Behaviour & Information Technology, Vol. X, No. X, Month 200X, xxx–xxx

Information structure and practice as facilitators of deaf users' navigation in textual Websites

I , FAJARDO*[1], J.J., CAÑAS[2], L. SALMERÓN[3] and J., ABASCAL[4]

[1] University of Manchester, UK

[1] University of Granada, Spain

[1] University of Valencia, Spain

[1] University of the Basque Country, Spain

*[1] Manchester Business School, University of Manchester , UK

Booth Street east, M15 6PB, Manchester, UK

email: ifajardo@manchester.ac.uk,

Phone: +44 (0) 161 30 63095

Fax: +44 161 30 63518

Abstract

Deaf users might find it difficult to navigate through Websites with textual content which, for many of them, constitutes the written representation of a non native oral language. With the aim of testing how the information structure could compensate this difficulty, twenty-seven prelingual deaf people users of sign language were asked to search a set of headlines in a Web newspaper where information structure and practice were manipulated. While practice did not affect deep structures (Web content distributed through 4 layers of nodes), wide structures (Web content concentrated in 2 layers) did facilitate users' performance in the last trial block and compromised it in the first trial block. It is argued that wide structures generate a textual information overload for deaf users, which decreases with practice. Thus, wide structures seem preferable for websites requiring frequent use, rather than for those intended for occasional interaction.

Keywords: web accessibility, information search, deafness, cognitive factors

1 Introduction

International organizations such as the International Organization for Standardization (ISO) and the Web Accessibility Initiative (WAI) offer web accessibility guidelines aimed at improving the use of the Internet by disabled people. However, accessibility guidelines often lack empirical validation (Ivory and Hearst 2001) and tend to focus on the physical and sensorial features of their deficiencies (e.g. Seeman 2002), not considering other cognitive, linguistic and cultural aspects.

This situation is particularly true for the guidelines developed to improve web use in prelingual deaf users, generally treated as *hearing users who cannot hear* which can lead to simplistic and often erroneous web design solutions like providing text or graphic equivalents to auditory content (e.g. Fajardo, Cañas, Salmerón and Abascal 2006). For instance, oral language constitutes a second language for a considerable number of deaf people, a fact which compromises the development of reading skills strongly associated to general oral language knowledge (Alegría and Leybaert 1986; see Musselman 2000, for a review). Although there is great variability, the levels of reading skill in the deaf population has been shown to be low in a diversity of cultures (e.g. Asensio, 1989, Alegría 1999, Goldin-Meadow and Rachel 2001, Leybaert, Alegría and Morais 1982), which seems to affect their educational and social integration. Villalba, Ferrer and Asensi (1999) observed that the most significant difference between a group of deaf people who passed university access exams in Spain and a group who failed them was the level of reading skills, similar in the former group to the level of a group of hearing people who also passed. Likewise, the lack of reading skills may affect Internet use making the research on strategies to overcome this situation decisive to ensure deaf people' access to information technology.

2 Information structure and practice as facilitators of Web interaction

A website or hypertext system is composed of a set of graphical, textual or audible information nodes connected by links. Unlike traditional media of information transmission (books, newspapers, journals), the structure of information nodes is not fixed and the user can access them in different ways. As the number of information nodes and links between them can be huge in a hypertext system, one of the main tasks facing the user is to find the information in an acceptable time without becoming disoriented.

Various models of visual search have been used to explain user behaviour in information search tasks in websites and menu-driven interfaces (Liu, Francis and Salvendy 2002, Pearson and Schaik 2003, Scott 1993). If we consider exclusively the visual search component, we may predict that the reading difficulties of prelingual deaf people will not interfere with their information search performance. However, these models appear to be unsatisfactory, because they do not embrace the complexity of web interaction. As was first noticed in the case of menu searches (Norman 1991), the visual search is just one component of any information search task (or 'interactive search task'; Payne, Richardson and Howes 2000): other factors, including lexical-semantic ones, may also be involved (Salmerón, Fajardo and Cañas 2005).

To perform navigation decisions, users must use 'information scents', that is, interpretations of the relevance of local cues such as textual links and images. If the user's goals and a local cue are associated, the local cue will be judged as being relevant to the goal and the user will select it to go forward through the hypertext structure. From this point of view, the use of the semantic content is the core process in the information search. This principle seems to underlie models such as SNIF-ATC (Scent-Navigation and Information Foraging in the ATC architecture), proposed by Pirolli and Fu (2003;

Fu and Pirolli, 2007), whose goal is to simulate users as they perform unfamiliar information-seeking tasks on the Web.

The Comprehension-based Linked model of Deliberate Search proposed by Kitajima, Blackmon and Polson (2000) also incorporates the idea of information scent. The authors describe four cognitive processes as central to the model: parsing, focusing, comprehension and selection. In at least three of these, users' knowledge could be relevant. When parsing, users can apply a top-down process controlled by their prior knowledge of the interface conventions to subdivide a page into manageable and meaningful parts. When comprehending, users elaborate the goal and the web objects under attention, based again on their knowledge of domain or interface conventions. Finally, to go forward users have to select an item judging the degree of relatedness (scent) between it and their elaborated goal. The output of the judgment must satisfy one of three competing constraints: similarity, frequency and literal matching. To satisfy similarity, users will use their semantic knowledge about the domain. Their episodic memory and visuospatial abilities could account for frequency and literal matching strategies respectively. Finally, in the Human-Web Interaction Cycle model, Farris (2003) also proposes that the system knowledge (of content or structure) guides the interaction, showing in a series of experimental data that the domain and the lexical understanding influence how users explore the system.

