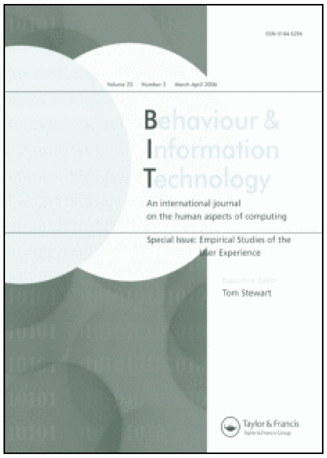


Behaviour & Information Technology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713736316

Improving deaf users' accessibility in hypertext information retrieval: are graphical interfaces useful for them?

I. Fajardo [a]; J. J. Cañas [b]; L. Salmerón [b]; J. Abascal [a]
[a] University of Basque Country. Spain
[b] University of Granada. Spain

To cite this Article: Fajardo, I., Cañas, J. J., Salmerón, L. and Abascal, J. , 'Improving deaf users' accessibility in hypertext information retrieval: are graphical interfaces useful for them?', Behaviour & Information Technology, 25:6, 455 - 467
To link to this article: DOI: 10.1080/01449290500331180
URL: http://dx.doi.org/10.1080/01449290500331180

Behaviour & Information Technology, Vol. 25, No. 6, November – December 2006, 455 – 467

Improving deaf users' accessibility in hypertext information retrieval: are graphical interfaces useful for them?

I. FAJARDO*†, J. J. CAÑAS‡, L. SALMERÓN‡ and J. ABASCAL†

†University of Basque Country, Spain
‡University of Granada, Spain

This paper explores the effect of substituting textual links for graphical ones on the performance of deaf signers in hypertext information retrieval (HIR). Both deaf and hearing users found more targets, were faster and became less disoriented in the *verbal* hypertext interface than in the *graphical* one. Deaf users were outperformed by hearing users in all conditions except in short paths with the graphical interface. The results and its applied consequences, which would be also relevant to other users with similar problems than those of deaf signers (elderly people, people with dyslexia, people navigating in a website using a foreign language or people with low literacy) are discussed in relation to the CoLiDeS model of web interaction (Kitajima *et al.* 2000) and to the overgeneralisation of 'Picture superiority effect' (Nelson *et al.* 1976).

Keywords: Web accessibility; Graphical user interface; Deafness; Semantic memory

1. Introduction

Web accessibility is the degree to which the Internet and its services are put at 'the disposal of all individuals, whatever their hardware or software requirements, their network infrastructure, their native language, their cultural background, their geographic location, or their physical or mental aptitudes' (Berners-Lee W3C). In the digital age, the web has become established as a fundamental repository of information and services but, aside from its ancient good intentions to facilitate information search (Bush 1945), it is currently far from being accessible to all users.

There exist diverse international organisations of standardisation (e.g. ISO, WAI) offering guidelines for providing accessible web designs. These guidelines are often insufficient and unsatisfactory, apart from the lack of empirical validation (Ivory and Hearst 2001). The shortage of satisfaction stems from an almost exclusive focus on the physical and sensorial features of the deficiencies (e.g. Seeman 2002), forgetting the cognitive limitations than they can involve. This is the case of the pre-locutive deafness (Marschark 2003) that will be the focus of this work.

We have started a research programme (named Cogniweb) with the pragmatic objective of finding ways to improve the accessibility for deaf people to the web or to hypertext systems, terms that we use as synonyms in this document. Our previous research has already shown that *deaf signers users*, those deaf people who use sign language as their primary and preferred language (Emmorey 2002), are inefficient and become disoriented while finding information in hypertext (Fajardo *et al.* submitted). We found that deaf signers' low reading comprehension abilities and their low prior knowledge on the topic were the main factors contributing to their performance.

In the research described in this paper we tried to test whether using graphical information could eliminate the problems deaf signer people have with verbal information. Different models proposed to explain hypertext interaction (e.g. Pirolli and Fu 2003, Kitajima *et al.* 2000, Farris 2003) consider that prior knowledge on the domain or interface conventions is rather important in the process of information search. This knowledge – mainly stored in the semantic memory of users – must be retrieved while navigating, in order to be used. Therefore, it is possible that the format in

*Corresponding author. Email: acbfabri@si.ehu.es

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01449290500331180

which information is presented in the hypertext (graphical or verbal) could affect the access to semantic knowledge and consequently, hypertext performance. The access to this knowledge seems to be more efficient through graphical clues, as shown by the almost universally accepted assumption that pictures access more readily to semantic memory than words, the effect known as *Picture Superiority* (Nelson *et al.* 1976). Therefore, replacing verbal information with the graphical equivalent could enhance deaf signers' interaction with hypertext.

However, there is also empirical evidence not supporting the picture superiority effect (PSE) in the literature of cognitive psychology (e.g. Snodgrass and McCullough 1986, Amrhein *et al.* 2002). In the field of human–computer interaction (HCI), the results regarding the comparison of graphical and verbal interfaces are also inconsistent (e.g. Dillon and Song 1997, Weidenbeck 1999), in spite of the generalised belief in the high efficiency of the human visuospatial processing (Landsdale 1988, Blankenberger and Hahn 1991, Dillon and Song 1997). In addition, individual differences in semantic memory could be reflected in hypertext performance and even interact with information format. In fact, deaf signers differ from their hearing peers in the access strategies and organisation of knowledge of semantic memory (e.g. Marschark 1998, McEvoy *et al.* 1999). Therefore, it seems necessary to take a closer look at the research that has been done on these topics to frame our work.

2. Literature review

2.1 *Semantic processing in hypertext*

The importance of semantic processing is a constant in different models of web or hypertext information retrieval (e.g. SNIF-ATC, CoLiDeS or HuWI). A hypertext system is composed of a set of graphical, textual, auditive or haptic information nodes connected by links. Unlike traditional linear texts, users can access information in different ways. As the amount of information nodes and links may be huge in a hypertext system, one of the main tasks of users is to find the information without becoming disoriented and within an acceptable time.

According to Kitajima *et al.* (2000) the core process underlying HIR is the comprehension of text and images. CoLiDeS, the model proposed by these authors, describes HIR as a reiterative process of parsing, focusing, comprehending and selecting information. The usage of semantic knowledge is rather important in all these processes. For instance, with the aim of defining their search goal in the comprehending process, individuals use their knowledge on the domain or interface conventions stored in semantic memory. In the selection process, individuals compare their elaborated goal with the available choices trying to satisfy one of three constraints that compete among them: similarity (semantic similarity between target and choices), literal matching (the target and an interface element match perfectly) and frequency (with which the user has encountered the elements in a path). As Kitajima *et al.* explain, the similarity comparison is a semantic comparison, that is, this process relies highly on semantic knowledge on the domain stored in users' long-term memory (LTM) – the permanent store of information of the cognitive system.

