialog opened.	FindOpen: "Carpet"
HighLight	Highlighted text on HPR window. It corresponds to text sent to Text-to-Speech engine.	Highlight: "Add to cart"
KeyDown	Key down event.	KeyDown: VirtKey=0×24 KeyData=0×147
KeyEvent	Key release event.	KeyEvent: wParam=0×24 lParam=0×1470001
Navigate	Navigation operation. Parameters are direction and unit of movement.	Navigate: direction=FIRST unit=ITEM
NewDocument	New page loaded.	NewDocument: http://www.rakuten.co.jp/
SetPOR	XPath notation of focused element.	SetPOR: /HTML[1]/BODY[1]/A[1]/text()[1]
SpeakText	HPR messages.	SpeakText: "Now connecting. Please wait..."
SpeechMarker	Behavior of Text-to-speech engine.	SpeechMarker: markType=3 markOffset=0
SpeechRate	Change of speech rate. Unit is WPM (Words per minute).	SpeechRate: 380WPM
TextEntryCancel	Cancel text input dialog.	TextEntryCancel:
TextEntryOK	Select OK on text input dialog	TextEntryOK: "23 2nd season"
TextEntryOpen	Open text input dialog.	TextEntryOpen: "DVD"

task completion times between sighted and blind shoppers. Experiment 2 (Section 4.4) aimed at comparing users' behavior on the original pages and accessibility-enhanced pages, which have logical content structures as described in Section 2.1.

4.2 Logging Home Page Reader (LogHPR) and Voice Access Behavior Diagram (VABD)

In order to measure and to analyze the fine-grained behaviors, we developed three tools and analysis methods. For measuring user's behavior, we modified the commercial voice browser IBM Home Page Reader (version 3.04) to automatically record the used navigation commands, menu operations, the text strings sent to the Text-to-Speech engine, and so on. This modified version is called the Logging Home Page Reader (LogHPR). Table IV shows the types of logged data. This is the first tool that allows observers to track this level of fine-grained behavior. The response-time latency for logging was not detectable by the users, and no subject reported noticing any latency.



a) Defined areas for main content in this page

```
// READY TO BUY?
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM[2]/TABLE[1]/TBODY[1]/TR[1]

// name
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM/B[1]

// manufacturer
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM/A[1]

// image
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM/TABLE[2]/TBODY[1]/TR[1]/TD[1]/A[1]/IMG[1]

// price and other information
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM/FONT[1]/TABLE[1]
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM/FONT[1]/SPAN[1]
/HTML[1]/BODY[1]/P[1]/TABLE[1]/TBODY[1]/TR[1]/TD[3]/TABLE[1]/TBODY[1]/TR[1]/TD[1]/FORM/FONT[1]/FONT[1]
```

b) XPath expressions for main content areas

Fig. 6. Example of defined main content area and XPath expression.

We also developed an analysis tool to calculate durations and time intervals related to the main content from the logged data and the XPath [Clark and Derose 1999] definition of the main content. For example, the reaching time to the main content area is the time from the opening of a page until reaching the main content area. This time interval shows the time required for the initial scanning task of searching for the content. These time intervals were analyzed by defining the content areas in the DOM for each page manually using XPath. Figure 6 shows an example of a main content definition. For search result pages, the search result lines were defined as the main contents, and product definitions were defined as the main content for product pages. In Figure 6(a), we defined the following elements as the main content areas of this page: name of the product, main image of the product, manufacturer, price, availability of the product, and "READY TO BUY?" (including the "Add to cart" button). The corresponding XPaths for each element are shown in Figure 6(b). For a search results page with no results, only the sentence reporting the lack of results was considered to be the main content.

In order to analyze users' behaviors, we also developed a visualization method, the Voice-Access Behavior Diagram (VABD). Figure 7 shows an example of a VABD diagram, which is a part of the logged data in Experiment 1 (see Section 4.3). VABD can be rendered as a sparse table structure. The

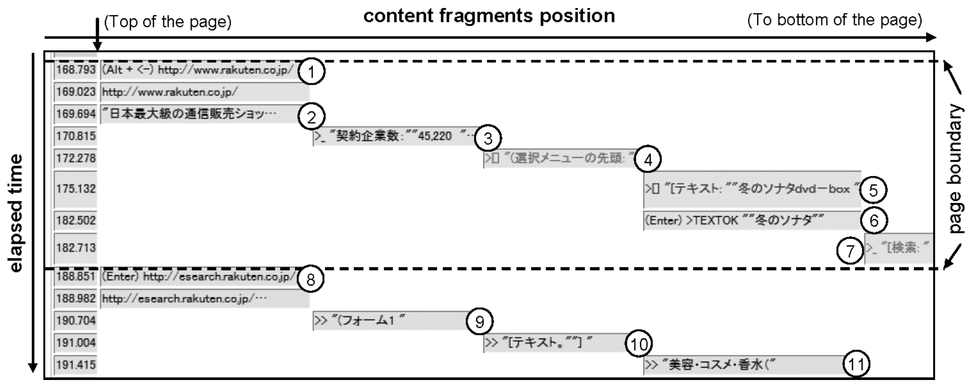


Fig. 7. Example of Voice Access Behavior Diagram (VABD).

different background colors of the cells show the pages. In this example, the cells before 1 are light blue, those between 1 and 7 are pink, and those between 8 and 11 are gray. (We refer each cell by the number, from 1 to 11.) As we mentioned in Section 2.1, webpages can be modeled as a one-dimensional sequence of short content fragments. This diagram tries to represent the user's movements as a sequence. The horizontal axis on the diagram shows the sequence of fragments. Each fragment is mapped to a column in the diagram, which corresponds to an "item" used in HPR (see Section 2.1). Empty columns (content fragments that a user did not visit) are not shown in Figure 7 to make the diagram concise. The vertical axis shows the elapsed time from the beginning of the experiment. The height of each row is proportional to the intervals between fragments.

The VABD shows when and how the user reached each content fragment. For example, Cell 1 shows a user navigating to a page (www.rakuten.co.jp) at 168.793 seconds (after the experiment started). The string "`Alt+<-`" means the user reached the page by pressing the ALT and Left arrow keys, which is the previous page key in HPR (and many other browsers). Then HPR started to read the content of the page at 169.694 (Cell 2). In Cell 3, the user pressed the Tab key (next link), denoted by ">_", and HPR read that content. After that, the user pressed 'o' (the next form key), and reached the top of the list box (Cell 4). In Cell 5, the user pressed 'o' again and reached the text box. This cell says "Text: fuyu no sonata dvd-box", where "Text:" is a message generated by LogHPR about the user a text box, and the rest is the content already input (the user had input that content the last time he visited the page before returning to the page). The user modified the content and hit the <Enter> key to confirm his input, denoted by "`>TEXTOK`" in Cell 6. The user tabbed and reached the "Search" button (Cell 7), then <Enter> to follow the link and navigate to the next page at 188.851 (Cell 8). We also defined symbols for other navigation commands.