A diversity of cognitive processes may be involved in the performance of a task with the complexity of a web information search task (e.g. Juvina and van Oostendorp 2004, Larson and Czerwinski 1998, Vicente, Hayes and Williges 1987). For instance, it has been proposed that users' reading skill and vocabulary are strongly related to their capacity to learn from hypertext (Foltz 1996). Foltz suggests that bad readers, whose reading process is not automatic and consumes a great amount of resources, may suffer interference by having to make choices of where to go next as they read. Vicente *et al*. (1987) found that user's reading comprehension and vocabulary were suitable predictors of performance in menu information retrieval tasks. Within a range of verbal factors, Juvina *et al*. (2004) showed that verbal working memory was an indicator of disorientation in different navigation tasks, including information searches. In that sense, Larson and Czerwinski (1998) observed that search task performance correlated more with verbal than with spatial memory span.

We may think that the involvement of verbal abilities in interacting with a system based on textual information like hypertext is simply obvious. However, these abilities may interact with some characteristics of the hypertext structure, changing their effect. This is the case of the breadth and depth structural dimensions (Larson *et al.* 1998, Lee and Tedder 2004). Depth is defined as the number of layers of nodes in the Web structure and breadth as the number of information items located in the same node of the structure. These dimensions may influence the verbal complexity of the hypertext structure and consequently its reading difficulty. In contrast to Lee *et al*. (2004), who argued that deep structures may overload working memory due to a great relational processing demand, Larson *et al*. (1998) considered that the interaction with deep structures would require less verbal memory resources than with wide ones, because the number of choices per node tends to be low (8 items in their experiment).

Although it is not clear how such web structure dimensions and verbal abilities interact, there seems to be a consensus that breadth is preferable to depth in hypertext and menu search tasks (Snowberry, Parkinson and Sisson 1983, Norman, 1991, Fraser and Locatis 2001, Zaphiris, Shneiderman and Norman 2002). The optimum number of choices per node or page differs among studies; for instance, menu structures with 8 choices per page are optimal according to Norman (1991) and Larson *et al*. (1998), whereas Fraser *et al*. (2001) found that an information space with a single page of 48 choices was better (in terms of effort and time) than a deep version of the space which displayed the choices of the four main sections on a main menu that linked to submenus with 3 choices of breadth. The study of Snowberry et al (1993) suggests that even a menu hierarchy of 64 choices of breadth and 1 level of depth makes participants be faster and more accurate than a hierarchy of 2 choices of breadth 6 levels of depth.

Finally, there are other task-related factors which may influence the user's performance, such as the format and familiarity with the labels used as links in the hypertext interface (e.g. McDougall, Curry and Bruijn 1999, Payne *et al*. 2000,) or the amount of practice with the task (Wiedenbeck 1999). Wiedenbeck (1999) utilized a direct manipulation interface (a simulated electronic mail program), which was similar to wide and shallow hypertext structures where the majority of choices are displayed simultaneously on the same screen, manipulating the format of the choices (textual labels, icons, icons plus textual labels). The author found a superiority of textual labels and labels plus icons over the icons-only condition in the first block of 10 tasks, which disappeared in the fourth block of 10 tasks. Wiedenbeck argued that users had learnt the meaning of icons by the fourth block and could take advantage of the positional consistency of items in the interface, which makes the nature of the representation of the item less important. Alternatively, the positional consistency of items could have no effect on deep webs, due to a higher structural complexity which makes them more opaque. In one of the scare cases that this interaction between information structure and practice has been investigated, Snowberry et al (1983) found, accordingly to our previous proposition, a higher search time improvement through practice for the widest hierarchy (64^1) than for the rest of hierarchies (8^2, 4^3, 2^6).

Therefore, the literature review highlights two factors, information structure and practice, which could be interacting and modulating deaf users web search performance but have not been tested directly with this type of users. Thus, with the general aim of finding strategies of improving deaf people navigation in non native language Websites, the experiment reported here was designed to test the influence type of web structure and practice on an information search task performed by deaf people users of sign language. Deaf users were required to search for information in three different Website structures (deep vs. mid-wide vs. wide structures). The performance was measured along two blocks of practice. We expected to corroborate for deaf users the effect of superiority of wide structures over deep ones (Norman 1991; Fraser *et al*. 2001, Zaphiris, Shneiderman and Norman 2002) and were additionally interested in potential interactions with practice (Wiedenbeck 1999).

3 Experiment

The experimental hypotheses were as follows: 1) Wide structures facilitate the information search task for deaf users in terms of percentage of targets found, search time and disorientation; 2) An increase in the amount of practice with the website improves search accuracy and efficiency, especially in the case of wide structures.

3.1 Method

3.1.1 Participants

Twenty-seven prelingual deaf signing people from the Basque Country's Federation of Deaf People's Associations (Euskal Gorrak) participated in this study. The group was composed of 12 women and 15 men whose first language was Spanish Sign Language and whose average age was 25.0 years. The totality of deaf participants had completed Spanish primary education (which consists of 6 academic years from the age of six to twelve years old). They had a mean of 5.1 (SD = 3) years of experience of using computers and spent a mean of 0.7 (SD = 0.8) hours/day using the computer and 0.4 (SD = 0.5) hours/day navigating on the Web. Data from three participants were removed from the analysis owing to unsuccessfully recorded data.

In order to control the potential individual differences in reading skill among the deaf people composing the sample of this study, such a variable was measured. All components of reading skill could not be measured due to time limitations, so we focused on semantic and syntax theoretically highly involved in the process of navigation decisions making as discussed in the section 2 of this document. We used the CLT-Cloze Test (Suárez and Meara 1992) which provides readers with a short text with blanks where some of the words should be and asking them to generate responses to fill in the blanks. Thus, it measures the sensibility of users to syntactic and semantic cues offered by the text.

High scores in the CLT-Cloze Test indicate that readers apprehend the semantic and syntactic structure of the text and select the appropriate words for the blanks. The test consisted of two standard parts (A and B) whose application lasted 35 minutes each. In order to reduce the experiment length, only Part A was used in this experiment, which consisted in 4 texts with 12, 8, 14 and 13 blanks respectively. Each correct word (replacements identical to the original text) scored one point; therefore, the maximum score in the Cloze Test was 47. Although the Spanish Cloze has not been validated with deaf people, others authors have used it before as a measure of reading skills with a sample of Spanish deaf students (Ferrer, Romero, Martinez et al. 2003). CLT-Cloze Test's instructions were given in Spanish Sign Language by an interpreter and, in order to ensure that they had understood them, the participants were asked to perform 1 practice text before the 4 experimental texts.