In agreement with the theoretical models, empirical data support that general vocabulary knowledge or prior knowledge on the domain, predicts success on information retrieval (Vicente *et al.* 1987, Salmerón *et al.* 2005, Fajardo *et al.* submitted). Results show that normally high reading skills and high prior knowledge are associated with better performance. Additionally, the structure of information in hypertext could influence the difficulty of the semantic similarity judgement involved in HIR (in terms of CoLiDeS). The higher links in a tree structure can vaguely identify the user goal, especially if the user goal is not placed in the higher nodes of the tree (Norman 1991). As predicted by network models of semantic memory, this semantic ambiguity would increase as a function of the distance between the user goal and the higher nodes of the hierarchy (e.g. Collins and Loftus 1975). Therefore, we can expect that performance will drop in lower levels of the hypertext structure.

User semantic knowledge could be differentially elicited by the available information in the hypertext (proximal cues or objects). Theoretically, pictures (e.g. graphical icons) are superior to words (e.g. verbal links) in a semantic task such as the 'similarity judgement' involved in hypertext search tasks (CoLiDeS). In the next section, we present a brief revision of picture effect in both basic cognitive research and HCI research.

2.2 *Pictures vs. words in cognitive psychology*

The PSE has been reported in a great variety of semantic tasks (e.g. Pellegrino *et al.* 1977) and episodic tasks (Paivio and Csapo 1973, Kinjo and Snodgrass 2000). The Sensorial–Semantic Model proposed by Nelson *et al.* (1976) is one of the models that tries to explain this effect and postulates an abstract and a-modal store of semantic representations. Pictures, in addition to having a larger sensorial distinctiveness, may access this abstract store faster than words, which must be phonologically preprocessed. Dual Code Theory (Paivio 1971, 1991) brings an alternative explanation to PSE as it postulates that verbal and pictorial information would be represented in different formats in LTM. One of Paivio's explicative hypotheses is that pictures would have more possibilities of being represented in both formats due to a referential processing: a picture would evoke a verbal label more easily than a word would evoke a picture (Paivio 1977, 1991). For this reason, pictures may facilitate performance in recognition or categorisation tasks.

However, there is inconsistent evidence on PSE in different types of semantic tasks. For instance, in a single-stimulus classification task (classify stimuli in one of two categories), Snodgrass and McCullough (1986) demonstrated that pictures superiority disappear when the two categories were visually similar (e.g. fruits and vegetables). These results suggest that pictures will not facilitate a faster access to semantic representations, PSE being a visual and not a semantic effect. However, Lotto *et al.* (1999) suggested that the results obtained by Snodgrass and McCullough may be due to an artefact as visual similarity and conceptual similarity of pictures vary normally together. They showed that when the effects are separated experimentally, the conceptual similarity of between-category affected both pictures and words while visual similarity only affected pictures, reversing the effect of superiority (words better than pictures). In opposition to Snodgrass and McCullough, these data suggest that the PSE could be both semantic and visual. However, in order to observe the PSE, a high visual and conceptual dissimilarity between categories may be necessary. In the context of HIR, these findings would mean that the effectiveness of icons might be conditioned by both the degree of semantic processing demanded by the task and the visual distinctiveness among icons (McDougall *et al.* 2000).

2.3 *Graphical interfaces effects in HCI*

The facilitation of graphical information has also been observed in more complex tasks such as programming (Navarro-Prieto and Cañas 2001), text comprehension (see revision of Levie and Lentz 1982, Kruley *et al.* 1994), text comprehension from multimedia (Gyselink *et al.* 2002) and information retrieval in database (Blakenberger and Hahn 1991, Dillon and Song 1997).

On the other hand, other authors have found that graphical information (e.g. icons) shows disadvantages or no differences with regard to verbal information in complex tasks (Guastello *et al.* 1989, Benbasat and Todd 1993, Wiedenbeck 1999). Although the separation between the inherent properties of pictures or their implementation in context is not clear, the results of diverse authors point out that the effect of pictures may vary due to different variables related to factors such as the configuration and location of pictures in hypertext (Byrne 1993), distinctiveness (Arend *et al.* 1987), visual grouping (Niemlä and Saarinen 2000), articulatory distance (Hutchins *et al.* 1986, Blankenberger and Hahn 1991) or familiarity (McDougall *et al.* 1999). Benbasat and Todd compared icons and text in two types of interfaces: direct manipulation and menus. They found no differences between icons and text, but direct manipulation interfaces were better than menus. The authors suggest that these two variables are frequently confused in HCI research as icons tend to be used in direct manipulation interfaces while text is used in menus. However, in a direct manipulation interface, a simulated electronic mail program, Wiedenbeck did find a superiority of textual labels or labels plus icons with regard to the usage of only icons.

In addition to an unclear influence of pictures versus words and graphical versus verbal interfaces, individual or group differences in the organisation or access strategies to knowledge on semantic memory may influence the effect of pictures and words in HIR. In the next section we go into the difference in semantic memory between deaf signers and hearing people.

2.4 *Organization and access of LTM knowledge in deaf people*

Several empirical findings induce us to think that there are differences between deaf signer people and hearing people in the organisation and access strategies to the knowledge stored in LTM (e.g. Marschark 1998, Marschark and Everhart 1999, McEvoy *et al.* 1999).

To the extent that semantic knowledge is involved in HIR, the particularities of semantic processing in deaf signer people may be affecting their performance. However, we do not know if these particularities affect them in a positive or negative way or if they interact with the information format or organisation in the hypertext interface. As suggested by Marschark (1998), it is necessary to demonstrate if such differences *are so large that they qualitative or quantitative affect learning in any real sense* (p. 87).

3. Conclusions and predictions

As stated above, our pragmatic objective was to improve the accessibility of deaf people to hypertext. If deaf people have problems with textual information, we could substitute it for graphical information. Theoretically, accessing semantic information is necessary to perform the different phases of an HIR (Kitajima *et al.* 2000), which could be faster with pictures than with words and, perhaps, the only alternative for deaf people. Therefore, this solution should benefit both deaf people and people without problems with textual information processing. However, the picture superiority effect does not always happen, neither in traditional semantic tasks nor HCI tasks. In addition, the location of the targets in the hypertext structure may augment the number of selections and the difficulty of the semantic similarity judgement. Finally, differences between deaf and hearing people in the semantic knowledge organisation or access strategies could affect HIR.