We use foreground colors to express the characteristics of each fragment. The fragments with predefined "scent" words are colored blue, since these words may be recognized as fragments with information scents. For instance Cell 5

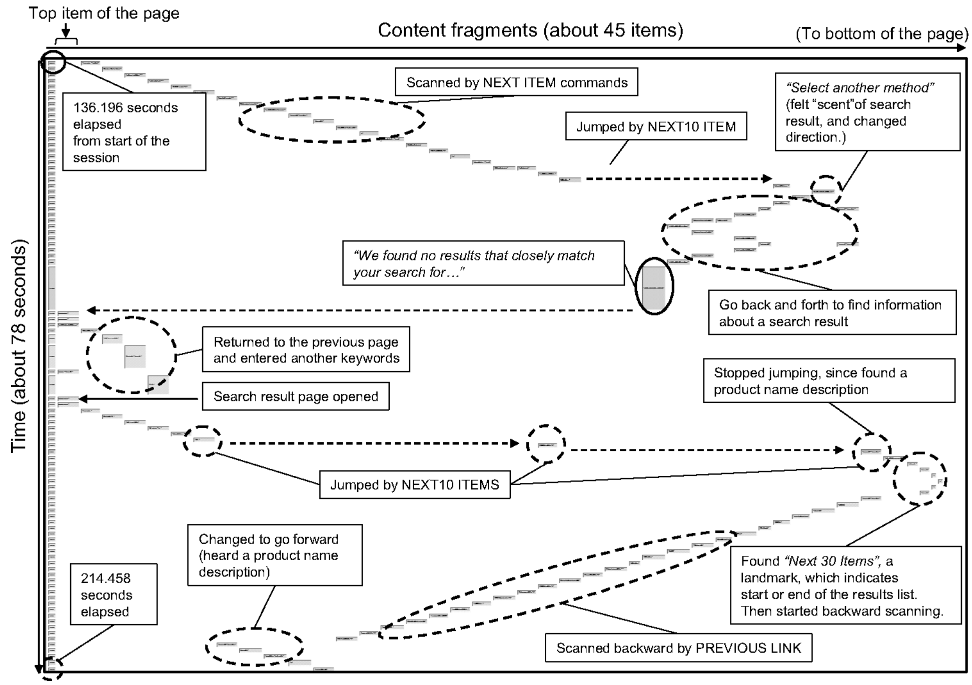


Fig. 8. Larger example of a Voice Access Behavior Diagram (VABD).

has a blue foreground, since the fragment is a part of the target string "DVD". The fragments with red foregrounds show the user is in the predefined main content of a page. Figure 8 shows a larger annotated example of VABD from Experiment 2. Other examples are presented in the next section.

4.3 Experiment 1: Comparison of Shopping Processes between Sighted and Blind

This experiment examined the entire process of online shopping, and compared the task completion times between sighted and blind shoppers.

4.3.1 *Tasks.* The task was designed to be as realistic as possible. Subjects were simply instructed to buy a DVD for a popular (Korean) TV program, "Fuyu no Sonata (Winter Sonata) DVD Volume 1". The task started when the test subject opened the top page of the shopping site, and finished when the subject selected the add-to-cart button. (See Appendix A for the detailed instructions.)

We conducted three experiments, with each experiment consisting of two sessions. One session was for Amazon.com, and the other was for the Rakuten online shopping site. Amazon.com was selected as a representative global shopping site, and Rakuten was selected as a representative domestic (Japanese) online shopping site. Both sites are among the largest online shopping sites in the world. Rakuten has more than 16 million items (as of April 2006). Amazon.com does not release the number of items, but it is well known as one of the largest. The subjects were divided into two groups. The first group

Table V. Participating Blind Subjects

Subject ID	Age Range	Job Type	PC Experience (years)	HPR	Online Shopping
P1	30–39	University research staff	15	OK	every month
P2	20–29	University student	7–8	OK	none
P3	30–39	Masseur	2-1/2	OK	none
P4	40–49	Accessibility tester	15	OK	twice a week
P5	30–39	University research staff	13	NO	twice a week

started with Amazon while the other started with Rakuten for Experiments 1 to 3. An experimental description was provided in HTML format for reference during each experiment.

4.3.2 *Procedures.* The experiments consisted of three phases, an interview session, a training session and experimental sessions, taking about 4 hours for each subject.

(1) *Interview*. We asked about 32 items to compile individual profiles, including their Web access experience, computer environment, and online shopping experiences. These results were referred to for analysis, and some of this information is shown in Table V.
(2) *Training Session*. Each subject was allowed to use the tested voice browser (LogHPR) freely for about 30 minutes. The training consisted of two phases: the first phase was for learning about the new features of the tested browser and the second phase was for navigating through the two online shopping sites used during the experiment. The regularly used navigation commands were reviewed to ensure that all of the subjects were aware of the necessary and useful navigational commands. During the online shopping navigation practice, they were allowed to ask questions. Some of subjects did not have any experience with online shopping. Such novice subjects were told that there is a search form in the page to look for a product. They were not given more specific information, such as the layout of the pages or effective ways to find products.
(3) *Experiments*. Experiments were performed by the subjects. The maximum time for each task was limited to fifteen minutes. If a subject could not perform a task during the fifteen-minute period, it was recorded as a failure. After each task, participants' protocol steps were coded by replaying the recorded sounds of their operations while performing the task. Generally speaking, a post-session protocol is not appropriate to obtain detailed process explanation. However, the talking-aloud method could not be used, since the verbal explanations might affect their minute-to-minute task processing when using the sense of hearing. Therefore, we decided to use the post-session protocol method with reviews of their task steps.

    An additional experiment for sighted participants was conducted beyond the scope of this basic procedure. This experiment consists of a short interview session and one experimental session without any training session.

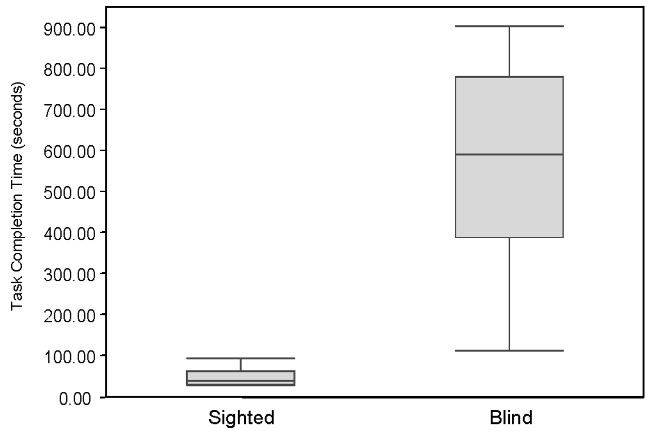


Fig. 9. Comparison of completion time length between sighted and blind.

