Original Article

A study on the usability of hierarchical structures in websites for the visually impaired

Junichi Iizuka (Research and Support Center, Tsukuba University of Technology, jiizuka@k.tsukuba-tech.ac.jp)
Akira Okamoto (Tsukuba University of Technology, okamotovla@nifty.com)
Kazunari Morimoto (Graduate School of Science and Technology, Kyoto Institute of Technology, morix@kit.ac.jp)

Abstract

Despite all of the information-gathering media, it remains unclear whether or not any website structures allow easy access to information for the visually impaired. In this study, visually impaired people who use screen readers were targeted, and the amount of informational content displayed on webpages and the hierarchical structures of the sites was examined. The results of a mental workload evaluation considering the search time and information access operations showed that the search time and the mental workload were considerably lower in the case of site structures that had less link text on the top page and a lot of second-layer content (information provided on the website). Furthermore, although the visually impaired subjects required a longer time to complete searches, their mental workload trend was similar to sighted users with respect to the differences in the hierarchical structure. This indicates that a high-usability website that is efficient and requires little mental workload when searching for information can be created for both sighted and visually impaired persons.

Key words

website, accessibility, usability, mental workload, visually impaired

1. Introduction

For the visually impaired, personal computers are now essential devices for them to access information. Because they cannot see the information that is presented on the display, they need to receive the information through either a screen reader that reads the text on the screen aloud in a synthetic voice or through a Braille display device that outputs Braille in real-time. However, when using these devices, it is very difficult to understand frames, menus, charts, etc., that are frequently used for arranging the contents of a webpage, as well as the overall site structure. This is because when two-dimensional content is converted into a one-dimensional medium, such as voice or Braille display, the users require a strong memory to remember the information content on the page as well as its layout. Essentially, because a sighted person can quickly perceive the colors, shapes, and sizes of text, charts, and photographs on the display screen, the amount of information that must be remembered when searching is less compared to that for the visually impaired. The visually impaired cannot make these distinctions immediately. For example, when a screen reader sequentially converts the information into audio, the person is relying only on sound. With a screen reader, it is difficult to quickly re-hear only the selected content, and it is necessary to remember the point where the information to be played back was presented. Therefore, it is considered that when a visually impaired person uses a computer, that person is forced to have a strategy for remembering more content than a sighted person would need to.

Thus, a method for the visually impaired to obtain information from websites (ensuring accessibility) is required; and for "web accessibility," concrete guidelines are listed in the JIS standard X 8341-34 and the W3C's WCAG 1.05 (Japan Standards Association, 2010; WCAG 2.0, 2008). However, most of these standards only describe the way to take in content by screen reader. It can be thought that it is not sufficient to simply be able to search the contents; it is also extremely important for people to be able to "easily" obtain information (considerations of usability) through a one-dimensional means such as sound, from the entire breadth and depth of a two-dimensional website.

There have been a series of studies and proposals on web usability for sighted people. If websites have the same amount of information (web content), then a page should not have a considerable amount of content and the overall hierarchy (layers) of the site should not be very deep (Miller, 1981. Snowberry et al, 1983. Larson et al, 1998. Nielsen, 1999). For example, with 64 content items, it is possible to go from a broad, shallow website with all the contents presented on the top page (structure of 64 items on one layer) to a top page with two text links, and whichever is chosen, it will lead to two more text links; thus, there is a narrow, deep website that always offers a straight choice between two things (a six-layer hierarchy of $2 \times 2 \times 2 \times 2 \times 2 \times 2$). As a result of studying the changes in the site structure, it has been reported that the search time with a two-layer structure is the shortest.

In contrast, when it comes to websites that are easy for the visually impaired to use and finding preferable site structures, there is nothing to be found, and provisions for a specific site structure have not been established in JIS X8341-3 or W3C's WCAG 1.0. Furthermore, it is unclear as to whether or not visually impaired persons can go over the same information in detail, listening to a synthesized voice, on websites that sighted persons can glance through within a short time. For the visually impaired, the larger the amount of content displayed on a single page, the longer it took to find the desired information, but movement up and down the layers of the site hierarchy was performed easily through keystrokes. However, since it is troublesome to go back up in the hierarchy if a wrong selection is made, for the sake of efficiency, accurate mobility throughout the hierarchy is important. In this study, various websites of different breadths and depths (amount of information per page and number of layers in the hierarchy) were created with the goal of revealing the types of website hierarchical structures that would be easier for visually impaired persons to use in terms of efficient information retrieval and mental workload.

2. Ease of website use

2.1 Evaluating website usability

In order to evaluate the usability of websites, the available indices of effectiveness was examined and it was decided to use the two following indices:

2.1.1 Behavioral index

As a behavioral index, evaluating the usability of websites information retrieval time was used. Other measures such as the error rate in information retrieval, and the side issue of multi-tasking while listening to music, were also considered, but the time it took to find information on a website was the standard measure for evaluating information retrieval efficiency, and was deemed to be a suitable means for measuring the operational efficiency of the visually impaired participants.