Therefore, due to this unclear empirical evidence we thought it necessary to perform one experiment to test whether it was true that graphical hypertext interfaces improved performance of deaf people in HIR and if there was an interaction between the type of interface (textual of graphical) and the type of user (deaf and hearing users) as well as the knowledge they have. Furthermore, we included path length as another factor in the experiment. The path length to find a target in the hypertext structure may affect the complexity of semantic processing (the longer the path, the higher the number of semantic decisions and the semantic ambiguity is) and could interact with the variables of interest: type of interface and type of users.

4. Experiment

4.1 *Method*

4.1.1 Participants. Twenty-one deaf signer users (DS) participated in the experiment from the Federations of Deaf People Associations of the Basque Country (Euskal Gorrak) and Granada (FAAS) and 24 hearing users (H) from the University of Granada (in exchange for experimental credits or an economical remuneration). The DS group was composed of 10 women and 11 men and the H group was composed of 18 women and six men. The first language of the DS was Spanish Sign Language (with the Spanish acronym *LSE*).

4.1.2 Design and Material. The study followed a 2 × 2 × (3) quasi-experimental design. The independent variables were *interface format* (graphical vs. verbal), *type of user* (DS vs. H users) and *path length* (short, medium and long). The dependent variables of web accessibility were the percentage of *correct answers* (targets found), the *response time* (total time to find the target from the homepage) and the *disorientation* in the hypertext structure measured with the formula of *lostness* (Smith 1996). The variable *path length* was referred to the minimal number of nodes that the users have to visit to find the target. In this way, in short, medium and long paths the users had to visit three, four or five nodes, respectively, to find the targets. We used the hypertext of a newspaper built for a previous research (Fajardo *et al.* submitted). The newspaper was composed of eight main sections and 82 subsections. The total number of nodes was 52. The structure of the hypertext was hierarchical, with five levels of depth and three items of weight per node.

In the graphical interfaces the verbal links were substituted by icons. The icons were selected from a set of 153 in a previous study of semantic distance, that is, how direct the relational graphic representation – function is (Hutchins *et al.* 1986)[1], where a group of 28 individuals (11 DS and 17 H), different to the experimental group, participated. This study was performed in advance of the main experiment to drive design of the graphical interface. The set of 153 icons was extracted from the Internet using the *Google* browser (option of picture retrieval). For each verbal link (representing a section of the newspaper), three icons in black and white were selected. The participants in the previous semantic distance study had to answer in a 5-points scale, where 1 meant low relational icon-referent and 5 meant high relational icon-referent (see figure 1).

We selected the sets of icons with higher relational icon-referent, that is, with lower scores in semantic distance (see Appendix A). The semantic distance average of this set was 3.75 (sd = 0.48). There was no significant difference in semantic distance between DS and H users. In figure 2, we can see a sample of a hypertext node in both graphical and verbal conditions.

4.1.3 Task and procedure. Users read the general instructions of the experiment (in the case of DS, the instructions were explained in sign language). The main task of the users was to find sections of a newspaper implemented in hypertext. Previous to this experimental search task, the participants completed a *relatedness* judgement task with the aim of evaluating and controlling the prior knowledge on the concepts that composed the newspaper and the relation between them. In this task, users had to evaluate in a 6-points scale the relationship between 55 pairs of concepts extracted from the hypertext (words or icons depending on the experimental condition). '1' meant a low relationship between concepts and '6' meant a high relationship. Once the *relatedness* judgement task was finished, users were asked to search for 12 targets in the newspaper hypertext (four per level of path length). After reading the instructions of the search task (in the case of DS, the instructions were explained in sign language), users completed a session of training in the search

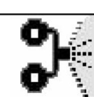
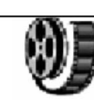
Concept	***Cinema***		
Pictures			
Picture–Concept relation From 1 (low) to 5 (high)			

Figure 1. Example of the Distance Semantic Questionnaire. Users had to indicate below each icon, the degree of relationship between the graphic representation and the concept that appeared above (e.g. Cinema). 'Low relation' (1) meant long semantic distance and 'high relation' (5) meant short semantic distance.

[1]The concept of 'semantic distance' has different meanings depending on the context. In icon usability research, it is used in the sense defined by Hutchins *et al.* (1986), while in the area of semantic memory research the concept is referred to the degree of relatedness between the nodes of a net. In the present document, we use both, but which of the two meanings we are referring to it is always indicated.

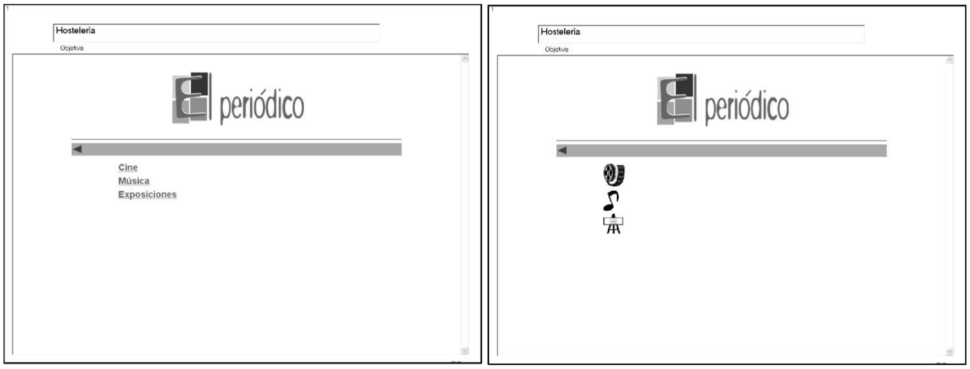


Figure 2. The left picture represents a node of the digital newspaper with verbal links. In the right picture the verbal links were substituted with graphical links.

task supervised by the researcher. Each target was presented individually in the format corresponding to the experimental condition of the user (icon or word) and the search started from the main menu. For example, in the condition *verbal interface*, users received the message 'Find the next target: Cinema'. The target *cinema* was in a short path; therefore, the user had to visit three nodes: 'menu -> culture -> cinema'. Users had one minute to find each target. The same order of target presentation was used in the graphical and verbal interface.

Next our experimental hypotheses are operationalized:

H1: DS users are more efficient in HIR (in terms of targets found, total response time and disorientation) in the graphical interface than in the verbal interface.
H2: Path length negatively affects performance of both types of users in HIR. The longer the path to find targets, the smaller the average of targets found is, the bigger the response times and disorientation is.
H3: As path length affects the semantic similarity judgement and pictures improve such process, the performance in the graphical interface will be less affected by the increase of path length than the performance in the verbal interface.