The participants obtained a mean of 10.2 correct answers (SD = 5.4), that is, scarcely a 22% of accurateness. This measure was used as a covariate variable in the analyses performed to test our hypotheses due to significant correlations between it and some of the Web search's dependent variables as it is described in the section of results.

3.1.2 Design

The study followed a 3 x 3 mixed design, where *Web Structure* was manipulated between-groups (Wide, Mid-Wide and Deep) and *Practice* was manipulated within-subject (Block 1, Block 2 and Block 3). Therefore, deaf users were divided into 3 groups and each performed the information search task under only one of the Web structure conditions through 3 blocks of trials. The dependent variables of the information search task were the average total response time to find a target, the percentage of correct answers (target found) and knowledge acquisition. The total response time was computed for correct answers only and measured the time that a user needed to find each target from its presentation. The average total response time in correct answers was computed by subject and this value was used in the statistical analysis.

3.1.3 Materials

Newspaper Website. With the aim of designing the content and sections of the digital newspaper, the most important Spanish newspapers were reviewed. As a result of this review, 8 main sections and 82 subsections were selected, which were used as the main content of this newspaper (see examples of section labels in Appendix A). The next step was to create three different structures of the Website containing the 90 sections. The three structures were in a hierarchical format: the 8 main sections were at the top of the hierarchy and each one had 12 subsections below it, except the *Opinion* and *Political* sections, which had only 3 and 11 subsections respectively. On the basis of the hierarchical semantic structure described, the sections were organized into three different physical structures. In order to do that we manipulated the dimensions of *Depth* (number of layers of nodes in the Web structure) and *Breadth* (number of information items located in the same node of the structure). As a result of combining such dimensions, we obtained three web structures, named Wide, Mid-Wide and Deep Structure (examples of distribution of section content through the layers in Mid-Wide and Wide Structures can be seen in Appendix A and graphical representations of the 3 structures can be seen in Figure 1).

Insert Figure 1 around here

Wide Structure: The total number of nodes (pages) was 63. This structure had one depth level. The home page had 62 items (sections) of breadth. Pages at level 1 (the first layer of nodes under the home page) had 3 items of breadth.

Mid-Wide Structure: The total number of nodes was 80. This structure was composed of two depth levels (see Appendix B). The home page had 8 items (sections) of breadth. Pages at level 1 had 12 items of breadth (except one with 11 items), of which only 9 (or 8) were clickable. Pages at level 2 had 3 items of breadth.

Deep Structure: The total number of nodes was 104. This structure was composed of three depth levels. The home page had 8 items (sections) of breadth. Pages at level 1 had 3 items of breadth. Pages at level 2 had 3 items of breadth (except one with 2 items). Pages at level 3 had 3 items of breadth.

The researchers generated three headlines with their corresponding texts for each subsection; thus, there were 246 headlines altogether. For instance, one of the 3 headlines in the “Medicine” section was “The risk of auto-medication”.

The sections and subsections of each hypertext node were hyperlinks which allowed users to access a new information node. Finally, each node of the hypertext had a menu button in the top right corner of the screen which allowed users to return to the home page and a back button in the top left corner for going back one navigation step.

A pilot study performed with 30 hearing participants, showed the following percentages of correct answers for deep, mid-wide and wide structures: 83 %, 87 % and 93 %. These results showed that the task was feasible at a high level of performance in the three types of Web structures.

Relatedness judgment task. The relatedness judgement task was used to measure the users' Knowledge Acquisition, that is, to evaluate the knowledge that users had of the newspaper content after the search task or the way in which the users organised the conceptual structure of the newspaper in memory (Acton, Johnson and Goldsmith 1994). Therefore, after the information search task, the participants were required, in accordance with their experience of newspaper reading, to make a judgment about the relationship between 40 pairs of concepts, for instance, "sports" and "culture" or "politics" and "weather forecast", which usually appear in a normal newspaper (we used the names of the sections and subsections pertaining to the experimental material). Five additional pairs of concepts were used as practice trials. The participants used a scale of 6 points (where "1" meant no relationship and "6" a strong relationship) presented in a written format. Two concepts were considered to be related when they were both in the same main section of the newspaper (e.g. "Football" and "Cycling"). If two concepts were not in the same main section, they were considered to be non-related (e.g. "Football" and "Politics"). The users' scores in non-related pairs of concepts were subtracted from related pairs. If a participant had a good level of Knowledge Acquisition about the digital newspaper at the end of the search task, his or her score would be high (close to 5). In this way, we obtained two measures which we called Knowledge Acquisition.

3.1.4 Apparatus

The experiment was administrated by means of the *EWEB* tool (Basque Country University and University of Granada 2002, López 2004) which is composed of three different modules. The Session Design Module supports evaluators in creating experiments in an XML-based language which feature experiment types (within-subject, factorial, etc.), web logs to be captured (time, visited pages, etc.), task models (search task, free navigation) and surveys (questionnaires, card sorting, etc). The Session Monitor and Guidance Module sets the different tasks to be executed by the user for a particular experiment, recording users' navigation data and exporting it to a portable file or directly to a database, preserving the experiment structure in the recorded data. The Session Analysis Module supports the data visualization and analysis (for a more detailed explanation of the tool functioning, see Basque Country University and University of Granada 2002, López 2004).

3.1.5 Procedure

Participants were tested in groups of 4 to 7 on individual computers in the same room. The instructions for the Relatedness Judgment Task, Web Information Search Task and Reading Comprehension Test were translated into Sign Language before each task was performed. Users were randomly assigned to one of the three levels of Web Structure. To begin with, participants performed a practice session (six trials) of the Web information search task in the structure corresponding to their experimental condition, then they completed the experimental session.