2.1.2 Subjective index

As a subjective evaluation of fatigue, methods such as NASA Task Load Index (NASA-TLX) and the Subjective Workload Assessment Technique (SWAT) are generally used for sighted persons. In this study, it was decided to use NASA-TLX because participants could easily understand and respond to verbal explanations, and also because visually impaired participants were not required to look at evaluation items. In addition, since the participants were being mentally evaluated for the first time using NASA-TLX, it was considered that the participants, both sighted and visually impaired, were able to confirm the terminology used for measurements at all times.

2.2 Evaluation indices and hypotheses

In general, usability is expressed as the ease of understanding something. In this study, website usability was defined as a website being both straightforward and easy to use. Furthermore, an easy-to-understand and easy-to-use website has two necessary qualities that must be met: (1) whether the content can be searched within a short period of time and (2) whether the content being sought can be approached with some sense of certainty, i.e., whether there is a sense of ease in finding what is being looked for. The measures were defined as described in Section 2.1 for the two corresponding evaluation indices, with the former being the behavioral index of search times for information and the latter being the subjective index of the NASA-TLX evaluation.

Here, in order to develop a hypothesis, the characteristics of the visually impaired using a computer were considered. From the results of previous studies (Iizuka et al., 2007) and observation of the use of computers by the visually impaired, it appeared that when there is a considerable amount of content presented on a single webpage in parallel, because the visually impaired cannot predict what is displayed, each piece of content must be confirmed one at a time without skipping over any parts. Thus, verifying the contents takes time, except for the headings in which contents are classified by genre. Therefore, many visually impaired people struggle with webpages that use a parallel presentation of information. On the other hand, with keyboard shortcuts, things such as headings and familiar menus can be easily jumped to, and the categories can be skipped in between. Similar movements can be used to go up and down the hierarchical structure. The following hypothesis was examined on the basis of these properties.

Hypothesis: For visually impaired persons using screen readers, the fewer the text links that appear on each webpage, the shorter is the amount of time that it takes to select information. Accordingly, content on a website arranged in a "narrow and deep hierarchical structure" requires a lesser mental workload. In contrast, when there is a considerable amount of content on the top page and the site has a "broad and shallow hierarchical structure," it takes time to search for information and comparatively the mental workload increases. That is, the site structure preferable for the visually impaired is different from that for sighted people, and has a relatively shallow and broad hierarchical structure.

3. Experimental methods

3.1 Website creation

3.1.1 Hierarchical structure

The hierarchical structures used for these experiments were produced using website structures previously published in a paper by Zaphiris (Zaphiris, 2000). The search theme was things related to a Mediterranean country. Five model site structures that arranged sixty-four items of content from narrowly and deeply to broadly and shallowly were studied.

Model 1: $2 \times 2 \times 2 \times 2 \times 2 \times 2$; six-layer structure with two text links on all layers

Model 2: $4 \times 4 \times 4$; three-layer structure with four text links on all layers

Model 3: 4 × 16; two-layer structure with four text links on the first layer, each of leading to 16 text links on the second layer

Model 4: 8 × 8; two-layer structure of eight text links, each of which leading to a layer of eight text links

Model 5: 16 × 4; two-layer structure of 16 text links on the first layer, each of which leading to four text links on the second layer

Not only do two text links appear on Model 1's first layer (top page), but the following layers also have two link options, forming a "narrow and deep six-layer hierarchical structure." In contrast, Model 5's (16 × 4) top page has 16 text links, and the second layer has four content items, forming a "wide and shallow two-layer hierarchical structure." Screenshots including the top page and the final layer of these hierarchical structures are displayed in the Appendix.

3.1.2 Webpage content placement

Next, when viewing the website structure, the alignment of the content on the page needs to be simplified so that searching with a screen reader would be easy (Figure 1). In particular, the first line designates the name of the layer in the hierarchy, the second line has the topic path (breadcrumb list), and the third line shows the selected content.

Location, area, coastline and marine areas of Cyprus
Current location; Top> 2nd layer> location, area, coastline, marine area
Location
Area
Coastline
Marine area

Figure 1: Example of webpage display

For this trial, the same web contents (information) as in Zaphilis' paper were used because a majority of Japanese people are fairly unfamiliar with the country of Cyprus and it was considered that there would be less influence from prior knowledge of the participants on the execution of searches. In addition, it was decided to use general terminology and the use of jargon for presenting text links and content was avoided.

3.2 Participants

Nine students (average age: 21.0 ± 0.9 years) with visual impairments enrolled at Tsukuba University of Technology participated in the experiment. Because they were either unable to see the computer screen's display or it was difficult for them to do so, when they used a screen, they made regular use of a screen reader. They had a fast input speed and were familiar with the use of writing software and information retrieval through a browser. It should be noted that the screen reader used in the experiment was one that they were familiar with and used every day (PC-Talker 7, Kochi System Development, Inc.).