The participants were given exactly the same instructions as the blind participants (Appendix A). We did use the think-aloud method for sighted participants only, and recorded the protocol by using a laptop computer.

4.3.3 *Instrumentation.* Two laptop computers were used. One was used for the main computer with the voice browser (LogHPR) used by each subject to perform the experiments. The other computer was used for recording the subject's voice and the TTS output. A video tape recorder was used to record the entire experimental session, in particular including screen shots and the hands of the subjects. For the sighted experiment, only Internet Explorer® was used.

4.3.4 *Participants* . Five blind subjects participated in this experiment. Table V summarizes data about the blind subjects referred to during the analysis of the experimental results. They are in their 20s to 40s. One is a student, three are employed university graduates, and the other graduated from a school for the blind and works as a masseur. Four of them had more than one year of experience using HPR, since we wanted HPR users who have sufficient skill to shop on online-shopping sites. HPR is the most common voice browser among blind internet users in Japan [Watanabe et al. 2002], so they are not exceptionally skilled subjects. None of them had any hearing impairments, based on their own reports.

The number of subjects is relatively small, but most experiments that are conducted using blind subjects usually have only four to six subjects [Stevens 1996; Challis and Edwards 2000]. It is difficult to obtain more blind subjects because of mobility problems and skill and employment qualifications.

Five sighted subjects also participated. All of them are researchers at our laboratory, from 28 to 45 years of age. Two of them had never tried online shopping before, even though they are advanced Internet users.

4.3.5 *Results and Analysis.* Figure 9 shows the box plot for task completion time for blind participants and sighted participants. It is clear that there are significant differences between the sighted and blind (t-test, $t = 6.28$, $p = 0.0001$). The average completion time was 46.3 seconds for sighted and 563.6 seconds for blind, so the blind participants were more than 10 times slower than the sighted participants.

All of the sighted subjects reported that there were many search results if they only searched for the DVD using part of the title, "Fuyuno Sonata" in the first task. They changed their search strategy by typing two or more words as search strings or by selecting the category for DVD. This is one example of the effectiveness of visual parsing of a page at a glance. They grasped the page layout and noticed that there were many things there. Then they noticed the text stating the total number of result items. The average number of the mouse clicks was 5.2, with a maximum of 7 and a minimum of 3. Scrolling operations were excluded from these numbers.

4.4 Experiment 2: Comparison of Original Pages and Accessibility-Enhanced Pages

This experiment aimed at comparing users' behavior on original pages and accessibility-enhanced pages, with logical content structures as described in Section 2.1.

4.4.1 *Tasks and Procedures.* These two experiments were done only by the blind participants, Exp 2-1 for the original pages and Exp 2-2 for the accessibility-enhanced pages. A search results page, which was the result of a search on the top page of a site, was given to each subject, and the subject accessed the page and selected one appropriate item from the list. Before the experiment, the pages for Exp 2-2 were stored locally and their accessibility was improved. The repaired problems were:

- Alternative texts were added for all images.
- Skip-links were added at the top of each page. The destinations were the first item in the result list on a search results page, and the first mention of the item name on an item page.
- Item numbers ("item 1", "item 2", . . .) were added to each item in a result list on a search results page.
- Heading tags were added to all of the pages. The headings were inserted to preserve the original subregion structure of each page. If no appropriate heading text existed, a text was inserted.
- No content was removed. All business (promotional) tools were preserved.
- The order was not changed. One exception was the "add-to-cart" button area on the item pages on Amazon.com. The button was positioned before the item description and was unrelated to any accessibility definition, so we decided to move it to the area just after the item description by changing the HTML tag order. This order was actually consistent with the visual appearance of the pages.

Each of the enhanced pages was evaluated using aDesigner (see Section 3.3) and main-content reaching times were reduced to less than 10 seconds by these

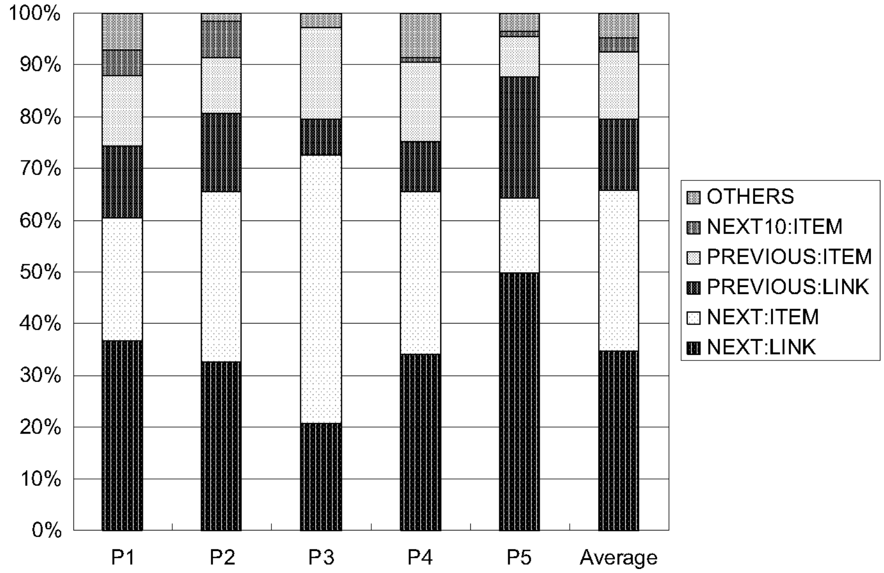


Fig. 10. Ratio of navigation commands (subject).

modifications. The subjects were told that the Exp 2-2 pages had been enhanced and that heading tags and intra-page links might have been inserted into the pages. (See Appendix A for the detailed instructions.)

The search keywords were "IH (Induction Heating) rice cooker" for Exp 2-1 and "Hot Carpet (electric carpet)" for Exp 2-2. These items are popular home appliances in Japan. Each subject was instructed to select an item that matched certain conditions, such as "a rice cooker for 5.5 cups of rice" for Exp 2-1 and "a 1.8 m $\times$ 1.8 m carpet" for Exp 2-2. These conditions were set so that the number of items in each list would be from three to four for each experiment. The subjects were allowed to open the item links, and to access item pages. They were also allowed to go back to the search results page by using the "back" function of the browser (ALT and the Left arrow in HPR). Each task started when the subject opened the search results page and finished when the subject selected an add-to-cart button.

4.4.2 *Procedures, Instrumentation and Participants.* Exp 2-1 and Exp 2-2 were done along with Exp 1, so the procedure is the same as in Section 4.3.2. We utilized the same system as in Exp 1 (See Section 4.3.3), and the same blind subjects (Table V) participated in Exp 2.