Participants were required to search twenty-one headlines in the digital newspaper, that is, seven headlines for each of the 3 Blocks. Each target headline was presented on an independent Web page with a hyperlink to the main menu page. Users had to read it (e.g. "Tom Cruise won a Golden Globe Award") and search for a match by navigating through the section of the hypertext in any order they chose. The twenty-one trials started from the main menu page and the order of presentation was fixed for all of the experimental conditions. If users found the correct headline and clicked on it, an announcement page was displayed, consisting of the message "Correct Answer" and a link to another page, where the next headline to be searched was presented. If users clicked on an incorrect headline, an announcement page with the message "Incorrect Answer" and a link to return to the previous page and continue the search was displayed. The feedback messages were presented in a textual format. The participants had 2 minutes to complete each target search; if the time limit expired, then the next headline was displayed. All of the newspaper's pages had links to the main menu and to the pages that immediately preceded and followed them in the hierarchy. The number of targets found and the time required to complete each search was registered and their means were used as dependent variables of the search task.

The three Blocks of Web information search trials were consecutively completed. Next, participants performed the Relatedness judgement task (the numbers on the scale and the pairs of concepts were presented in a textual format). To finish, they completed Part A of the CLT-Cloze Test.

3.2 Results

Analyses are described as follows. First, we analysed the relation between users' efficiency in the Web information search task and the component of reading skills measured by the CLT-Cloze test, by means of a correlation analysis. Next, we tested main effects and interactions between Web structure and practice.

3.2.1 Correlations between reading skills and Web information search

A summary of Web performance measures is reported in Table 1. We performed three Pearson product-moment correlation analyses, between CLT-Cloze test' scores and percentage of Targets Found, Response Time and Knowledge Acquisition respectively. The participant reading skills correlated significantly with Percentage of Targets Found ($r=0.64$, $p<0.001$) and with Knowledge Acquisition ($r=0.65$, $p< 0.001$). The correlations were positive, that is, the higher the level of the component of reading comprehension measured by the CLT-Cloze test of deaf users, the more targets found (see Figure 2) and the higher the knowledge acquisition about the web structure. Response times, although in the direction expected ($r= -0.23$, $p< 0.27$), did not reach significance levels.

Insert Table 1 about here

3.2.2 Effects of web structures and practice

The first hypothesis stated that wide hypertext structures would facilitate the performance of prelingual deaf people in the Web information task.

Insert Table 2 about here

We performed several ANCOVAs with *Web structure* (Wide vs. Mid vs. Deep) and *Practice* as factors and reading skills as a covariate variable for each dependent variable, except for Knowledge Acquisition, where the factor *Practice* was not introduced, because this measure was obtained only at the end of Block 3 (see Tables 2 and 3 for means and F-test results). The main effect of *Web Structure* was not significant for any dependent variable. The effect of *practice* was significant for correct answers: deaf users found more targets in Block 3 than in Block 1 and Block 2 (Table 3). However, the effect of *practice* was not significant for this variable.

Insert Table 3 about here

As Figure 3 shows, the second-order interaction between *Web Structure* and *Practice* was significant for Percentage of Targets Found but not for Response Time.

On the one hand, looking at the effect of practice on each structure, post hoc comparisons showed that practice facilitated users' performance in wide and mid-wide structures but not in deep structures; that is, users found significantly more targets in Block 3 than in Block 1 when navigating in wide and mid-wide structures but not in deep structures (see Table 2, last two columns), where they found the same percentage of targets along the three blocks of trials. On the other hand, the percentage of targets found by deaf users was higher when searching in deep than in wide and mid-wide structures in Block 1 ($F(1,20)= 7.85$; $MSE=209.16$; $p<0.01$). By contrast, in Block 3, deaf users found around 10% more targets in wide and mid-wide structures than in deep structure, ($F(1, 20)= 5.62$; $MSE= 241.4$; $p<0.03$).

Insert Figure 3 about here

4 Discussion

The goal of this study was to test the facilitative effect of type of hypertext structure and practice on information search tasks in a textual, non native language Website for prelingual deaf users. Regarding the effect of web structures, results showed that the breadth superiority effect (facilitation of wide vs. deep structures) which usually appears in hearing users navigating in native language Websites (e.g. Fraser *et al*. 2001, Snowberry, Stanley and Parkinson, 1983, Zaphiris et al 2002), occurs in deaf users only after some practice with the website. As Larson *et al*. (1998) suggested, this result may be due to an excessive number of textual choices in the wide structure, which overloads the verbal capacity of users in the first trial block. This explanation could be particularly viable in our study since the textual content were in a non native language for the deaf users which would reduce their processing resources. Our wide structure condition had 62 links in the home page, even more than in the case of the wide conditions of Larson *et al*. (16 and 32 respectively), where better verbal span predicted faster search time. In the case of the deep structure, the number of choices in the homepage was 8, so that the users could make selections among a number of alternatives, which did not overload their memory capacity. In fact, the number of choices in our deep structure coincides with the optimum breadth proposed by Norman (1991).

However, the breadth superiority effect did happen for deaf users after several trials. One possible explanation is that the more trials completed, the smaller is the set of choices in the homepage, which could especially favour their performance in wide and mid-wide structures. With practice, added to the reduction of information load to a smaller set of choices, the relational processing demands would be minor in wide than in deep hypertext structures, which according to Norman (1991), augment the uncertainty about the location of a target, consequently increasing the semantic or relational processing demands between nodes and interfering with the performance. Therefore, although our wide structure with 62 choices was superior to the optimum breadth (8 choices), after each new trial users could discard more choices (those already selected). They would use their episodic knowledge on successful and unsuccessful choices per target in earlier trials to guide the subsequent searches. For instance, users may have recognized that a choice led to the current goal in a previous session of trials and then selected it, or vice versa (Payne et al. 2000). Additionally, as suggested by Wiedenbeck (1999), the facilitative effect of practice may be due to the learning of the position of labels after some blocks of trials. If so, then the absence of the practice effect in the case of deep structures could be explained by the fact that learning the position of labels in such structures would be more difficult because they have a more opaque structure. Further research is necessary in order to determine how many trial blocks would be required to obtain a practice effect in the deep web structure used in our experiment.