Furthermore, all information retrieval tasks performed by the visually impaired subjects in the experiments were also performed by the sighted subjects, and the differences between the trials were examined. The sighted participants had a similar level of personal experience of using computers and included 13 students aged 20.8 ± 3.6 years from a nearby technical school. These students used computers daily for creating reports and gathering data, and they were all familiar with word processing software and browsers.

The informational content used in the experiment was considered something that Japanese students, irrespective of their vision, would be unfamiliar with, and the terminology used would be accessible to everyone with little influence from participants language ability or existing knowledge. Furthermore, as shown in Figure 1, the arrangement of the content on the page was very simple, so there was little possibility of the screen reader having trouble with the content. Moreover, as will be described later, the results of a memory test showed that both the visually impaired persons and the sighted persons showed the same degree of retention. Therefore, it can be believed that there will be no problem in comparing and examining the data of the visually impaired and the sighted participants.

3.3 Screen reader

The current computer operating systems and a lot of application software are developed under the assumption that the buttons and menu items are selected by using the mouse cursor. However, people with impaired vision cannot confirm the position of the mouse cursor on the screen, and even when it is possible to use a keyboard, operating a computer can be very difficult for them.

A screen reader, as the name suggests, is software that converts the information on a display screen into synthesized speech. Operations which are normally performed by a mouse can be performed with a keyboard, and it is possible to have the full text read from the beginning or use shortcut keys to read only the headings and move between webpages. If the website has a designated order written into the html source code, the software can read in that order from the top, and the string that provides a link can be recognized through a change in the quality of the voice reading the text. Essentially, if a word on the display is a text link to another layer, a visually impaired person can recognize it as different from the other content through the reading voice's characteristics.

Since the participants with visual disabilities used screen readers routinely, when conducting the experiments, they were allowed to use their own operational conditions, i.e., no special parameters were set regarding the speed at which the text was read or the preferences for key operation.

3.4 Memory test

The memory skills of the participants are considered to have a significant impact on the efficiency of information retrieval from a website. When unable to find the target contents while looking through a site, thinking about whether or not the page has already been viewed, what category was chosen earlier, etc., are related to the ability of the user to mentally retain the past browsing history, and are thought to impact not just the search time but also the mental workload. Therefore, major differences in the memory skills of the participants were checked. Although there are existing memory tests, such as the kit of factor-referenced cognitive tests (Ekstrom et al., 1976), many of these involve answering while looking at printed documents or a display. For the visually impaired persons, a memory test that required the person to look at it could not be used. Therefore, a new software for measuring memory skills that corresponded to a screen reader was developed.

The procedure for the memory test is summarized as follows. The software is launched, and a number input area is displayed at the center of the screen; the experiment is started by pressing the enter key. A beeping sound rings; then multiple strings of numbers are read by the screen reader's synthetic voice. The strings of numbers are not displayed on the screen. Each number is read aloud at one second intervals, and in order to clarify when the reading of the numbers has concluded, there is a final beep at the end. When finished, the participants are asked to enter the string of numbers in the number input area. It should also be noted that until the enter key is pressed to begin the reading of the next string, input mistakes can be corrected.

The experiment starts with the presentation of three digits. If the answer is correct, the length of the string is increased one digit, and if the answer is incorrect, the length of the string is decreased one digit. The participants were not told whether or not an answer was correct. The test was completed when the answer status changed five times from either correct to incorrect or incorrect to correct. These series of tests were conducted twice. This was because it was thought that the participants might adopt a different strategy for memorization that would allow them to remember more on the basis of the results of the first trial. The participants' memory capacity was defined as a relatively large number of correct answers from the two trials. The operation history for the memory test was recorded onto the computer for data analysis.

3.5 Website search and search time

Before the experiment, the participants were informed that the experiment would involve retrieving information of Cyprus from websites that used five different hierarchical structures. The websites were presented at random, but the participants were not told about the type of hierarchical structure they would be working with. In order to enable the visually impaired participants to answer easily, they were given the search terms they would be searching for in Braille, and were allowed to read at any time during the experiment.

The cue for starting the experiment was the presentation of the top page of the website, at which point the search for information began. The participants listened to the synthetic voice of the screen reader and searched. The participants were asked to give a cue verbally as soon as the specified information was found, and the time between the start of the search and when the cue was given was recorded as the information search time. Meanwhile, data relating to the following four areas was simultaneously recorded on both audio and video (Figure 2): participant operation, transitioning between each webpage, screen reader voice output, and participant vocalizations.

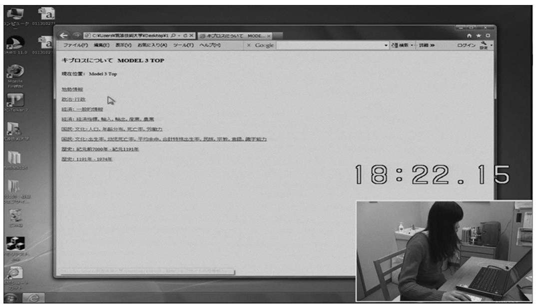


Figure 2: Example screen of recording

Note: Sighted persons also looked at webpages and made similar searches.