4.2 *Results*

We performed an analysis of variance (ANOVA) by item instead of by subjects for each dependent variable (correct answers, total response time and disorientation) to filter the possible variability between targets. *Interface format* (graphical vs. verbal), *type of user* (DS vs. H users) were introduced in the analysis as intra-item variables and *path length* (short, medium and long) as between-items variables.

4.2.1 Type of users by type of interface format. The effect of interface was significant for Correct Answers $F(1.9) = 31.31$; Ms_e; 199.3; $p < .0003$) (see figure 3), Response Time ($F(1.8) = 13.09$; $Ms_e = 56, 7$; $p < .0068$) and Disorientation ($F(1.8) = 29.90$; $Ms_e = 0.02$; $p < .0006$).

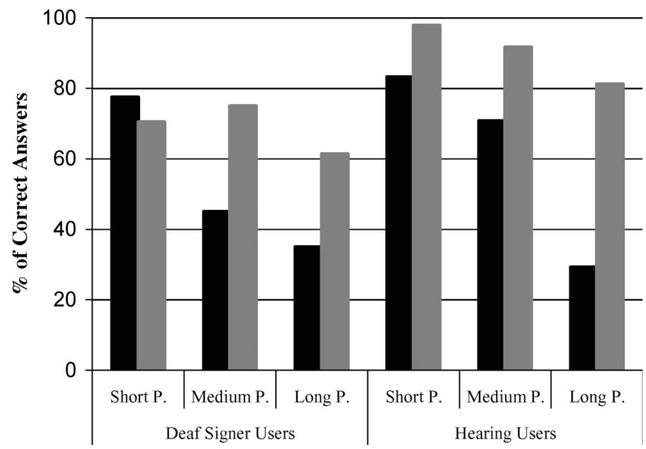


Figure 3. Percentage of correct answers in each path length for both deaf signer users (left) and hearing users (right). The black lines represent performance in the graphical interface and the grey lines represent performance in verbal interface.

However, contrary to H1 both types of users found less targets (57% vs. 80%), were slower (26″ vs. 18″) and became more disoriented (0.3 vs. 0.1) in the graphic interface than in the verbal hypertext interface. The difference among users was significant for correct answers ($F(1.9) = 8.51$; $Ms_e = 316$; $p < .0171$), and response time ($F(1.8) = 10.13$; $Mse = 48$; $p < .0130$). DS users found fewer targets (61% vs. 76%) and were slower (25″ vs. 18″) than H users. There was no interaction between interface format and type of user for any dependent variable.

4.2.2 The effect of path length. In table 1, we can see the average score for each experimental condition in each dependent variable. Supporting our H2, the effect of path

Table 1. Averages of Correct Answers, Response Time and Disorientation in the HIR task for each experimental condition.

		Correct Answers		Response Time		Disorientation	
Users	Path Length	Graphical I.	Verbal I.	Graphical I.	Verbal I.	Graphical I.	Verbal I.
Deaf	Short	77.5 (18.9)	70.5 (20.2)	22.3 (14.8)	14.3 (2.2)	0.3 (0.3)	0.07 (0.1)
	Medium	45 (20.8)	75 (26.1)	27.1 (17.7)	18.4 (4.2)	0.2 (0.1)	0.05 (0.1)
	Long	35 (26.5)	61.4 (23.9)	42.4 (4.7)	26.6 (3.5)	0.5 (0.1)	0.1 (0.2)
Hearing	Short	83.3 (0.0)	97.9 (4.2)	11.2 (3.3)	10.4 (3.7)	0.16 (0.1)	0.1 (0.1)
	Medium	70.8 (14.4)	91.7 (6.8)	19.4 (10.7)	14.8 (4.3)	0.2 (0.2)	0.1 (0.1)
	Long	29.2 (24.1)	81.25 (12.5)	33.4 (4.0)	21.8 (8.3)	0.5 (0.1)	0.2 (0.1)

length was significant for correct answers ($F(2.9) = 4.77$; $Ms_e = 801$; $p < .0388$) and response time ($F(2.8) = 9.19$; $Ms_e = 103$; $p < .0085$). According to our predictions, users found more targets in shorts than in medium and long paths (82% vs. 71% vs. 52%). In addition, users were faster in short paths than in medium and long paths (15″ vs. 20″ vs. 31″). The effect of path length was not significant in the case of disorientation in spite of the pattern of results agreed with our hypothesis, that is, the users became less disoriented in short than in medium and long paths (0.15 vs. 0.16 vs. 0.3).

In the case of correct answers, a significant interaction of 'interface format x path length', $F(2.9) = 6.41$; $Ms_e = 199$; $p < .0186$, showed that path length affected performance only when users search on graphical interfaces, but not when doing so on verbal interfaces. Therefore, the prediction of H2 was not totally supported as, in spite of differences existing between the levels (long = 71%; medium = 83%; y short = 84%), the effect of path length was not significant for the verbal interface. However, for the graphical interface, contrarily to H3, which predicted a smaller effect of path length over this type of interface, the difference between long paths (32%) and medium (58%) and long paths (80%) considered together the result was significant ($F(1.9) = 11.2$; $Mse = 658$; $p < 0.01$). Furthermore, the difference between interfaces was only significant in medium (IG = 58% vs. IV = 83%) and long paths (IG = 32% vs. IV = 71%) respectively, $F(1.9) = 13$; $Mse = 199$; $p < 0.005$ y $F(1.9) = 30.9$; $Ms_e = 199$; $p < 0.000$), but not in short paths (IG = 80% vs. IV = 84%).

A significant 3-way interaction, interface format x path length x type of user for correct answers ($F(2.9) = 5.24$; $Ms_e = 64$; $p < .0310$) pointed in the same direction. For H users, performance on the verbal interface was always better than in the graphical, although the differences were only significant in medium ($F(1.9) = 9.2$; $Ms_e = 85$; $p < 0.01$) and long paths ($F(1.9) = 57.4$; $Ms_e = 95$; $p < 0.000$). Weakly supporting our H1, DS users found more targets in the short paths with the graphical interfaces than with the verbal interface (although the difference was not significant). The opposite pattern was found in the rest of levels of path length, the differences between interface formats being significant (medium path = $F(1.9) = 10.3$; $Ms_e = 174$; $p < 0.01$ and long path = $F(1.9) = 8$; $Ms_e = 174$; $p < 0.02$). Therefore, it seemed that the advantage of verbal interfaces over graphical appears when semantic processing becomes more difficult, for both types of users.