On the other hand, we measured reading skills in order to control its variability among our deaf participants sample. We observed that for deaf people who use oral language as a second language, reading comprehension levels as measured by the CLT-Cloze test correlates highly (r=0.65) with Web navigation performance, especially in the case of accuracy measures (percentage of target found and knowledge acquisition): the better the deaf user's reading skills, the higher the percentage of targets found and the better the knowledge acquisition.

Users with restricted reading skills, a frequent handicap associated with the prelingual deaf population (Leybaert *et al.*1982, Alegría 1999), could not understand or misunderstood both the verbal labels of the sections composing the websites and the target headlines (which, in this case, were long sentences). For this reason, they could find only a low percentage of them. However, a percentage of the variability in the search task (35%) is not explained by the scores in the Cloze test. As we said before, this test focuses on syntactic and semantic aspects of reading, but it does not measure other important elements such as vocabulary level, which usually predicts reading performance in deaf people (Moores, Kluwin, Johnson *et al.* 1987, in Spanish, Augusto, Adrián, Alegría, et al., 2002) and could account for part of the variance in the hypertext search task. The study of those and other user variables such as verbal and spatial span (Larson *et al.* 1998), interest (Alexander, Kulikowich and Jetton 1994) or word knowledge will be the subject of future research with deaf users.

From an applied point of view, these findings suggest that in order to design web sites with textual content for prelingual deaf people it is important to consider the type of Web structure and practice. Based on our conclusions, we would offer some hypertext design guidelines:

1. When the interaction with a specific website or hypertext is frequent, it is recommended to design wide structures to facilitate the visual search and match process and reduce the uncertainty on the location of a target (Norman, 1991).
2. When the interaction with a specific website or hypertext is incidental or novel, it is recommended to organize the textual information in deep structures, reducing the number of choices per node (no more than 8) to prevent information overload during decision-making.
3. However, guideline 2 implies certain risks:
    a. The uncertainty on the target location is increased (Norman, 1991) because depth augments the number of semantic judgements and the ambiguity of top-level choices. In this case, a reasonable solution may be to locate the most important textual links or information in the shallower layers of the hypertext.
    b. The relational processing demands are augmented (Lee *et al*. 2004). In this case, it is useful to provide users with aids that support the comprehension of the content and the relational processing and to obtain information scents from proximal cues. For instance, it is possible to measure the semantic distance between labels by means of the Latent Semantic Analysis (LSA) technique (Soto 1999, Tamborello and Byrne 2005). LSA is an automatic statistical method for representing the meaning of words and text passages, allowing the comparison of some units of a piece of information (link, paragraph, or whole text) with an adjoining unit of the text to determine the degree to which the two are semantically related. LSA provides a measure of the degree of argument overlap between texts, which is assumed to reflect the level of coherence

between them (Foltz, Kintsch and Landauer 1998). Therefore, by means of this technique designers could select a set of Web labels with low semantic distances among them.

As a future research line, we consider it necessary to properly test the hypothesis of wide Web structures overloading deaf users' working memory and to explore the strengths of prelingual deaf people in some cognitive areas in order to improve their web performance. Potential advantages in visuospatial information processing of deaf signers (e.g. Wilson and Emmorey 1997, Arnold and Mills 2001) could be exploited in websites and hypertext systems by utilizing maps or other visuospatial aids. In addition, the efficiency of videos of sign language as information scent cues in the Web is an essential area for future research which is starting to be exploited (Fajardo, Parra, Cañas et al. 2008). Finally, the results and guidelines derived from them should not be generalize without previous empirical testing to users with other types of disabilities such as blindness or learning disabilities since their cognitive, physical and cultural characteristics greatly vary with respect to deaf population.

Acknowledgments

This research has been partially supported by the COGNIWEB project "Estudio empírico de patrones de acceso a la web para mejorar la accesibilidad cognitiva de usuarios sordos" that was sponsored by IMSERSO/Ministerio de Salud (within the Spanish Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica).The first author has been supported by a pre-doctoral grant from the University of the Basque Country–Euskal Herriko Unibertsitatea.

The authors wish to thank all the deaf people who collaborated in the experiments, and specially the Deaf Associations: Federación Vasca de Asociaciones de Personas Sordas (Euskal Gorrak). A special mention is deserved by Joseba Palacios and all the other people from the company K.I.-Comunic.

Appendices

Appendix A. Example of distribution of section content through the layers Mid-Wide and Wide Structures

A.1

Example of the structure of the *Culture* section in the Mid Wide Structure		
Main Page	Level 1 of depth	Level 2 of depth
CULTURE	**Cinema**	
	Horror	Headline1,Headline2,Headline3
	Drama	Headline1,Headline2,Headline3
	Comedy	Headline1,Headline2,Headline3
	Music	
	Singles	Headline1,Headline2,Headline3
	Concerts	Headline1,Headline2,Headline3
	Bands	Headline1,Headline2,Headline3
	Exhibitions	
	Painting	Headline1,Headline2,Headline3
	Sculpture	Headline1,Headline2,Headline3
	Photography	Headline1,Headline2,Headline3

A.2

Example of the structure of the *Culture* section in the Narrow and Deep Structure			
Main Page	Level 1 of depth	Level 2 of depth	Level 3 of depth
CULTURE	Cinema	Horror	Headline1,Headline2,Headline3
		Drama	Headline1,Headline2,Headline3
		Comedy	Headline1,Headline2,Headline3
	Music	Singles	Headline1,Headline2,Headline3
		Concerts	Headline1,Headline2,Headline3
		Bands	Headline1,Headline2,Headline3
	Exhibitions	Painting	Headline1,Headline2,Headline3
		Sculpture	Headline1,Headline2,Headline3
		Photography	Headline1,Headline2,Headline3

References

ACTON, W. H., JOHNSON, P. J. and GOLDSMITH, T. E., 1994, Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology*, **86**, 303-311.