3.6 Measuring mental workload

First, consent was obtained after describing to participants the objective of the experiment, the experimental methods, the concept of mental workload, and the measurement method used with NASA-TLX. Then, a preliminary test of a website search was performed. Simple search challenges were completed on the Tsukuba city website.

In the experiment, there were two search topics presented for each model, and when the search was completed, the mental workload was measured. First, evaluation values were obtained for each of the six subscales of the NASA-TLX: mental demand, physical demand, temporal demand, performance, effort, and frustration; for each subscale a 10-point evaluation was conducted verbally with the participants. Next, the weight for the rating for each of the six subscales was determined by presenting them two at a time to the participants and selecting which was important as a factor in the task's workload (binary comparison). Then, the values for the six subscales and the weighting from the binary comparison were multiplied and the sum provided a comprehensive evaluation value.

Although there are various Japanese translations of the subscales, in this study, it was decided to use those of Haga (Haga, 2001) because they are easy to understand.

4. Experimental results

4.1 Memory test

The results of the reproduction numbers for the visually impaired (nine students, A–I) are shown in Figure 3.

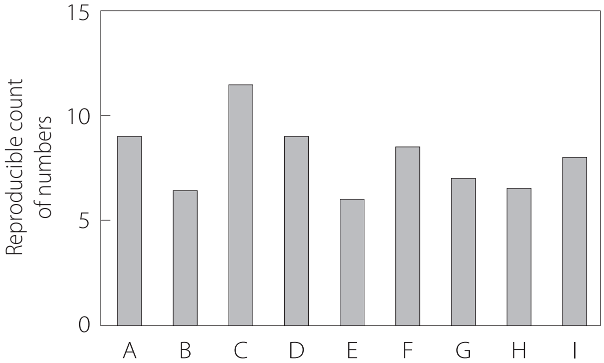


Figure 3: Reproducible count of numbers for visually impaired

The vertical axis shows the largest number of correct answers in reproducing the read numbers. The mean value was 8.0 ± 1.7. Furthermore, as a reference, the mean value for the sighted students that participated in the experiments was 7.0 ± 0.6. Both results fell within the range of Miller's magical number 7 ± 2, which is the human limit for information that can be remembered for a temporary period of time. When examining the differences between the reproduction numbers for the visually impaired and the sighted, no significant difference with 5 % level of significance was observed. Therefore, it was concluded that there was no significant difference in the memory capabilities of the sighted participants and the visually impaired participants.

4.2 Search time

The mean search times for each model of visually impaired participants are shown in Figure 4.

The mean search time was determined by taking the average of the required time to complete two search challenges for each model. Model 3 (4 × 16) took the least amount of search time at 20 s. Model 1 (2 × 2 × …) took the second shortest amount of time, but at over three times the length at 67 s. The longest search time was for Model 2 (4 × 4 × 4) at 96 s. The order of the mean search time from the shortest to the longest was as follows:

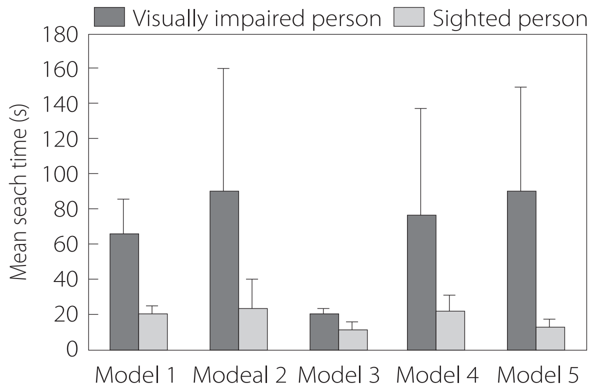


Figure 4: Average search times for visually impaired and sighted persons

Model 3 (4 × 16) < Model 1 (2 × 2 × …) <
Model 4 (8 × 8) < Model 5 (16 × 4) ≦
Model 2 (4 × 4 × 4)

Next, the mean search times for sighted people were examined. All the search times were around 20 s, and the order from the shortest to the longest was as follows:

Model 3 (4 × 16) < Model 5 (16 × 4) <
Model 1 (2 × 2 × …) ≦ Model 4 (8 × 8) <
Model 2 (4 × 4 × 4)

The visually impaired subjects took a long time in all models. Of the models, Model 5 (16 × 4) was particularly time consuming for the visually impaired subjects and took about seven times as long when compared with the sighted subjects.