In the case of response time and disorientation, there were no interactions of path length with any other independent variable. That is, response time and disorientation increased in parallel to the path length for both types of interfaces and users. It is possible that theses variables were only sensitive to the physical increase of the number of pages in long paths and not to the augmenting semantic decision complexity.

4.2.3 Prior knowledge. In the relatedness judgement task, users had to judge two types of concept pairs: related and non-related. The scale goes from 1 (low relation) to 6 (high relation). The scores in non-related pairs were subtracted from the scores in related pairs. In this way, values close to 6 were interpreted like high prior knowledge and values close to 0 were interpreted like low prior knowledge. We used these values as a dependent variable of an ANOVA with types of users and formats of judged concepts (icon vs. word) as independent variables. There was a significant effect of concept format, $F(1.37) = 38.11$; $MS_e = 0.43$; $p < .0001$. All users in general had better prior knowledge on the verbal labels (0.42) than on the icons (1.7).

In contrast to what could have been expected, derived from the theoretical revision of deaf peoples' memory functioning, there were neither significant differences between DS and H users in prior knowledge nor interaction with the type of prior knowledge evaluated (icons or verbal labels). However, we analysed the differences in prior knowledge between users for each pair of items (icon or word) instead of using the average scores of all items. We performed two multivariate analyses of variance (MANOVAs) to compare the previous knowledge on the relation between each pair of concepts (see results in figure 4 and appendix B).

One MANOVA was performed for related pairs of concepts and another one for unrelated pairs. In the case of unrelated pairs, we found a significant interaction between the type of user and the type of related pair,

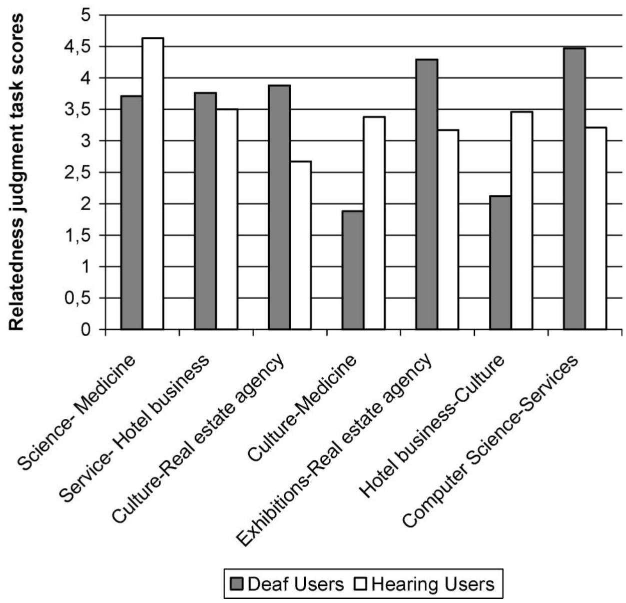


Figure 4. Pairs of experimental hypertext concepts where the prior knowledge of deaf and hearing users was significantly different.

$F(39.1521) = 1.49$; $Ms_e = 2.4$; $p < .0276$. For example, DS found less relation than H between the pairs of concepts: Culture – Medicine ($F(1.39) = 6.2$; $Ms_e = 3.6$; $p < 0.02$) and Hotel business – Culture ($F(1.39) = 4.9$; $Ms_e = 3.6$; $p < 0.03$). However, DS found a higher relation than H between the pairs: Culture – Real estate agency ($F(1.39) = 4.1$; $Ms_e = 3.6$; $p < 0.05$), Exhibitions – Real estate agency ($F(1.39) = 4.2$, $Ms_e = 2.99$; $p < 0.04$) and Computer Science – Services ($F(1.39) = 4.4$; $Ms_e = 3.6$; $p < 0.04$). With regard to related pairs, the two-way interaction between the type of user and the type of related pair was also significant, $F(14.546) = 1.72$; $Ms_e = 2.9$; $p < .0483$. Ds found less relation than H between the related pairs: Science – Nutrition, Science – Medicine, Exhibitions – Culture and Service – Hotel business, though only this last difference was significant, $F(1.39) = 10.8$; $Ms_e = 3.9$; $p < 0.002$. In related pairs, Ds only found more relation than H in the case of pair Nutrition – Computer Science; however, the difference was not significant.

Finally, the search in the condition verbal interface/long path was similar to the search in the condition of deep hypertext structure in our previous experiment (Fajardo *et al.* submitted). For this reason, we expected that the results of both experiments in those conditions to be similar to DS and H users. In order to test this hypothesis we performed several ANOVAs by subjects with experiment (Fajardo *et al.* vs. this experiment) as an independent variable for each type of user and each dependent variable. As expected, the analysis showed that the averages of correct answers were almost identical between experiments for both DS, $F(1.16) = 0.49$; $MS_e = 469.7$; $p = 0.5$ (DS = 68.7% in Fajardo *et al.* vs. 61.3% in this experiment) and H users, $F(1.18) = 0.19$; $MS_e = 315.8$; $p = 0.7$ (83.1% vs. 79.5%). The average of total response time to find a target was also similar between experiments for both DS, $F(1.15) = 0.45$; $MS_e = 151.4$; $p = 0.5$ (33.9 vs. 29.8) and H users, $F(1.18) = 1.34$; $MS_e = 50.8$; $p = 0.3$ (25.9 vs. 22.2). Finally, the difference between experiments with regard to disorientation was not significant either for neither DS, $F(1.16) = .28$; $MS_e = 0.05$; $p = 0.6$ (0.22 vs. 0.16) nor H users, $F(1.18) = 0.00$; $MS_e = 0.01$; $p = 0.9$ (0.14, vs. 0.14).

5. Discussion and conclusion

The main hypothesis of this research states that, as deaf signer users are inefficient in HIR tasks, mainly due to their low verbal and reading competencies, the substitution of textual links for graphical links may reduce such inefficiency. This hypothesis derives from the following arguments: 1) the semantic similarity judgement is an important process during information search in hypertext, according to HIR models, and 2) pictures have a privileged access to semantic memory, according to a central assumption in cognitive psychology. Additionally, several findings in the field of HCI seem to support the advantage of graphical interfaces with regard to verbal interfaces.

