

Enhancing accessibility for visually impaired users: the Munch exhibition

Nicoletta Di Blas and Paolo Paolini, HOC-DEI, Politecnico di Milano, Italy; Marco Speroni, University of Lugano, Switzerland and Angelo Capodieci, MEDÌ, Lecce, Italy

Abstract

The Web enhances the role of visual communication: users look at pages to get content and to select links and operations. The visual nature of the Web involves at least three aspects: content (images, graphics and text are all meant to be looked at), page organization (e.g. locating different pieces of content, with their relationships) and navigation (e.g. locating links, and guessing their meaning from the label or icon, the hint, the position, etc). The most widely used technique for allowing visually impaired users to access the Web is based on screenreaders, i.e. SW tools capable of reading HTML pages. The W3C (<http://www.w3.org>) has published important recommendations, under revision right now, to help designers to develop readable Web pages. This paper will argue that W3C recommendations and also screenreader's technology are not sufficient to ensure an efficient - even less satisfactory - Web experience.

It is, in fact, possible to develop Web pages fully compliant with W3C standard, yet hardly manageable through screenreaders. The HOC laboratory of Politecnico di Milano and the TecLab of University of Italian Switzerland are developing a new set of guidelines, expanding those of W3C, based on the assumption that a Web experience can be compared to a dialogue between a user and a machine. The first exploitation of this research is a Web site developed in cooperation with the Staatliche Museen of Berlin in Spring 2003. Unsolicited user reactions show that the research's effort is moving in the right direction, but also that we are still far from a fully satisfactory solution.

Keywords: accessibility, usability, screenreader, W3C, dialogue.

1. The Context Of Interaction

In principle, Web users should at a quick glance be able to understand the page structure, spot the section and the specific pieces of information they are looking for, identify relevant links, click on them. It is common experience, however, that often this is not case. But it *should* be true for well designed Web sites where content, navigation and layout are carefully planned..

Content: Web pages are crowded with images, graphics, text and multimedia all meant to be *looked at*, (except audio comments.) Web content displays at least three characteristics:

Richness: while a text is always linear; that is, one element follows another, a page can host very many different pieces of content at the same time, sometimes even too many (see *Figure 1*).



Figure 1: Amazon Home Page (<http://www.amazon.com>)

Redundancy: the same information is often repeated on the same page or on different pages of the same site. In <http://www.nga.gov>, the site of the National Gallery of Washington (see Figure 2) the list of the landmarks (important sections of the site), can be found on almost every page (on the left frame) and, at the same time, at the bottom of each page.



Figure 2: NGA Web site (<http://www.nga.gov>) Redundancy of links

Contextualization: The precise meaning of a text often depends on where it is placed: very short labels; for example *description*, usually mean something like *description of the work of art, the image, close to this link*.

Page layout : page organization conveys a meaning of its own, separating different areas of content, highlighting the main content of the page, driving the user's attention to some relevant piece of information. The layout of the above mentioned page visually shows the role of the different parts: the main content, the different types of links and their different relevance or meaning. Visual features such as the font, colour, size, type, the items' positions, help the user not to lose orientation: we call all these characteristics *graphic's*

semantics, in that they are all meant primarily to (help) convey meaning rather than to play a purely aesthetic role. Graphic semantics play a crucial role in all the above aspects. For example, in the National Gallery's list of landmarks, the last two links are both orange: this means they are *of the same kind*; in fact, they both concern special events.

Navigation: the possibility of moving from one page to another is based on links. As shown in another page of the same site - see Figure 3 - the user must use *visualization* for locating links, guessing their meaning, understanding the overall positioning (e.g. in a guided tour), and moving the pointer of the mouse over one of them.



Figure 3: NGA Web Site (<http://www.nga.gov>)
Visualisation for locating links and guessing their meaning

2. The Problem

There are at least three basic situations where the interaction between a human being and the application *should not* rely on visual perception:

- when the user can't see well or can't see at all (an obvious and socially very relevant situation!)
- when the user has a device with a very small screen (a mobile device, typically).
- when the user is in a situation (e.g. driving a car, walking in a city or in a museum) where looking at a screen is not desirable or possible, since it is dangerous or it is preferable to look elsewhere (e.g. to the countryside, the buildings, the monuments, the paintings, etc.).

Since situations two and three often go together, communication from the application to the user, and possibly also vice versa should be mainly oral, or exclusively oral (when driving), rather than visual.

The focus of this paper is on situation one, lack of clear sight - from the technical point of

view the worst situation, since vision is completely ruled out. Good solutions for case one, therefore, should be easily adaptable to (and adoptable for) situations 2 and 3.