4.3 Mental workload

Figure 5 shows the average evaluation values for the NASA-TLX for the visually impaired subjects. Model 3 (4 × 16) had the lowest score at 17, followed by Model 1 (2 × 2 × …),

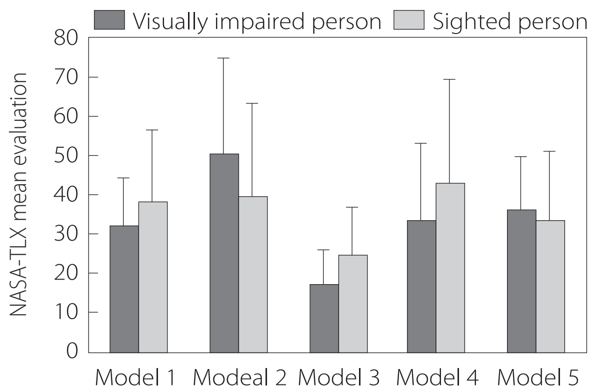


Figure 5: NASA-TLX average evaluation values for visually impaired and sighted persons

Model 4 (8 × 8), and Model 5 (16 × 4), that were all about the same; the greatest mental workload was for Model 2 (4 × 4 × 4) at 51. When looking at the average evaluation values for the NASA-TLX for the sighted subjects, Model 3 (4 × 16) had the lowest score, and in contrast, Model 4 (8 × 8) had the highest score. Still, the trend in the variation of the mean values of the NASA-TLX caused by the differences in the model was similar between the visually impaired and the sighted groups.

5. Discussion

5.1 Regarding search time and mental workload

5.1.1 Relationship of hierarchical structure and search time

Of the five models, Model 3 (4 × 16) took the least amount of time to complete searches. It was felt that this was due to the relatively small number of links, i.e., only four text links on the top page. Thus, the effort to move between the text links was also little. In addition, it was thought easy to select from among the four links the one with a genre closest to the search terms. Moreover, it seems that movement to the second layer was possible without reading through the remaining text links; in other words, without getting lost. Then, because the user could be sure that the search term would be in the contents displayed on the second layer, the reading voice could be set to a higher speed for searching. It was possible to ascertain these patterns by observing the video recordings.

Model 2 (4 × 4 × 4) was similar to Model 3 (4 × 16) with respect to having four text links on the top page. However, on the second and third layers there were also four text links, and when compared with the text links on the top page, the four text links on the second layer showed a greater degree of similarity in meaning. Therefore, in order to search for the text link closest to the desired search term, many of the participants listened to all four text links carefully before moving to the third layer. This is thought to be why it demonstrated a longer search time.

Model 4 (8 × 8) and Model 5 (16 × 4) have the same two-layer structure as Model 3 (4 × 16), but because there are many text links on the top page, users had to go back and forth many times to confirm the content in order to select the correct text link, and sometimes selected the wrong text link and returned to the top page after having moved to the second layer. This is considered to be why these models had a longer search time.

In the case of Model 1 (2 × 2 × …), because only two text links appear on each layer, it was initially thought that choosing between two text links for the one close to the search term would be efficient despite its deep hierarchy. However, in this experiment, the result was that it took a significantly longer search time than Model 3 (4 × 16).

When looking at the video recordings, the participants were observed spending time making their selections on each layer, and as they progressed further down the hierarchy, they had the tendency to take an accordingly longer amount of time. This is because the meaning of the text links when moving down the hierarchy becomes increasingly similar, and deciding which to choose in order to trace the search term becomes difficult. If the wrong text link is chosen, then it is necessary to go back up one level higher in the hierarchy. Therefore, particularly when compared with the shallow Model 3 (4 × 16) hierarchy, it seemed that the participants were very cautious about deciding which text link they should choose.

5.1.2 Relationship of hierarchical structure and mental workload

Model 3 (4 × 16) required the least amount of mental workload. A sense of accomplishment resulting from finding search terms within a short time can be pointed to as the reason for this. It is also believed that this is because the participants could build confidence through a steady approach to the search objective. This is attributed this to the fact that the relatively few four text links on the initial top page makes it possible to select with confidence a text link that is close to the search term. This is corroborated by the results of the interviews: "with a broad classification on the first layer, I was able to make a choice without questioning it." Furthermore, as mentioned in the previous section, immediately after moving to the second layer the participants were sure that the search term would be there, and when they found the desired search term, a sense of happiness was observed, which is thought to also have been a source for reducing the mental workload. This is backed by the opinion expressed in an interview that "I was happy that the search term I was looking for was on the second layer." Although all the movements and speech of the participants were recorded, since the visually impaired individuals were concentrating on listening to the voice of the screen reader, they did not speak at all during their searches. However, even though there was no protocol analysis based on speech, a "relieved" appearance was observed in all the participants at the moment a search term was found, suggesting a low level of mental workload.

On the other hand, in Model 2 (4 × 4 × 4), the second layer's text links, which suggest the existence of lower layers, show a higher degree of similarity between them than the text links on its top page. Therefore, in order to select the correct link text, the four text links need to be checked many times, and the added pressure to be careful in selecting the text links on the second layer is thought to be a major cause of the increase in the mental workload.

Model 4 (8 × 8) and Model 5 (16 × 4) have the same two-layer structure as Model 3 (4 × 16), but because there were many text links on the top page, users had to go back and forth many times to confirm the content.