5.1 *Verbal interface superiority effect of path length*

In contrast to our main prediction (H1), this experiment shows that there are advantages of the verbal hypertext interface over the graphical advantages, which mainly appear when semantic processing becomes more difficult in long paths (as predicted in H2, but contrary to H3) for both deaf signers and hearing users. The finding of verbal superiority agrees with the data of other authors with regard to direct manipulation interfaces or information retrieval (Guastello *et al.* 1989, Benbasat and Todd 1993, Wiedenbeck 1999), who find that users are better with verbal or mixed interfaces (icons plus verbal labels) than with the graphical versions. However, other authors had not taken into account the potential interaction between the interface formats with the path length. In the experiment, the path length was shown to be relevant to establish the difference between interfaces: the longer the path, the harder the semantic processing and the larger the interference with the search in the graphical interface.

5.2 *Prior knowledge influence*

Knowledge on pictures and words may be a variable strongly related to the results. If users have less prior knowledge on pictures than words used in interfaces, the semantic processing may be interfered in the former case.

To test this hypothesis we re-analysed the data, taking into account the prior knowledge on interface elements (icons or words depending on the condition), which were elicited from the users in the relatedness judgement task. All users in general had better prior knowledge about the verbal labels than the icons. For each dependent variable, we performed an analysis of covariance (ANCOVA) with the type of users, interface format and path length as independent variables and prior knowledge as a covariate variable. The previous interface format effects disappeared for all dependent variables with the introduction of the covariate variable in the analysis, which suggests that prior knowledge on the interface elements and their relations is an important contributor to the difference among interface formats. This data supports the hypothesis that semantic memory is involved during HIR as the models of HIR suggest (e.g. Pirolli and Fu 2003, Kitajima *et al.* 2000) and that this process is more relevant when the information is located in deep layers of hypertext nodes since the top choices are more ambiguous and unrelated with the target. Consequently, the models of visual search (e.g. Scott 1993, Niemelä and Saarinen 2000, Liu *et al.* 2002, Fleetwood and Byrne 2003, Pearson and Schaik 2003), in spite of being helpful to predict the search process in each node of the hypertext, do not seem to be enough to explain and predict the users' interaction with hypertext interfaces and other complex systems where semantic processing is involved.

5.3 *Differences between deaf and hearing users*

Additionally, the data on prior knowledge enables us to explain the general superiority of H users over DS users found in this experiment. In general, DS users found more relationship between unrelated pairs and less relationship between related pairs than H users, which may have made the semantic similarity judgements during the HIR difficult for DS users. The knowledge difference between users could be explained by the findings of Marschark and Everhart (1999), and McEvoy *et al.* (1999) who found that deaf and hearing people utilised different strategies to access their semantic knowledge about words. For instance, in the study performed by Marschark and Everhart (1999), deaf signers and hearing children used different strategies on the 20 questions game. In this game, players must find an objective in a set of 42 pictures by asking a maximum of 20 questions. Marschark and Everhart found that hearing children asked more 'constraint questions' which, apart from eliminating more choices each turn, denoted a categorical knowledge organisation of users. However, deaf children asked more specific questions such as, 'Is it the cow?' According to the authors, these results may be due to differences in either the organisation of knowledge or the strategies of information retrieval from LTM.

With the aim of directly measuring the amount and organisation of verbal concepts (Lexicon) in deaf signer people, McEvoy *et al.* (1999) used the classical task of *controlled association norms* (Deese 1965). In this task users had to say the first word that came into their mind after the individual presentation of a word (80 words in total). Although the results of deaf signer and hearing participants were quantitatively similar, the qualitative analysis revealed that the within group coherence in the answers was higher for hearing than for deaf participants. In addition, deaf individuals left more items in blank than hearing individuals, which was interpreted as a small number of concepts available in the former.

Other authors have found that deaf people have a problem with tasks involving multidimensional decisions where a relational processing of information is necessary. Deaf people would tend to store concepts and details more than the relationship between them. The problem with the relational processing not only occurs in verbal tasks such as reading (e.g. Banks *et al.* 1990, Marschark *et al.* 1993), but also in visuospatial tasks (Otten 1980). Once again, this data may be interpreted as evidences of differences in the access strategies or organisation of knowledge in LTM between deaf and hearing individuals.

In that way, although deaf signer people would have more visuospatial capacity to maintain and manipulate visual (e.g. icons) and spatial information derived from the sign language usage, when there is also semantic information involved as in the HIR task, their different organisation of knowledge in LTM would prevent the facilitation.

5.4 *Methodological shortcomings*

In spite of several methodological shortcomings of this experiment, the data is opposite to the PSE and highlights a re-examination of the theories that postulate picture advantages in access of semantic memory. Amrhein *et al.* (2002), after finding null results in a series of classical semantic memory experiments, suggest that alternative models as proposed by Snodgrass and McCullough (1986) should be considered to explain the PSE. The hypothesis of Snodgrass and McCullough postulates that the individuals would apply two strategies in parallel in the process of categorisation with pictures, a visual and a semantic strategy, while, in the case of words, it is only possible to apply the semantic strategy, as words are not visually distinctive. The visual strategy would be faster than the semantic strategy, which would explain the PSE. However, when the visual distinctiveness among pictures is low, users must use the semantic strategy in the categorisation of pictures. In this case, words are better than pictures. One of the explanations argued by Snodgrass and McCullough is that the visual similarity among categories hinders the usual process of categorisation, which would not happen with words. However, the PSE has been contrasted with a normalised set of pictures and words in cognitive psychology research while, in our experiment, we only control the semantic distance of icons. Therefore, further research is necessary to test if the absence of PSE in our experiment is due to a non-comparable set of pictures and words in a series of measures such as the visual distinctiveness or the typicality and frequency of the concepts represented, which could vary between or within the set of icons and words utilised.

5.5 *Provisional guidelines*

Finally, we are able to hint at diverse guidelines for the design of hypertext based on these results, which obviously must be more extensively researched. For instance, the substitution of verbal links by icons may interfere instead of favouring HIR tasks when users do not have enough prior knowledge on the meaning and functions of those icons in the hypertext system. In such cases, it would be recommended to explicitly teach users the meaning and function of icons in the specific hypertext. In addition, it is important to take into account that the path length augments the number of semantic judgements and the ambiguity of top-level choices, which may mainly affect icons. For this reason, a reasonable solution may be to locate the most important icons in the shallower layers of hypertext nodes or design wide structures instead of deep. While wide verbal hypertext structures have proved to be negative for deaf signer people (Fajardo *et al.* submitted), wide graphical hypertext may improve the performance as icons seem to facilitate direct manipulation (Benbasat and Todd 1993).