ALEGRÍA, J., 1999, La lectura en el niño sordo: elementos para una discusión [Reading in the deaf child: elements for discussion]. In: VV. AA. Actes du colloque: *Lenguaje Escrito y Sordera. teóricos y derivaciones practicas*, (Universidad de Salamanca: Publicaciones Universidad Pontificia de Salamanca), pp.59-76.

ALEGRÍA, J. AND LEYBAERT, J. 1986. Adquisición de la lectura en el niño sordo: un enfoque psicolingüístico. Investigación y Logopedia. *III Simposio de Logopedia* (pp. 221-232). Madrid: CEPE.

ALEXANDER, P. A., KULIKOWICH, J. M. and JETTON, T. L., 1994, The role of subject matter knowledge and interest in the processing of linear and non-linear texts. *Review of Educational Research*, **64**, 201-252.

ARNOLD, P. and MILLS, M., 2001, Memory for faces, shoes and objects by Deaf and Hearing Signers and Hearing Non-signers. *Journal of Psycholinguistic research.* **30**, 185-195.

ASENSIO, M. 1989. *Los procesos de lectura en los deficientes auditivos*. Tesis doctoral, UAM. In AA.VV. (1990). Resúmenes de Premios Nacionales de Investigación e Innovación Educativa. Madrid: CIDE, pp. 205-246.

AUGUSTO, J.M., ADRIÁN, J.A., ALEGRÍA, J. and MARTÍNEZ DE ANTOÑANA, R. 2002. Dificultades lectoras en niños con sordera, *Psicothema, 14*, nº 4, pp. 746-753

Basque Country University and University of Granada, 2002, *Cogniweb: Estudio empírico de patrones de acceso a Internet de las personas sordas* [Empirical study of web access profiles to improve deaf users' cognitive accessibility], (Tech. Rep. No. 3). San Sebastian, Spain: Basque Country University.

FAJARDO I., CAÑAS J. J., SALMERÓN L. and ABASCAL J., 2006, Improving deaf users' accessibility in hypertext information retrieval: are graphical interfaces useful for them? *Behaviour & Information Technology,* **25**, 455 – 467.

FAJARDO, I., PARRA, E., CAÑAS, J.J., ABASCAL, J., LÓPEZ, J.M. and GEA, G. 2008, Web Textual Hyperlinks supported with Sign Language Videos, Communication accepted in the *Workshop on Cognition and the Web* to be celebrated in Granada 24-26 April, 2008.

FARRIS, J. S. 2003, The Human-Web Interaction Cycle: A Proposed and Tested Framework of Perception, Cognition, and Action on the Web. PhD thesis, Kansas State University.

FERRER, A., ROMERO, R., MARTÍNEZ, Mª.C., ASENSI, Mª.C. and ANDREU, A. 2003. *Access through support: The Spanish MÀS Module*. Presented to Making Access Succeed. Más Project Conference in Leicester (UK). September 11, 2003.

FOLTZ, P.W., 1996, Comprehension, Coherence, and Strategies in Hypertext and Linear Text. In: *Hypertext and Cognition*, Rouet, J.F., Levonen, J.J., Dillon, A., Spiro, R (Eds.) (Mahwah, New Jersey:LEA), pp. 109-136.

FOLTZ, P. W., KINTSCH, W. and LANDAUER, T. K., 1998, The measurement of textual coherence with Latent Semantic Analysis. Discourse Processes, **25**, 285-307.

FRASER, L. and LOCATIS, C., 2001, Effects of link annotations on search performance in layered and unlayered hierarchically organized information spaces. *Journal of the American Society for Information Science and Technology*. **52**, 1255-1261.

FU, W. and PIROLLI, P. 2007. SNIF-ACT: A Cognitive Model of User Navigation on the World Wide Web, *Human-Computer Interaction, 2*, 355-412.

GOLDIN-Meadow, S. and MAYBERRY, R.I., 2001 How Do Profoundly Deaf Children Learn to Read?, *Learning Disabilities Research & Practice,* **16**, 222-229

IVORY, M.; HEARST, M., 2001, State of the Art in Automating Usability Evaluation of User Interfaces. *ACM Computing Surveys*, **33**, 1-47.

ISO 9241-171 -*Ergonomics of human system interaction-Guidance on software accessibility*. Available online at: http://www.iso.org

JUVINA, I., VAN OOSTENDORP, H., 2004, Individual Differences and Behavioral Aspects Involved in Modeling Web Navigation.In *ERCIM 8th Workshop "User Interfaces For All"*, June 2004, Vienna, Austria, (Vienna, Springer.), pp. 28-29.

KITAJIMA, M., BLACKMON, M. H., & POLSON, P. G. A, 2000, Comprehension-based model of Web navigation and its application to Web usability analysis. In *People and Computers XIV*, S. MCDONALD, Y. WAERN and G. COCKTON (EDS.) (SPRINGER), pp. 357-373.

LARSON, K., CZERWINSKI, M., 1998, Web page design: Implications of memory, structure, and scent for information retrieval, In *Proceedings of CHI'98 Human Factors in Computing System*, (ACM press), pp. 25-32.

LEE, M. J., & TEDDER, M. C. 2004, Introducing expanding hypertext based on working memory capacity and the feeling of disorientation: Tailored communication through effective hypertext design. *Journal of Educational Computing Research,* **30**, 171-195.

LEYBAERT, J., ALEGRIA, J., MORAIS, J., 1982, On automatic reading processes in the deaf. *Cahiers de Psychologie Cognitive*, **2**, 185-192.