Currently, visually impaired users access the Web thanks to software tools called *screen-readers*. Since one of the most commonly used is JAWS (available at <http://www.freedomscientific.com>), we will often refer to its specific features, with immediate applicability to other similar tools.

Screen-readers read aloud, with a synthesized voice, the screen contents (actually reading the associated HTML file), describing icons, texts, new windows opening - making interaction with the application possible.

Screenreaders provide a tentative solution for visually impaired people, with a number of unsolved problems:

- They read everything, including elements of HTML that are useful for visualization only (and do not convey real meaning to the listener).
- They have (by default, at least) a simplistic listening strategy, top-to bottom/left-to-right, making it difficult and boring to wait for the relevant piece of information. (Try reading aloud a page of a daily newspaper adopting the same strategy, and measure how long it takes to get something relevant.)
- They fail to convey the overall organization of the page, with the relative priorities of the different parts.
- They interweave reading content with reading links, with total confusion for the listener. The good exception is that the listener can get the links list (in alphabetical order), without the content; but can't get the content without the links. In addition, even the list in alphabetical order is not effective (see *Figure 4*). What if many links begin with the same word? Or if they're in an interrogative form, for example, all beginning with "where can I find"? Again, this means time and patience in waiting for the link meaning to clarify, or wrong and time-consuming moves in the site (Theofanos & Redish, 2003).

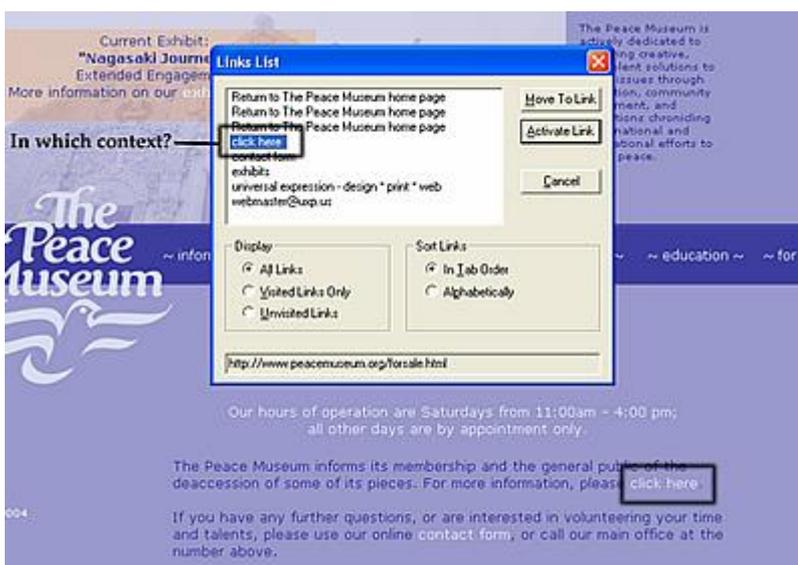


Figure 4: Home Page of The Peace Museum Web site (<http://www.peacemuseum.org>)

List of links using JAWS

The selection mechanisms of the links are difficult and cumbersome. While in theory it is possible to confirm the selection while listening to a link, in practice, due to synchronization problems (of the audio with the current position on the page), it never works.

Page layout and graphic semantics are both completely lost: the metallic voice of the screenreader reads one by one all the pieces of information on the page with the same emphasis and tone (the landmarks, the main content, the service links), as if they all share the same degree of importance.

The richness and redundancy of the page, therefore, become cumbersome obstacles to a good outcome. The final effect is that navigation across Web page becomes an annoying and tiring task, rather than an engaging and pleasant activity. In many cases the user is forced to surrender after a short attempt.

It is obvious, from the previous discussion, that a generic HTML page can be scarcely readable or not readable at all. The solution is to provide designers with tips and hints on rules for making Web pages readable. General awareness of the problem is increasing: in Italy, for example, a piece of legislation is being approved to force public administration to make Web sites accessible to, among others, visually impaired users.

3. The W3C standard

The World Wide Web Consortium (W3C) published Web Content Accessibility Guidelines 1.0 in May 1999; version 2.0 is currently under development (<http://www.w3.org/TR/2003/WD-WCAG20-20030624/>). The recommendations of WCAG are part of the WAI - Web Accessibility Initiative.

The basic recipe of the WCAG consists of 4 major guidelines for accessibility: the site must be

1. Perceivable: make content perceivable by any user'
2. Operable: ensure that interface elements in the content are operable by any user'
3. Understandable: make content and controls understandable to as many users as possible'
4. Robust: use Web technologies that maximize the ability of the content to work with current and future accessibility technologies and user agents'

Each guideline is further qualified by a number of checkpoints (18 in the current draft version of WCAG 2.0), defined as *core* and *extended*. (In version 1.0, checkpoints were assigned priorities from 1 to 3 instead.) Consider the following.