The obtaining of empirically validated cognitive accessibility guidelines of web design for deaf signer people has an enormous applied repercussion, not only for such a community of users. For instance, cognitive problems of deaf people related to memory and language (e.g. short-term verbal memory or reading comprehension) would also affect other types of users such as elderly people, people with dyslexia, people navigating in a website written in a foreign language, or people with a low cultural level, which could also benefit from such guidelines.

In addition, the exposed applied problems offer us the coverage to generate and test hypotheses on the cognitive functioning of deaf people in general (for instance, how knowledge about the world stored in LTM is structured by the experience, McEvoy *et al.* 1999) and in the specific context of interaction with hypertext. The interaction with hypertext is a new and complex way of interaction, which is changing the habitual ways of knowledge acquisition and this may be influencing the manner in which our cognitive system works (Dix *et al.* 2003). For this reason, it is important that cognitive science is able to provide explanatory answers and generate predictive models on the interaction with hypertext. In addition, as Wright (1993) suggests, hypertext may be an important probe field of psychological theories, which try to explain how to integrate a range of cognitive processes (e.g. attention, comprehension, memory).

References

Amrhein, P.C., McDaniel, M.A. and Waddill, P.J., 2002, Revisiting the picture superiority effect in symbolic comparisons: do pictures provide privileged access? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **28,** pp. 844–857.

Arend, U., Muthig, K. and Wandmacher, J., 1987, Evidence for global feature superiority in menu selection by icons. *Behaviour and Information Technology*, **6,** pp. 411–426.

Banks, J., Gray, C. and Fyfe, R., 1990, The written recall of printed stories by severely deaf children. *British Journal of Educational Psychology*, **60,** pp. 192–206.

Benbasat, I. and Todd, P.A., 1993, An experimental investigation of interface design alternatives: icon vs. text and direct manipulation vs. menus. *International Journal of Man-Machine Studies*, **38,** pp. 369–402.

Berners-Lee, T., World Wide Web Consortium. Available online at www.adelaide.edu.au/webguide/topics/accessibility/(accessed 30 December 2005).

Blankenberger, S. and Hahn, K., 1991, Effects of icon-design on human–computer interaction. *International Journal of Man-Machine Studies*, **35,** pp. 363–377.

Bush, V., 1945, As we may think. *The Atlantic Monthly*, **176,** pp. 101–108.

Byrne, M.D., 1993, Using icons to find documents. In *Human Factors in Computing Systems: Proceedings of INTERCHI'93*, M.D. Byrne (Ed), pp. 446–453 (Reading, MA: Addison-Wesley).

Collins, A. and Loftus, E., 1975, A spreading-activation theory of semantic processing. *Psychological Review*, **82,** pp. 407–428.

Deese, J., 1965, *The Structure of Associations in Language and Thought* (Baltimore, MD: The Johns Hopkins Press).

Dillon, A. and Song, M., 1997, An empirical comparison of the usability for novice and expert searchers of a textual and a graphic interface to an art–resource database. *Journal of Digital Information*, **1,** pp. 12–20.

Dix, A., Howes, A. and Payne, S.J., 2003, Post-web cognition: evolving knowledge strategies for global information environments. *International Journal of Web Engineering Technology*, **1,** pp. 112–126.

Emmorey, K., 2002, *Language, Cognition and the Brain. Insights from sign language research* (Mahwah, New Jersey: LEA).

Fajardo, I., Cañas, J.J., Salmerón, L. and Abascal, J., 2004, Web Accessibility Problems of Deaf Signer Users: the influence of Language Modality, Web Structure and Reading Abilities (submitted).

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.

Fleetwood, M.D. and Byrne, M.D., 2003, Modeling the visual search of displays: A revised ACT-R/PM model of icon search based on eye-tracking and experimental data. In *Proceedings of the Fifth International Conference on Cognitive Modeling*, F. Detje, D. Dörner and H. Schaub (Eds), pp. 87–92 (Bamberg, Germany: Universitas-Verlag Bamberg).

Guastello, S., Traut, M. and Korienek, G., 1989, Verbal versus pictorial representations of objects in a human–computer interface. *International Journal of Man-Machine Studies*, **31,** pp. 99–120.

Gyselinck, V., Cornoldi, C., Ehrlich, M.-F., Dubois, V. and de Beni, R., 2002, Visuospatial memory and phonological loop in processing texts and illustrations. *Applied Journal of Cognitive Psychology*, **16,** pp. 665–685.

Hutchins, E.L., Hollan, J.D. and Norman, D.A., 1986, Direct manipulation interfaces. In *User Centered System Design*, D.A. Norman, and S.W. Draper (Eds), pp. 87–124 (Hillsdale, NJ: Lawrence Erlbaum Associates).

Ivory, M. and Hearst, M., 2001, State of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, **33,** pp. 1–47.

Kinjo, H. and Snodgrass, J.G., 2000, Does the generation effect occur for pictures? *American Journal of Psychology*, **113,** pp. 95–117.

Kitajima, M., Blackmon, M.H. and 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), pp. 357–373 (UK: Springer).

Kruley, P., Sciama, S.C. and Glenberg, A.M., 1994, On-line processing of textual illustrations in the visuospatial sketchpad: evidence from dual-task studies. *Memory and Cognition*, **22,** pp. 261–272.

Landsdale, M.V., 1988, On the memorability of icons in an information retrieval task. *Behaviour and Information Technology*, **7,** pp. 131–151.

Levie, W. and Lentz, R., 1982, Effects of text illustrations: a review. *Research Educational Communications and Technology Journal*, **30,** pp. 195–232.

Liu, B., Francis, G. and Salvendy, G., 2002, Applying models of visual search to menu design. *International Journal of Human Computer Studies*, **56,** pp. 307–330.

Lotto, L., Job, R. and Rumiati, R., 1999, Visual effects in picture and word categorization. *Memory and Cognition*, **27,** pp. 674–684.

Marschark, M., 1998, Memory for Language in Deaf Adults ad Children. *Scandinavian Audiologist*, **27,** pp. 87–92.

Marschark, M., 2003, Cognitive functioning in deaf adults and children. In *Oxford handbook of deaf studies, language, and education*, M. Marschark and P.E. Spencer (Eds), pp. 464–477 (New York: Oxford University Press).

Marschark, M. and Everhart, V., 1999, Problem-solving by deaf and hearing students: twenty questions. *Deafness and Education International*, **1,** pp. 65–82.

Marschark, M., de Beni, R., Polazzo, M. G. and Cornoldi, C., 1993, Deaf and hearing-impaired adolescents' memory for concrete and abstract prose: effects of relational and distinctive information. *American Annals of the Deaf*, **138,** pp. 31–39.