4.4.3 *Results and Analysis*

4.4.3.1 *Used Navigation Commands and Scanning Strategies.* Figure 10 shows the total ratio of used navigation commands for Exp 1 and 2. The most-used navigation commands are NEXT and PREVIOUS LINK, and NEXT and PREVIOUS ITEM. These four commands cover about 90% of the commands for all of the participants. This shows that participants mostly used item scanning

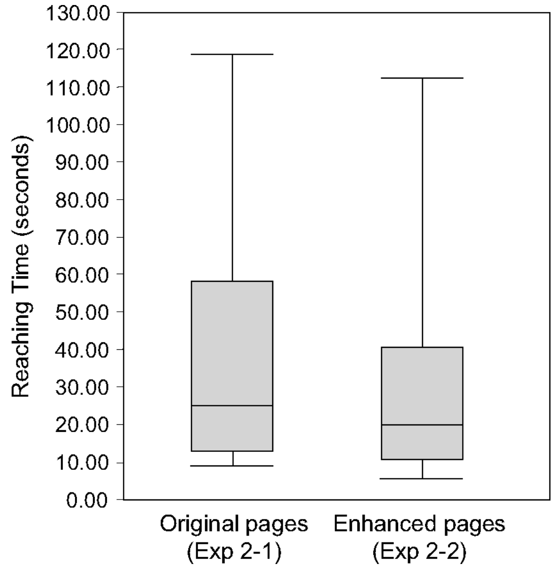


Fig. 11. Comparison of reaching time to result lines in search result pages between Exp 2-1 and Exp 2-2.

and link scanning methods. The 10-ITEM skip navigation commands were also frequently used for scanning.

Among these four commands, the ratios varied for each subject. For example, P3 mostly used the NEXT ITEM and PREVIOUS ITEM commands for scanning. We found from his comments that he was afraid to skip any content, since he frequently experienced skipping over important fragments in the page by using the LINK navigation commands in his daily Web experience. P5 mostly used LINK navigation commands. She had extensive experience with online shopping sites, so she was good at creating mental models based on the links. In the experiment, she could use landmarks in a page to create her own mental model. Intra-page search was rarely used. Some participants tried, but they faced problems because of variations of the key expressions. For example, "1.8 meters by 1.8 meters" can be described as "1.8 m by 1.8 m", "1.8 × 1.8", or "1.8 m × 1.8 m".

4.4.3.2 *Comparison of Original and Enhanced Pages.* We had expected that reaching times would be improved by using heading tags and skip-links. However, surprisingly, there was no difference in the results for the reaching times to the main content area. Figure 11 shows a box plot of the reaching times for the result lines after opening the search result page and after submitting the keywords for Exp 2-1 and Exp 2-2. There were no significant differences. Another indicator for the scanning task is the usage of navigation commands. On average, the use of HEADING navigation commands increased (0.0% to 3.0%), but they varied among participants. There was no consistent and significant difference in the usage of navigation commands. The task completion time was also not significantly different between Exp 2-1 and Exp 2-2, as shown

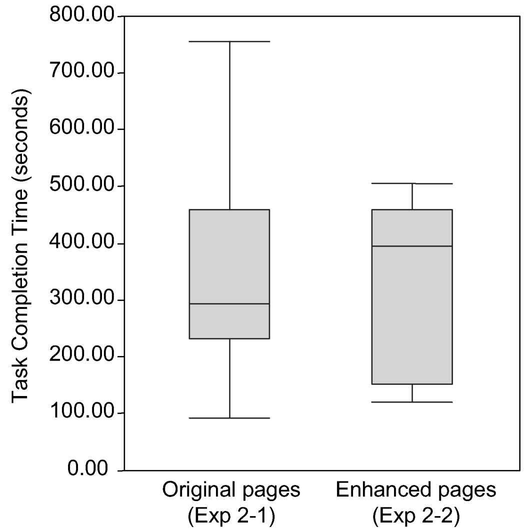


Fig. 12. Comparison of task completion times between Exp 2-1 and Exp 2-2.

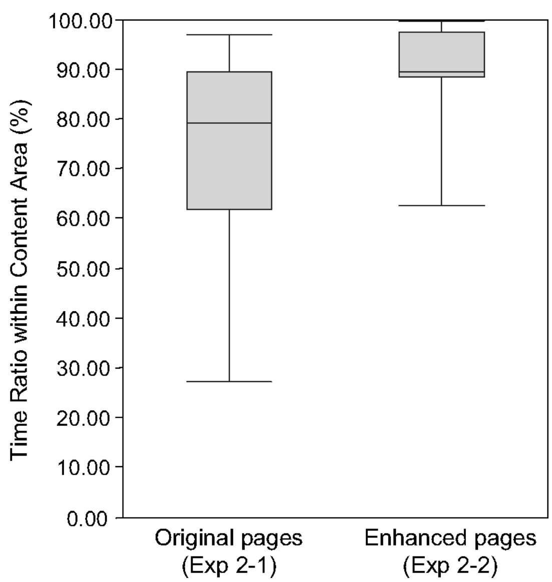


Fig. 13. Ratio of time spent within content area on search result pages.

in Figure 12. The average time for Exp 2-1 was 341 seconds and Exp 2-2 was 335 seconds, so the participants spent almost the same amount of time for one session on average.

On the other hand, the results of the ratios of time spent in content areas showed significant differences. Figure 13 shows a box plot for the ratios for search result pages and Figure 14 shows the ratios for the product pages. It is obvious that participants spent significantly longer time in the content areas