Model 1 ($2 \times 2 \times \ldots$) had a six-layer hierarchy, but the participants did not know this depth in advance With each successive layer of two choices, the frustration among the participants would become apparent. The participants would be increasingly anxious pertaining to their selection of the correct choice in the preceding layers. In interviews after the searches, there was an opinion expressed that "it was not possible to know if a choice was correct, and there was a feeling of uneasiness that there would probably be a need to return to a higher level of the hierarchy." Essentially, despite there being only two text links to select from in each layer of the hierarchy and the participants being able to read these options within a short period of time, it was felt that anxiety outweighed ease of selection.

5.1.3 Comparison with sighted users

There was a significant difference in the mean search times between the sighted participants who could quickly read the information on the screen and the visually impaired ones who needed to listen to the voice of the screen reader go over each line one at a time. This is an inevitable fundamental factor of the sighted people's ability to process the two-dimensional information instantly, whereas the visually impaired people must use one-dimensional information, in this case, a synthetic voice. This barrier is extremely difficult for finding a solution without the development of innovative technologies for efficient dimension conversion.

However, as shown in Figure 5, the difference in the average evaluation of NASA-TLX for the sighted and the visually impaired was as little as about ten points. These are very interesting results. Despite the visually impaired taking up to seven times as long for search times as the sighted, the mental workload when compared with that of the sighted people did not display a significant difference. That is, shortening the search times of websites for the visually impaired is not necessarily important, and for the visually impaired, it seems more important to produce an easy-to-understand website.

Furthermore, the trends for the magnitude of mental workload for each of the models were similar for the visually impaired and the sighted participants, showing a strong correlation ($r = 0.71$). These results are also very interesting. This suggests that a special hierarchical structure does not need to be considered for the visually impaired, and instead, common structures used for the sighted people can be used.

Thus far, when it came to the visually impaired, to improve browsing on websites such as those for schools for the blind, it was assumed that the user would be using a screen reader, and the implementation of figures and tables was avoided as far as possible with a purely text data website being created separately. However, creating two websites and consistently synchronizing the content between them takes a considerable amount of time and money. Therefore, it is unrealistic to expect general websites to produce separate sites for the visually impaired. The findings of this study suggest that it is possible to build a website with a hierarchical structure that is easy to use for both the visually impaired and the sighted people.

For this study, 64 items of content were evaluated using five different types of hierarchical structures, but even small websites actually in use have several orders of magnitude more content and complicated site structures. Therefore, for further advancement, a website with a larger amount of content and a more complex hierarchical structure needs to be created and examined. Furthermore, the affinity of the language between the search terms and the text links could possibly affect the search efficiency. For example, if a participant is researching "capital city," the determination of whether this word is classified "geography and politics" link or "economy" link may depend on the knowledge and education of the participant. The different thinking processes of the website creators may also play a role. For more accurate assessment of optimum hierarchical structure, examining this type of terminology is also an important issue for the future.

It is thought that for the evaluation indices of mental workload, it is necessary to consider physiological indices, such as any correlations to saliva amylase, heart rate, and sudation (sweating). Furthermore, when the desired content is not found immediately, users get lost while moving through the layers of a website and a long period of time elapses. Thus, "being lost" is expected to have an effect on the mental workload. In a different manuscript (Iizuka et al, 2014), it was reported that this condition of "being lost" or "lostness" occurs while searching for information. Further research on "lostness" along with more detailed analysis of the search times for websites should be conducted.

6. Conclusion

In order to learn how site structure affects information retrieval for the visually impaired, 64 content items were arranged into five categories of websites, ranging from broad and shallow site structures to narrow and deep site structures, and the time that it took visually impaired subjects to search for contents by using a screen reader as well as the mental workload involved was measured. The results showed that for websites having a hierarchical structure with fewer text links on the top page and larger number of contents in the second layer required the least amount of search time and mental workload. This was very similar to the results obtained for the sighted subjects. Furthermore, comparing the two-choice, six-layer narrow and deep site structure with the two-layer site structure, which was the structure that allowed for the shortest possible search time, a difference of almost three times in the search time was found. Moreover, when comparing the results obtained for the visually impaired people to those for the sighted people, although the searches were more time consuming, the differences in terms of the mental workload by hierarchical structure was little. Therefore, it can be concluded that when creating a high-usability website, it is possible to make it efficient with a low mental workload when used by either sighted or visually impaired persons searching for information.

Acknowledgments

While advancing this study, a variety of guidance and advice was received daily, and for this, we would like to express our sincere gratitude to Dr.Akira Ichikawa, professor emeritus of Chiba University, and Dr.Yasuo Horiuchi, associate professor of the graduate school of Chiba University. Furthermore, we would like to thank all the students from Tsukuba Gakuin University, Tsukuba Nursing School, and Tsukuba University of Technology who embraced the opportunity to participate in these experiments.