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,** pp. 487–519.

McDougall, S., de Bruijn, O. and Curry, M., 2000. Exploring the effects of icon characteristics: concretness, complexity and distinctiveness. *The Journal of Experimental Psychology: Applied*, **6,** pp. 291–306.

McEvoy, C., Marschark, M. and Nelson, D.L., 1999, Comparing the mental lexicons of deaf and hearing individuals. *Journal of Educational Psychology*, **91,** pp. 1–9.

Navarro-Prieto, R. and Cañas, J.J., 2001, Are visual programming languages better? The role of imagery in program comprehension. *Int. J. Hum.-Comput. Stud.* **54,** pp. 799–829.

Nelson, D.L., Reed, U.S. and Walling, J.R., 1976, Picture superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, **2,** pp. 523–528.

Niemelä, M. and Saarinen, J., 2000, Visual search for grouped and ungrouped icons. *Human Factors*, **42,** pp. 630–635.

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

Ottem, E., 1980, An analysis of cognitive studies with deaf subjects. *American Annals of the Deaf*, **125,** pp. 564–575.

Paivio, A., 1971, *Imagery and verbal processes* (New York: Holt, Rinehart and Winston).

Paivio, A., 1977, Images, propositions and knowledge. In *Images, perception and knowledge* [Western Ontario Studies in the Philosophy of Science, 8], J.M. Nicholas (Ed.), pp. 47–71 (Dordrecht, Netherlands: Reidel).

Paivio, A., 1991, Dual-coding theory: retrospect and current status. *Canadian Journal of Psychology*, **45,** pp. 255–287.

Paivio, A. and Csapo, K., 1973, Picture superiority in free recall: imagery or dual coding? *Cognitive Psychology*, **5,** pp. 176–206.

Pearson, R. and 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,** pp. 327–353.

Pellegrino, J.W., Rosinski, R.R., Chiesi, H.L. and Siegel, A., 1977, Picture-word differences in decision latency: an analysis of single and dual memory models. *Memory and Cognition*, **5,** pp. 383–396.

Pirolli, P. and Fu, W-T.F., 2003, SNIF-ACT: a model of information foraging on the World Wide Web. In *Proceedings of the Ninth International Conference on User Modeling*, P. Brusilovsky, A. Corbett and F. de Rosis (Eds), pp. 45–54 (Johnstown, PA: Springer).

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,** pp. 471–475.

Scott, D., 1993, Visual search in modern human–computer interfaces. *Behaviour and Information Technology*, **12,** pp. 174–189.

Seeman, L., 2002, Inclusion of cognitive disabilities in the web accessibility movement. In *Proceedings of the Eleven International World Wide Web Conference*, Honolulu, Hawaii, USA, 7–11 May.

Smith, P., 1996, Towards a practical measure of hypertext usability. *Interacting with Computer*, **8,** pp. 365–381.

Snodgrass, J.G. and McCullough, B., 1986, The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **12,** pp. 147–154.

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), pp. 225–229 (New York: Springer-Verlag).

Weidenbeck, 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,** pp. 68–82.

Wright, P., 1993, To jump or not to jump: strategy selection while reading electronic texts. In *Hypertext: A Psychological Perspective*, C. McKnight, A. Dillon and J. Richardson (Eds), pp. 137–152 (West Sussex, England: Ellis Horwood Limited).

Appendix A

Sets of icons with lower scores in semantic distance, which were selected to serve as links in the graphical interface condition.

Appendix B

Average scores in the judgement relation task for each pair of hypertext relevant concepts. The bold words sign the pairs of concepts where the scores of deaf and hearing users were significantly different.

Pairs of Concepts		Type of User	
Related Pairs		Deaf	Hearing
Computer Science	Medicine	2.47	2.63
Real State Agency	Hotel Business	2.76	3.00
Music	Exhibitions	3.53	3.25
Nutrición	Computer Science	2.94	2.08
Medicine	Nutrición	3.41	4.33
Music	Paintings	3.12	4.21
Science	Nutrición	5.71	5.79
Science	**Medicine**	**3.71**	**4.63**
Exhibitions	Paintings	4.35	4.42
Exhibitions	Culture	3.76	4.08
Computer Science	Science	3.82	4.63
Music	Culture	2.24	4.29
Paintings	Culture	4.12	4.13
Services	**Hotel Business**	**3.76**	**3.50**
Services	Real State Agency	3.82	3.79
Un Related Pairs			
Science	Exhibitions	3.65	3.92
Science	Music	3.00	3.04
Science	Services	2.24	1.88
Culture	**Real State Agency**	**3.88**	**2.67**
Culture	Computer Science	4.24	3.71
Culture	Science	3.59	4.25
Culture	**Medicine**	**1.88**	**3.38**
Exhibitions	Services	2.24	2.33
Exhibitions	Hotel Business	1.82	2.33
Exhibitions	**Real State Agency**	**4.29**	**3.17**
Hotel Business	Science	2.00	2.08
Hotel Business	Medicine	2.35	2.50
Hotel Business	**Culture**	**2.12**	**3.46**
Hotel Business	Nutrición	3.94	4.54
Hotel Business	Computer Science	2.12	2.00
Computer Science	Music	3.24	3.83
Computer Science	Exhibitions	2.71	3.21
Computer Science	**Services**	**4.47**	**3.21**
Real State Agency	Computer Science	4.35	3.63
Real State Agency	Paintings	2.59	2.96
Real State Agency	Science	2.18	1.67
Medicine	Exhibitions	1.65	1.96
Medicine	Music	1.59	2.00
Medicine	Real State Agency	2.88	2.54
Medicine	Paintings	2.24	2.00
Music	Real State Agency	2.82	2.75
Music	Services	2.29	2.21
Music	Hotel Business	2.76	2.13
Nutrición	Exhibitions	2.18	2.54
Nutrición	Services	2.47	2.54
Nutrición	Real State Agency	2.00	2.38
Nutrición	Music	1.41	1.50
Paintings	Nutrición	1.71	1.67
Paintings	Hotel Business	1.71	1.67
Paintings	Computer Science	3.06	2.75
Paintings	Science	2.82	3.13
Services	Medicine	4.06	3.83
Services	Paintings	2.12	1.96

(*continued*)

Appendix B (*Continued*).

Pairs of Concepts		Type of User	
Related Pairs		Deaf	Hearing
Services	Culture	2.82	2.96
Culture	Nutrición	2.94	4.00