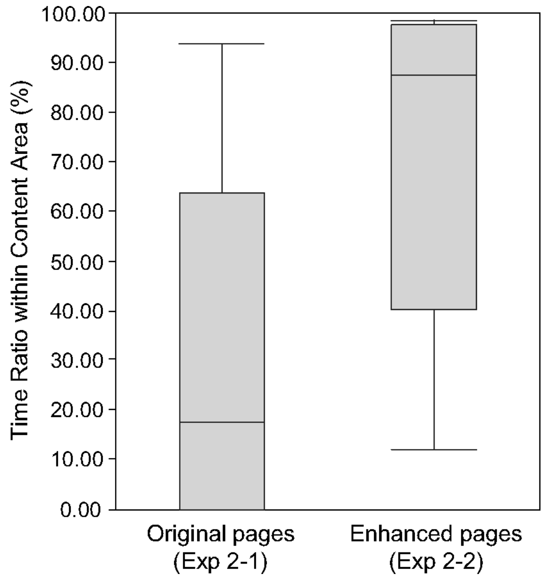


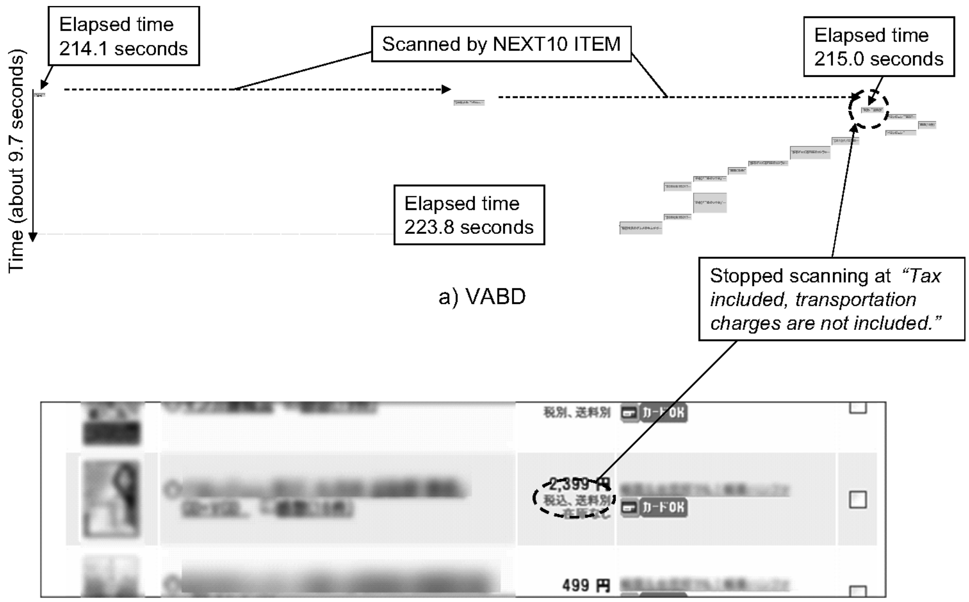
Fig. 14. Ratio of time spent within content area on product pages.

in both cases (t-test, $t = 2.31$ & $p < 0.05$ for search result pages, and $t = 4.00$ & $p < 0.01$ for product pages). This result indicates that participants could explore the necessary content for the tasks without losing their way among unrelated content. In the search results pages in Exp 2-2, the participants could easily recognize the separator of items with the item number indicators ("Item 1", "Item 2", ... in the enhanced pages). Therefore, they reported that they could easily remember the numbers of the candidate items, and then move among these items by using the navigation commands. In Exp 2-1 tasks, the participants sometimes lost their positions and strayed into unrelated content areas. We will discuss these issues more in Section 4.5.1.

4.4.3.3 *Effects of Information Scent.* We analyzed the effects of information scent on participants' behavior by investigating the VABD diagrams. We investigated the following events when a participant changed navigation patterns, and found various examples of the effects of scents.

(1) Scanning was stopped (and then the participant crawled around) to check the content.
(2) A link was clicked before reaching the main content.
(3) The direction of navigation changed.

Figure 15 shows a typical example of content with a strong scent, for which a participant stopped scanning on the fragment. This is an example of case (1). In this case, the participant succeeded in finding the target content. The fragment said "Tax included, shipping charges are not included." The participant had already learned (in the session), that this fragment was frequently used in the search result list (Figure 15(b)). Therefore, right after hearing the



b) Screen shot of the corresponding position to a)

Fig. 15. Example of the effect of information scent.

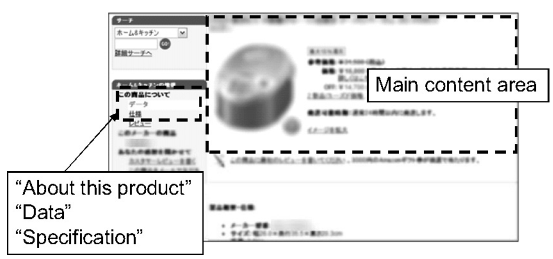


Fig. 16. Example of the effect of an (incorrect) information scent.

fragment, the participant stopped scanning with the NEXT 10 ITEMS commands and started scanning with PREV/NEXT LINK commands. Participants always learned various landmarks around the main content in each experimental session, and built a "local" mental model around the main content.

Figure 16 shows an example of an unsuccessful use of a fragment with a strong scent. This is an example of (2). A link in a link list was selected before reaching the main content area. In the link list, there were links with the text, "About this product, Data, Specification", so the participant expected that some detailed description of the target product might exist in the linked page. Unfortunately, the linked page was a similar product page with only a small amount of additional information. In this case, the participant came to the product page by selecting an item in the search results page, so the product information existed within the page. However, the participant was lured by a link with a strong scent and went to the wrong page. Various other fragments with scent were found in the results. One example is lists of numbers, such as "1 2 3 4 5 . . . ". This type of list of numbers is frequently used for navigation to another page of the same search results, where the "1" without a link indicates "this" page, and the other numbers with links lead to other result pages. This list usually appeared right before the result lines or right after the results, so some participants used the list as scented text. Another example is the scent of footers. The footer of a page usually includes some unique information, such as "company info" or "privacy notice". Participants used these texts for determining that they had gone past the main content area.

The use of scent in nonvisual access is characterized by its fine granularity. Participants did not listen to all of the text of a fragment. They just checked the first part of each fragment for 300 to 700 msec. Though this time length was not statistically confirmed in the data just based on this sampling investigation, it is obvious that blind participants could determine the scent within 1–5 words, and changed their navigation pattern. For sighted users, they determined the scent using the "full page" content, and checked 10 to 100 times larger numbers of words. This difference suggests the utility of "word-level design of content" for the blind instead of the page-level design for sighted users.

4.5 Discussion

4.5.1 *Nonimprovement of Reaching Time with Heading Tags and Skip Links.* In Section 4.4.3, the results indicated that the reaching times to the main content were not improved significantly. This result did not match our expectations. We studied the VABD diagrams and protocols, and found that the participants lacked experience in navigating "accessible" sites and they did not have practical experience in building mental models for these "navigable" pages.

(1) *Misunderstanding of the Roles of Heading Tags*. Participants expected that they could directly reach the main content area by using heading navigation. Heading tags indicate the titles of the subregions in a page, so the content will not be directly reached by using only the heading navigation commands. Heading navigation did not lead them to the exact content but to points just before the content. However, they usually searched for fragments of the main content (fragments with the "scent" of content) by using the scanning navigation commands. Some participants reported that they could not determine where the search result was located, even with a heading labeled "Search Results". We think that such a participant may be having trouble forming a mental model of the page based on the heading structure.