References

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). *Kit of factor-referenced cognitive tests, auditory number span test (MS-1)*. Educational Testing Service

Haga, S. (2001). *Theory and measurement of mental workload*. Japan Publication Service

Iizuka, J. and Ichikawa, A. (2007). Considerations of efficiency and mental workload of search task on website by the visually impaired persons. *Technical Report of IEICE, WIT2006-114*, 55-60.

Iizuka, J. and Okamoto, A. (2014). Consideration on mental workload of visually impaired persons in web searching (2) Measurement sale of lostness in web searching. *Proceedings of the Human Interface Symposium 2014*, 701-706.

Japan Standards Association (2010). *JIS X 8341-3:2010, Guidelines for older persons and persons with disabilities – Information and communications equipment, software and services – Part 3: Web content*. Japanese Industrial Standards Committee.

Larson, K. and Czerwinski, M. (1998). *Web page design: Implications of memory, structure and scent for information retrieval, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press.

Miller, D. P. (1981). The depth/breadth tradeoff in hierarchical computer menus. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting October*, Vol. 25, No. 1, 96-300.

Nielsen, J. (1999). *Designing Web Usability*. New Riders Publishing.

Snowberry, K., Parkinson, S. R. and Sisson,N. (1983). Computer display menus. *Ergonomics*, Vol. 26, No. 7, 699-712.

WCAG 2.0. (2008). http://www.w3.org/TR/WCAG20/.

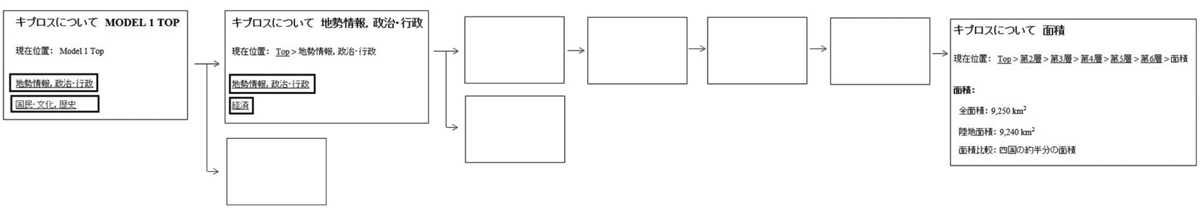
Zaphiris, P. G. (2000). Depth vs. breadth in the arrangement of web links. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 44, No. 4, 453-456.

(Received: December 4, 2015; Accepted: December 26, 2015)

Appendix Hierarchical structures and screenshots

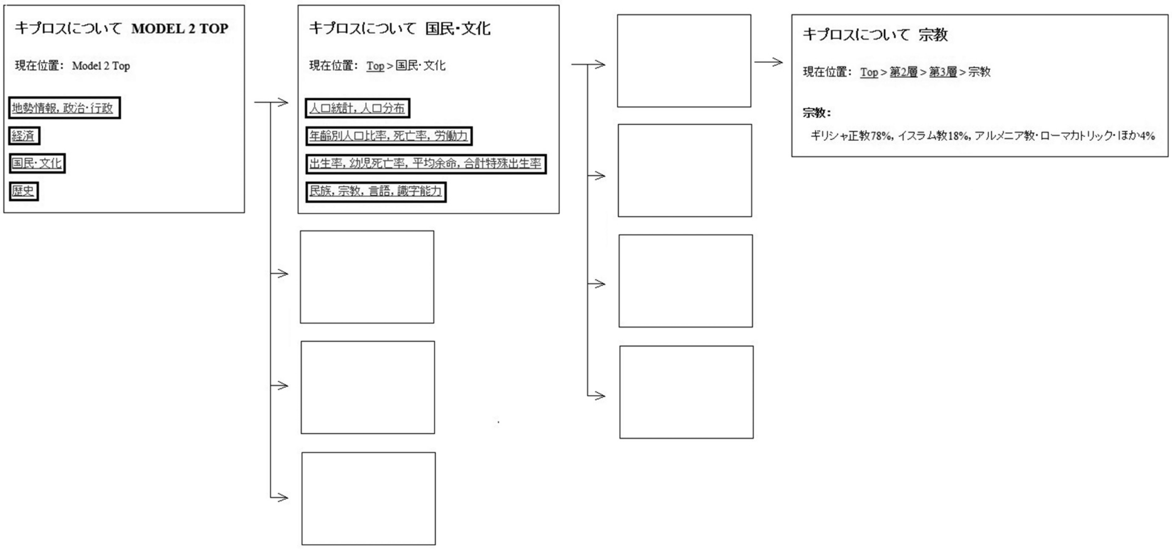
(1) Model 1: Six-Layer Structure of 2 × 2 × 2 × 2 × 2 × 2

Two text links are displayed on the top page, the second layer also has two links, and the two-choice structure continues for six layers. In the last layer (the sixth), there is a description for the content.