LIU, B., FRANCIS, G., SALVENDY, G., 2002, Applying models of visual search to menu design. *International Journal of Human Computer Studies,* **56**, 307-330.

LÓPEZ, J.M., 2004, Development of a tool for the Design and Analysis of Experiments in the Web. In: *Proceedings of The 5th Spanish Human Computer Interaction Conference-Interacción*, 3-7 May 2004*,* Lorés, J., Navarro, R. (eds.):. (Lleida: Springer), pp. 434-437

MCDOUGALL, S. CURRY, M. and de BRUIJN, O., 1999, Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity and semantic distance for 239 symbols. *Behavior Research Methods, Instruments, and Computers,* **31**, 487-519.

MOORES, D., KLUVIN, T., JONSON, R., COX, P., BLENNERHASSERTT, L., NELLY, L., EWOLDT, C., SWEET, C., and FIELDS, L., 1987, *Factors predictive of literacy in deaf adolescents with deaf parents. Factors predictive of literacy in deaf adolescentsin total communication programs*. Prject No. NIH-NINCDS-83-19. Final report to national institute of Neurological and Communicative Disorders and Stroke. Washington, DC:Gallaudet University.

MUSSELMAN, C. 2000. How do children who can't hear learn to read an alphabetic script?: A review of the literature on reading and deafness. *Journal of Deaf Studies and Deaf Education, 5*, 9 - 31.

NORMAN, K., 1991, *The Psychology of Menu Selection: Designing Cognitive Control of the Human/Computer Interface* (Norwood, NJ.:Ablex Publishing Corporation)

PAYNE, S., RICHARDSON, J., and HOWES, A., 2000, Strategic use of familiarity in display-based problems solving. *Journal of Experimental Psychology*: Learning, Memory and Cognition, **26**, 1685–1701.

PEARSON, R., VAN SCHAIK, P., 2003, The effect of spatial layout of and link colour in web pages on performance in a visual search task and an interactive search task. *International Journal of Human-Computer Studies,* **59**, 327-353.

PIROLLI, P., and FU, W.T.F., 2003, SNIF-ACT: a model of information foraging on the World Wide Web. In *Ninth International Conference on User Modeling (UM2003), 22*-26 June 2003, Johnstown, Pennsylvania (Berlin: Springer Verlag), pp. 45-54.

SALMERÓN, L., CAÑAS, J.J. and FAJARDO, I, 2005. Are expert users always better searchers? Interaction of expertise and semantic grouping in hypertext search tasks. *Behaviour and Information Technology,* **24**, 471-475.

SCOTT, D., 1993, Visual search in modern human–computer interfaces. *Behaviour and Information Technolog*y, **12**, 174–189.

SEEMAN, L., 2002, Inclusion of Cognitive Disabilities in the Web Accessibility Movement. In *Proceedings of the Eleven International World Wide Web Conference,* 7-11 May 2002 Honolulu, Hawaii, USA.

SNOWBERRY, K., PARKINSON, S., and SISSON, N., 1983, Browser soar: A computational model of a highly interactive task. In *Proceedings of CHI'92 Human Factors in Computing Systems*, (ACM Press), pp. 165-172.

SOTO, R., 1999, Learning and performing by exploration: label quality measured by latent semantic analysis. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, 15-20 May 1999, Pittsburgh, Pennsylvania, US, pp. 418-425

SUÁREZ, A. and MEARA, P., 1992, *CLT. Dos pruebas de comprensión lectora (Procedim. "cloze")* [Two tests of reading comprehension (Procedim. "cloze")]. (Madrid: TEA Ediciones, S.A).

TAMBORELLO, F., and BYRNE, M., 2005, Information Search: The Intersection of Visual and Semantic Space. In *Extended Abstracts of ACM CHI Conference on Human Factors in Computing Systems ,*2-7 April 2005, Portland, Oregon, US., *pp.* 1821-1824

VICENTE, K. J., HAYES, B. C., and WILLIGES, R. C. 1987, Individual differences in computer-based information retrieval, In *Ergonomics and Human Factors*, L. S. Mark, J. S. Warm, and R. L. Huston (Eds.), (New York: Springer-Verlag),.pp. 225-229.

VILLALBA, A., FERRER, A. and ASENSI, C. 1999. La comprensión lectora en universitarios con déficit auditivo. *Logopedia, Foniatría y Audiología*, XIX (1), 33-39.

WAI-Web Accessibility Initiative, 1999, Web Content Accessibility guidelines 1.0. Available online at http://www.w3.org/WAI/.

WIEDENBECK, S., 1999, The Use of Icons and Labels in an End User Application Program: An Empirical Study of Learning and Retention. *Behaviour and Information Technology,* **18**, 68-82.

WILSON, M. and EMMOREY, K., 1997, Working memory for sign language: a window into the architecture of the working memory system. *Journal of Deaf Studies and Deaf Education,* **2**, 121-130.

Table Caption

Table 1 Means and Standard Deviation of each measures of web information search performance for deaf users.

Table 2 Adjusted Means (with reading comprehension scores) and Standard Deviations of correct answers, response time and knowledge acquisition in each level of web structure and block of practice.

Table 3 F-test results of Information Structure and Practice for each dependent variable. Only significant simple effects are included.