(2) *Strong Tendency to Stick to Their Familiar Scanning Methods*. Participants switched back to their familiar scanning method after they found that heading navigation did not take them directly to the content. It was easy for them to build a mental model by using their familiar scanning commands. In particular, the Exp 2-2, pages had alternative texts and some additional text information, and therefore their usual scanning methods worked

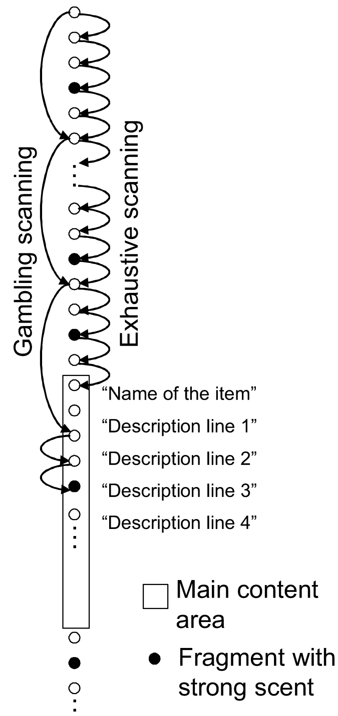


Fig. 17. Observed mental model based on results of Study 2.

better than with Exp 2-1 pages. They reported that the appropriate alternative texts gave them more landmarks and helped them to navigate within the content. Therefore, they still preferred to use the scanning navigation commands.

(3) *The Existence of Skip-Links was not Noticed*. No participants noticed the existence of the skip-links in Exp 2-2. As soon as they opened a page, they jumped from the top of the page to some other position. This is because they believed that no meaningful content exists at the top of a page, and they usually skip any content at the top. Therefore, there was no chance for them to become aware of the existence of the skip-links.

4.5.2 *Mental Models of Blind Users.* In Section 2.1, we proposed a metal model in the form of a one-dimensional string of content fragments. Before the experiment, we expected that participants might build a logical mental model as shown in Figure 1(d). Now, we would like to review the model based on results of Study 2. Figure 17 shows our model based on these results. They focused on landing in the main content area by using gambling scanning or by exhaustive scanning. Sometimes they used landmarks, which have strongly scented information, to find the boundaries of the content area. They memorized audio landmarks both "inside" and "outside" the main content area. This basic access methodology did not change, even if they accessed pages with heading tags. No participant tried to grasp the page structure by using logical navigation such as heading navigation commands.

We think that this is a "missing link" for Web accessibility activities. The logical content model has been focused for improving nonvisual navigability. However, the results clearly showed the importance of landmarks, which are fragments with strong information scent. They utilized even tags for the logical content as landmarks for their scanning. From this result, it is worth considering a new landmark-oriented model for nonvisual navigation. In this model, users could build their mental model as a sequence of landmarks, consisting of intended landmarks by a page author (e.g., heading tags) and unintended landmarks with information scent (e.g., "Tax included" in Figure 15). The intended landmarks work extremely well if they are appropriately embedded across a site. It is, however, rare situation and the subjects in this experiment were not used to using heading tags as landmarks. Therefore, they mainly used unintended landmarks, where every fragment, even a meaningless word, has landmark potential. We will discuss some solutions in Section 5.2.

5. DISCUSSION OF FUTURE POSSIBILITIES

5.1 Global Knowledge Sharing for Building Accessible Online Shopping Sites

Study 1 showed us the inaccessible status of online shopping sites. Such companies are always facing fierce competition among sites and it is sometimes difficult to pay attention to accessibility. However, some counties, such as the U.K., have been succeeding in increasing the number of accessible sites by enacting laws enforcing accessibility for all websites. From the technical point of view, it is important to share the experience of the site designers and developers who worked on building accessible online shopping sites. If their know-how for accessible representation of their promotional tools can be published as guidelines, it would be beneficial for online shopping site developers worldwide.

5.2 Enhancement to Support Landmark-Oriented Navigation

Experiment 2 in Study 2 showed that users tend to adhere to their familiar navigation methods based on unintended landmarks (see Section 4.5.2) and are conservative about using logical navigation methods. Therefore, we proposed a landmark-oriented navigation model in Section 4.5.2. In this section, we will discuss the possibilities to support users in terms of landmark-oriented navigation by enhancing current Web accessibility activities.

5.2.1 *Improvement of HTML/XHTML Specification.* New landmark tags could be one solution to improve the quality of intentional landmarks. The heading tag (H1, H2, H3, ...) is the most important landmark tag in the HTML specifications. However, heading tags do not cover all of the landmarks, such as the "first item in a search result list" or an "add-to-cart button." It is important to allow users to navigate among a sequence of landmarks, which includes all of the important fragments of content. One possibility is to add a "landmark" role to the new WAI Role Taxonomy [Seeman 2006]. This taxonomy has been developed for improving the accessibility of dynamic Web applications [Schwerdtfeger 2006]. It is an extensible taxonomy for adding additional semantics to Web content, so one of the best places to add a new tag. Yesilada et al. [2003] classified various types of nonvisual landmarks into 14 categories based on a traveling metaphor [Goble et al. 2000]. Examples of the resulting categories include Decision Points, Navigation Points, Directions and Travel Assistance. This classification can be a theoretical basis to build a taxonomy for intentional landmarks.

5.2.2 *Simplification of Navigation Interface.* The participants mainly used only four commands for navigation in Study 2 (Section 4.4.3.1). Theofanos and Redish [2003] also reported this tendency. This result suggests the importance of providing simplified intuitive navigation methods for voice browsers. SpeechSkimmer [Arons 1997] is an example of efforts that tried to make navigation in one-dimensional voice data simpler. A navigational controller was proposed, which mapped audio navigational commands to a touch pad. The forward and backward commands were mapped on the vertical axis, while the horizontal axis was used for changing the skimming labels. The research targeted the navigation of recorded voice data, but it might be a good reference for a drastic simplification of the navigation interface. TAJODA [Asakawa et al. 2004] is a jog-dial based navigation interface. It allowed forward and backward movement in the audio information using the dial movements. It also had dynamic speech rate adaptation. Spinning the dial faster, the speech rate automatically became faster, while spinning slower slowed down the speech rate. It is proposed that voice browsers should simplify the interface by referring to such related work.

5.2.3 *Automatic Suggestion of Navigation Methods.* It is clear that voice browsers try to provide richer navigation commands (see Section 2.3), but blind users seldom use those commands. In order to utilize intentional landmarks, it is necessary to effectively use such commands. We therefore propose to improve the functions of voice browsers to be capable of suggesting appropriate navigation commands for each page based intended landmarks in a page. Possible examples would be "There are heading tags in this page, so heading navigation commands will help you," and "There is a skip-navigation link to jump to the main content." Such functions could be possible by analyzing the Web content and by tracking the key sequences pressed by users for detecting their navigation strategies.