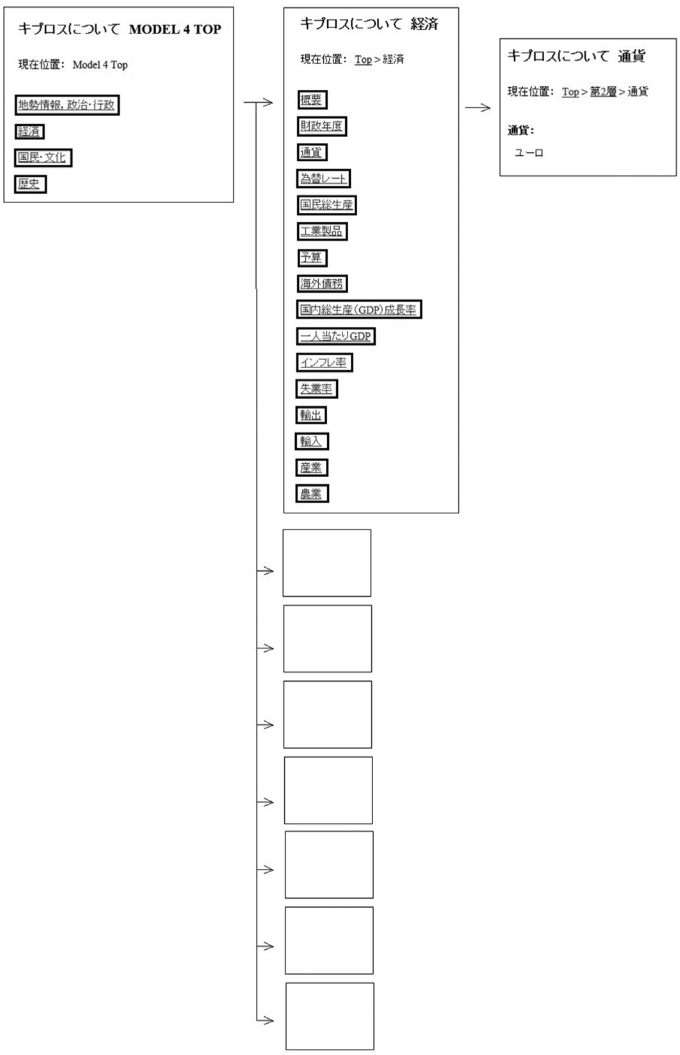
(2) Model 2: Three-Layer Structure of 4 × 4 × 4

Here is Model 2 (4 × 4 × 4). There are four text links on the top page and four text links on layers 2 and 3 with an explanation located on the final layer (fourth).



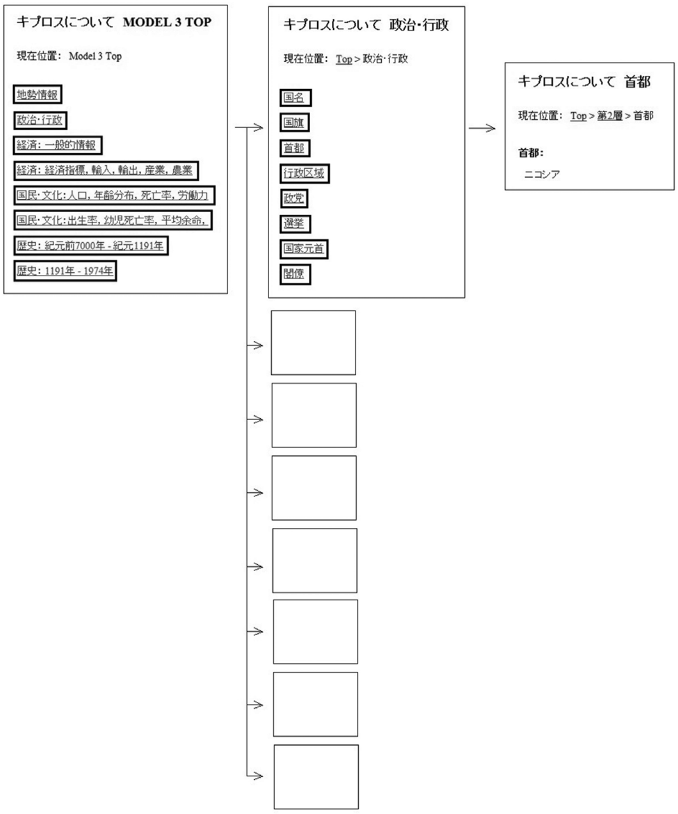
(3) Model 3: Two-Layer Structure of 4 × 16 (left image)

Here is Model 3 (4 × 16). Four text links appear on the top page, and 16 links are on the second layer with the final layer (third) having an explanation.



(4) Model 4: Two-Layer Structure of 8 × 8 (right image)

Here is Model 4 (8 × 8). It has a two-layer structure like Model 3, but the top page has eight text links and the second layer has eight text links. The final layer (third) has an explanation.



(5) Model 5: Two-Layer Structure of 16 × 4

Here is Model 5 (16 × 4). Like Model 3, it has two layers, but the top page has 16 text links and the second layer has four text links. The final layer (third) has the explanatory text.