Table 1

Dependent Variable	M	SD
Percentage of Targets Found	60.1	14.9
Response Time (s)	15.1	6.8
Knowledge Acquisition*	0.7	1.1

*Ranging from -*5* or low acquisition to *5* or high acquisition

Table 2

% Target found	Block 1		Block 2		Block 3				Block1 x Block 3	
Structure	M	SD	M	SD	M	SD	Total	N	F	p
Wide	33.3	16.0	58.7	18.1	79.4	17.7	68.5	9	81.3**	. 001
Mid-Wide	32.1	18.3	58.9	20.8	76.8	18.6	69.0	8	67.9**	.001
Deep	61.2	13.6	75.5	22.9	67.3	13.6	80.2	7	1.1	.3

Response Time (s)	Block 1		Block 2		Block 3			
Structure	M	SD	M	SD	M	SD	Total	N
Wide	42	16	53	14	43	12	56	8
Mid-Wide	59	41	56	14	55	17	75	7
Deep	78	99	56	49	43	44	60	7

Knowledge Acquisition			
Structure	M	SD	N
Wide	0.65	0.91	8
Mid-Wide	0.54	0.31	9
Deep	0.92	1.57	7

Table 3

Information structure	df	MSE	F	*p*
Percentage of Targets Found	2/20	455.5	.1	.9
Response Time (s)	2/18	.4	.28	.8
Knowledge Acquisition	2/20	.8	.31	.8
Practice	**df**	**MSE**	**F**	***p***
Percentage of Targets Found	2/42	159.7	40.7**	.001
Block 1 x Block 2	1/21	186.7	31.2**	.001
Block 1 x Block 3	1/21	117.3	105.4**	.001
Block 2 x Block 3	1/21	172.2	7.1*	.015
Response Time (s)	2/38	.08	1.1	.4
Information structure x Practice	**df**	**MSE**	**F**	***p***
Percentage of Targets Found	4/42	158.7	6.4**	.001
Response Time (s)	4/38	.8	1	.4

Figure Caption

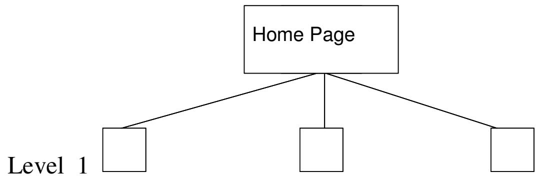
Figure 1 Sketches of the depth levels of the three experimental web structures. The Wide structure home page had 62 information items and each page at depth level 1 had 12 information items. The Mid-Wide structure home page had 8 items, pages at level 1 had 12 items and pages at level 2 had 3 items. The Deep structure home page had 8 items and pages at levels 1, 2 and 3 had 3 items each.

Figure 2 Percentage of correct answers as a function of reading comprehension abilities scores.

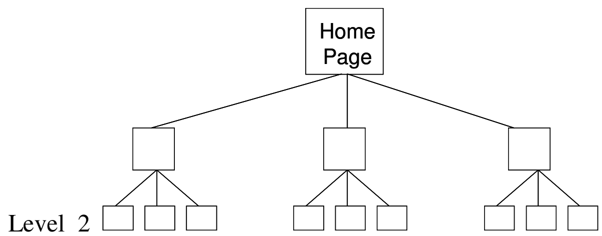
Figure 3 Interaction between Block of Practice and Web structure for the percentage of target found measure. Wide and Mid Wide structure had a facilitative effect only in the third blocks of trials.

Figure 1

B.1 Wide Web structure (62x3)



B.2 Mid-Wide Web structure (8x12x3)



B.3 Deep Web structure (8x3x3x3)

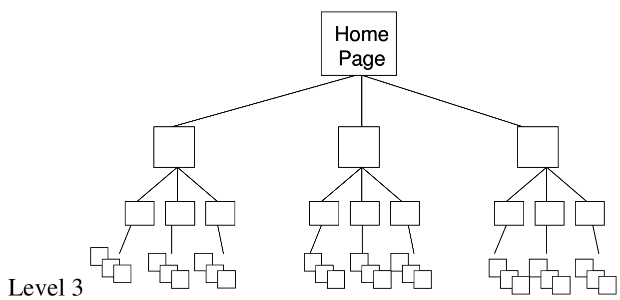


Figure 2

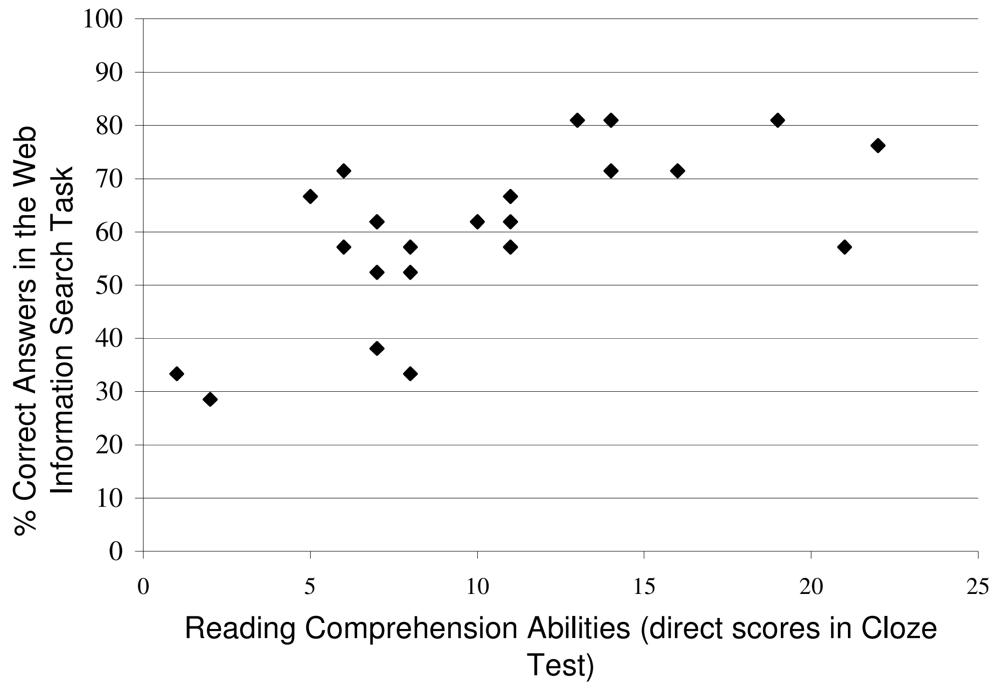


Figure 3