5.2.4 *Integration of Transcoding Functions.* We found that blind users are exposed to too many inaccessible pages without logical landmarks. It is important to increase accessible online shopping sites, but the ratio of accessible online shopping sites is not high. In order to increase the number of accessible online shopping sites, Web accessibility transcoding offers great possibilities. Web accessibility transcoding transforms existing (inaccessible) webpages into accessible pages on the fly without modifying the original content. One approach is automatic transcoding [Hanson et al. 2005; Takagi and Asakawa 2000], which aims at filtering webpages on the fly without additional information. Some techniques focused on improving navigability, but accuracy is an open issue. To provide more accurate assistance to users of nonvisual navigation, annotation-based transcoding systems were developed [Asakawa and Takagi 2000; Pontelli et al. 2002; Plessers et al. 2005; Yesilada et al. 2007]. These systems can provide advanced transcoding functions, such as inserting heading tags by using annotations that have additional semantic information about the target page layout. Yesilada et al. [2007] evaluated the improvement associated with an annotation-based transcoding system, and confirmed that there were large improvements in terms of mental workload and "mobility" on webpages. They also found examples of effective use of nonvisual landmarks in experiments. However, the cost of annotation authoring is a major problem inhibiting practical deployment of this technology [Takagi et al. 2002; Fukuda et al. 2003]. If these technologies can be improved and integrated into voice browsers, accessible online-shopping experiences will become more common for blind users.

5.3 Toward an Effective Usability Testing Framework for the Blind

In Study 2, LogHPR was used for collecting usability data, and this showed various aspects of nonvisual Web usability. Compared with usability testing for visual users, logged data for voice access users shows more detailed moment-to-moment data. For example, the ratios of content access times might be difficult to analyze for visual Web access users, even using an eye-tracking system. It is easy for nonvisual access data, since all of the Text-to-Speech output can be logged and also the positions can be logged using XPath. In other words, usability engineers can obtain all of the information about a user's operations and the related data. We think the logging HPR, its analysis tool, and the Voice Access Behavior Diagrams have great potential to contribute to building a new usability testing environment for Web accessibility professionals.

6. CONCLUSION

These studies aimed at understanding the navigability of real-world Web applications, especially for online shopping sites. In Study 1, we tried to give an overview of the current status of worldwide online shopping sites. We evaluated 30 representative sites, and we found a few nonvisually navigable sites, but most of the sites were poor as regards accessibility. The average reaching time for the first result line in item search result pages was 98 seconds for the U.S. and 178 seconds for Japan. For the U.S. and Japanese sites, skip-links were used in only two sites, and only three sites had high ratios (more than 80%) for pages with heading tags. These results supported our expectations for poor accessibility of shopping sites. We saw signs of improvements in the U.K., where new regulations appear to be affecting the levels of accessibility, and the U.K. results showed significantly better navigability. The average reaching time was 49 seconds, and half of the examined sites had high ratios (more than 80%) of pages with heading tags.

Study 2 aimed at investigating real users' behavior while accessing online shopping sites. We developed LogHPR, which is a modified version of a commercial voice browser, IBM Home Page Reader, to output users' operations and the browsers' internal data to log files. We also developed a log analysis tool and Voice Access Behavior Diagrams for visualizing the log files. Experiment 1 was conducted to measure completion times for purchasing items both for the sighted and the blind. The results showed us that blind users took more than 10 times longer compared with sighted users. This quantitative data clearly shows the difficulty of nonvisual access. Experiment 2 was conducted to compare usability improvements for original (inaccessible) sites and enhanced (accessible) sites. The results showed surprisingly small improvements for the accessible sites. The reaching time to the main content was not improved, even though accessible pages had skip-links and heading tags. The results showed that participants adhered to their familiar scanning navigation methods, and mental models were not built to understand the logical contents. We propose the concept of landmark-oriented nonvisual navigation model based on the results. In this model, both intended landmarks (e.g., heading tags) and unintended landmarks with information scent are utilized to navigate.

It is obvious that increasing the numbers of useful landmarks is important. At the same time, improvement of HTML specification is one possibility. Another possibility is improvement of voice browsers to support users to learn appropriate navigation methods for each page they access. We hope the results of this study will contribute towards improving the nonvisual Web access environment in the near future.

APPENDIX: INSTRUCTIONS FOR STUDY 2

Exp. 1

— Please buy the DVD, "Fuyu no Sonata DVD Volume 1"[3]
— "Fuyu" is written as a Japanese Kanji character. "Sonata" is written in Japanese katakana characters. "DVD, box and vol" are written in the Latin alphabet and there is hyphenation between "DVD" and "box".[4]
— There are DVD-box vol 1 and 2, but you need to find vol 1.
— It might be read as vol I instead of vol 1.
— It costs about 15,000 yen, for your reference.
— There is a similar DVD called "Fuyu no sonata soushuhen", so remember the target DVD name correctly.
— Perform the task on both Rakuten and Amazon shopping sites and use any function on both shopping sites.
— You should not search from something like the Google search engine.

Exp. 2-1

The starting page is a result page for IH rice cooker. Select your favorite rice cooker from the result page. There are two search result pages from Rakuten

[3]"Fuyu no sonata" is translated as "Winter Sonata". It is a Korean love story that is tremendously popular in Japan.
[4]It is important to let a subject know how each word is written, with Kanji, Katakana, or Roman letter. If they make a mistake, the search result will not be correct.

and Amazon. For the Rakuten result page, select one from among the products that holds three cups, while for Amazon, find one with 5.5 cups and add it to your shopping cart.

Exp. 2-2

The experiment is similar to Experiment 2-1. The starting pages are search results pages for Hot Carpets from Rakuten and Amazon. Select your favorite carpets among the products whose names contain "for two mats" and add each to a shopping cart. Unlike Experiment 2-1, these result pages have been modified by the experimenters by considering accessibility in voice browsers.

ACKNOWLEDGMENTS

We thank the Barrier Free Project and the University of Tokyo for their assistance in finding appropriate blind subjects for Study 2. We would like to thank all of the blind participants who participated in the experiments for Study 2. It took each of them about 4 hours, but they did the long experiments without any complaints. We thank Takashi Itoh at our laboratory, who modified HPR for the logging, and who gave us many suggestions for analyzing the log file.

REFERENCES

ARONS, B. 1997. SpeechSkimmer: A system for interactively skimming recorded speech. *ACM Trans. Comput.-Human Interact. 4*, 1, 3–38.

ASAKAWA, C. AND ITO, T. 1998. User interface of a home page reader. In *Proceedings of Assets '98: The 3rd International ACM Conference on Assistive Technologies* (Marina del Rey, CA, Apr.). ACM, New York, 149–156.

ASAKAWA, C. AND TAKAGI, H. 2000. Annotation-based transcoding for nonvisual Web access. In *Proceedings of Assets '00: The 4th International ACM Conference on Assistive Technologies* (Arlington, VA, Nov.). ACM, New York, 172–179.

ASAKAWA, C., TAKAGI, H., INO, S., AND IFUKUBE, T. 2004. TAJODA: Proposed tactile and jog dial interface for the blind. *Trans. IEICE, E87-D*, 6.

BARNICLE, K. 2000. Usability testing with screen reading technology in a Windows environment. In *Proceedings of CUU '00: The Conference on Universal Usability* (Arlington, VA, Nov.). ACM, New York, 102–109.

BLACKMON, M. H., POLSON, P. G., KITAJIMA, M., AND LEWIS, C. 2002. Cognitive walkthrough for the Web. In *Proceedings of CHI '02: The SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, MN, Apr.). ACM, New York, 463–470.

BLACKMON, M. H., KITAJIMA, M., AND POLSON, P. G. 2005. Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In *Proceedings of CHI '05: The SIGCHI Conference on Human Factors in Computing Systems* (Portland, OR, Apr.). ACM, New York, 31–40.

CABINET OFFICE JAPAN. 2004. Annual report on government measures for persons with disabilities 2004. (See http://www8.cao.go.jp/shougai/whitepaper/h16hakusho/zenbun/ (in Japanese)).

CARD, S. K., PIROLLI, P., WEGE, M. V. D., MORRISON, J. B., REEDER, R. W., SCHRAEDLEY, P. K., AND BOSHART, J. 2001. Information scent as a driver of Web behavior graphs: Results of a protocol analysis method for Web usability. In *Proceedings of CHI '01: The SIGCHI Conference on Human Factors in Computing Systems* (Seattle, WA, Mar.). ACM, New York, 498–505.

CHALLIS, B. P. AND EDWARDS, A. D. N. 2000. Weasel: A computer based system for providing non-visual access to music notation. In *Proceedings of the ACM SIGCAPH Computers and the Physically Handicapped*. ACM, New York, 1–12.

CHISHOLM, W., VANDERHEIDEN, G., AND JACOBS, I. 1999. *Web Content Accessibility Guidelines 1.0.* W3C Recommendation. (See http://www.w3.org/TR/WCAG10/.).

CHISHOLM, W., VANDERHEIDEN, G., AND JACOBS, I. 2000. *HTML Techniques for Web Content Accessibility Guidelines 1.0.* W3C Note. (See http://www.w3.org/TR/WCAG10-HTML-TECHS/.)

CLARK, J. AND DEROSE, S. 1999. *XML Path Language (XPath) Version 1.0.* W3C Recommendation (See http://www.w3.org/TR/xpath/.)

FUKUDA, K., TAKAGI, H., MAEDA, J., AND ASAKAWA, C. 2003. An assist method for realizing a Web page structure for blind people. In *Proceedings of UAHCI '03: The 2nd International Conference on Universal Access in Human-Computer Interaction* (June). Lawrence Erlbaum Associates, Inc., 960–964.

GOBLE, C., HARPER, S., AND STEVENS, R. 2000. The travails of visually impaired web travellers. In *Proceedings of the 11th International ACM Conference on Hypertext and Hypermedia.* ACM, New York, 1–10.

HANSON, V. L., BREZIN, J. P., CRAYNE, S., KEATES, S., KJELDSEN, R., RICHARDS, J. T., SWART, C., AND TREWIN, S. 2005. Improving Web accessibility through an enhanced open-source browser. *IBM J. Res. Devel. 44*, 3, 573–588.

JACOBS, I. AND BREWER, J. 1999. *Accessibility Features of CSS.* W3C Note. (See http://www.w3.org/TR/CSS-access.)

NCHS. 1996. National health interview survey on disability 1994–1995, NCHS (National Center for Health Statistics). (See http://www.cdc.gov/nchs/about/major/nhis_dis/nhis_dis.htm.)

PIROLLI, P. 2006. The use of proximal information scent to forage for distal content on the World Wide Web. In *Working with Technology in Mind: Adaptive Perspectives on Human-Technology Interaction*, A. Kirlik, Ed. Oxford University Press.

PLESSERS, P., CASTELEYN, S., YESILADA, Y., TROYER, O. D., STEVENS, R., HARPER, S., AND GOBLE, C. 2005. Accessibility: A web engineering approach. In *Proceedings of WWW '05: The 14th International Conference on World Wide Web* (Chiba, Japan, May). ACM, New York, 353–362.

PONTELLI, E., GILLAN, D., XIONG, W., SAAD, E., GUPTA, G., AND KARSHMER, A. I. 2002. Navigation of HTML tables, frames, and XML fragments. In *Proceedings of Assets '02: The 5th International ACM Conference on Assistive Technologies* (Edinburgh, Scotland, July). ACM, New York, 25–32.

SALAMPASIS, M., KOUROUPETROGLOU, C., AND MANITSARIS, A. 2005. Semantically enhanced browsing for blind people in the WWW. In *Proceedings of HYPERTEXT '05: The 16th ACM Conference on Hypertext and Hypermedia* (Salzburg, Austria, Sept.). ACM, New York, 32–34.

SEEMAN, L. 2006. Role taxonomy for accessible adaptable applications, See http://www.w3.org/WAI/PF/GUI/.

SCHWERDTFEGER, R. 2001. Dynamic Accessible Web Content Roadmap, See http://www.w3.org/WAI/PF/roadmap/.

STEVENS, R. 1996. Principles for the design of auditory interfaces to present complex, information to blind people. Ph.D. dissertation, University of York, York, UK.

TAKAGI, H. AND ASAKAWA, C. 2000. Transcoding proxy for nonvisual Web access. In *Proceedings of Assets '00: The 4th International ACM Conference on Assistive Technologies* (Arlington, VA, Nov.). ACM, New York, 164–171.

TAKAGI, H., ASAKAWA, C., FUKUDA, K., AND MAEDA, J. 2002. Site-wide annotation: Reconstructing existing pages to be accessible. In *Proceedings of Assets '02: The 5th International ACM Conference on Assistive Technologies* (Edinburgh, Scotland, July). ACM, New York, 81–88.

TAKAGI, H., ASAKAWA, C., FUKUDA, K., AND MAEDA, J. 2004. Accessibility designer: Visualizing usability for the blind. In *Proceedings of Assets '04: The 6th International ACM SIGACCESS conference on Computers and accessibility* (Atlanta, GA, Oct.). ACM, New York, 177–184.

THEOFANOS, M. F. AND REDISH, J. G. 2003. Bridging the gap: Between accessibility and usability. *ACM Interact 10*, 6, 36–51.

WATANABE, T. 2002. Windows PCs and internet usage survey for people with visual impairments in Japan. The National Institute of Special Education Research Note (See http://www.nise.go.jp/research/kogaku/twatanab/PCUserSurvey/Survey2002/Survey2002Jp.html (in Japanese)).

YESILADA, Y., STEVENS, R., AND GOBLE, C. 2003. A foundation for tool based mobility support for visually impaired web users. In *Proceedings of the 12th International Conference on World Wide Web* (Budapest, Hungary, May). ACM, New York, 422–430.

YESILADA, Y., STEVENS, R., HARPER, S., AND GOBLE, C. 2007. Evaluating DANTE: Semantic Transcoding for Visually Disabled Users. *ACM Trans. Comput.-Human Interact. (TOCHI) 14*, 3 (Sept.), Article 14.

Received June 2006; accepted June 2006 by Andrew Sears, Viki Hanson, and Brod Myers