Guideline 1: perceivable. Make content perceivable by any user.

Core checkpoint 1.1: (ensure that) *all non-text content that can be expressed in words has a text equivalent of the function that the non-text content was intended to convey.*

This is undoubtedly a very important checkpoint, the basic step for blind users, but not always easy to interpret. A text equivalent is specified as communicating the same information as the non-text content was intended to convey'. Let's see two examples:



Figure 5: MOMA Web site (<http://www.moma.org>)
Text equivalent of a non-text content

In the Web site of the Museum of Modern Art of New York (see *Figure 5*), the image of the chair displays a text equivalent, stating that we are facing a design object by Josef Hoffmann, namely a 'Sitzmaschine chair with adjustable back', designed more or less in 1905: but does it really make the image content more accessible? Moreover, apart from the string 'design object', it is a repetition of the text available to all just below: therefore, a blind user will have to listen to it twice!

It is definitely not an easy task to decide what exactly a text equivalent is; let's look again at the collection page on the National Gallery of Washington Web site. We are told that the Web tour of the week is about Gilbert Stuart (American, 1755-1828). This is all the information a blind user can get, while sighted visitors can immediately see (using the thumbnails available as an appealing preview) that Mr. Stuart was a *painter* and that he painted *portraits* (or at least the Web tour is mainly about portraits). They can also decide whether the tour is interesting or not (simply by looking at the examples given), or, if they're more knowledgeable, to which style Stuart can be attributed, etc. How much of this information should or could be conveyed through a text equivalent?

Guideline 2: operable. Ensure that interface elements in the content are operable by any user.

Extended Checkpoint 2.4: (ensure that) *structure and/or alternate navigation mechanism have been added to facilitate orientation and movement in content.*

The required success criteria for this checkpoint are the following:

- Hierarchical structure mark up
- Table of contents (or site map)
- Alternate display orders (or alternate site navigation mechanism)

This checkpoint is simplistic and therefore quite unusable. It is certainly true that an effort must be made to improve orientation, but a hierarchical structure or a site map are poor solutions that do not work for complex sites (see www.nga.gov, for example) in which

pieces of content are networked. The true solution is a rational design technique that allows effective organization of content, of user paths to the pieces of information, and of the associations among them. To split long texts into paragraphs, to provide a site map, or to ensure alternate display orders might improve details once a good site is designed, but they are completely useless if the site is badly designed.

Guideline 3: understandable. Make content and controls understandable to as many users as possible.

Extended Checkpoint 3.4: (Ensure that) *layout and behaviour of content is consistent or predictable, but not identical.*

Again this is a true, but simplistic, checkpoint. Consistency is very often, if not always, a desirable characteristic of a site, but it depends on the overall design methodology and not on the assessment of details such as the location of navigation elements or bars.

Moreover, for visually impaired users, the visual layout has nothing to do with the audio layout; therefore, the suggestion of putting navigational elements always in consistent locations is pointless. It would certainly be more important to tell the designer how to shape content and navigation patterns in a consistent manner - as, for example, the UWA methodology does (UWA Consortium, 2001a).

Guideline 4: robust. Use Web technologies that maximize the ability of the content to work with current and future accessibility technologies and user agents.

Extended Checkpoint 4.2: (Ensure that) *technologies that are relied upon by the content are declared and widely available.*

This recommendation looks more like a wish than a real guideline; Web designers do not control Web technology and its evolution (while W3C has a strong influence on both).

Therefore, although W3C guidelines represent an important contribution, they are not sufficient to ensure visually impaired people of efficient and satisfactory access to the Web. In our opinion, its basic recommendations (especially those about content) are important, correct and usable, whilst general recommendations are simplistic and of little use. Moreover, there is a strong concern about details, but little concern for, or demonstrated understanding of, conceptual issues: overall design of the Web site, organization of each page, reading strategy, etc. It is sufficient to take any Web page with a long list of items to realize that it would be impossible to make it readable. Whatever its relationship with the W3C standard, try to navigate a page using basically the button **back**; you soon realize how difficult it is for someone who can only listen to the site! The estimated use of **back** in a typical Web site, with respect to the total of navigation actions, is 40%!

4. Beyond W3C: The WED Project And A Practical Experience

The HOC laboratory of Politecnico di Milano, together with the Tec-Lab of the University

of Lugano (Switzerland), is currently developing a set of guidelines attempting to go beyond the state of the art represented by the WC standard.

The ambitious goal is to design Web sites *optimised* for visually impaired users, using, as a starting point, a robust design methodology - W2000 (UWA Consortium, 2001), the evolution of HDM (Garzotto & Paolini, & Schwabe, 1995) & (Paolini & Garzotto, 1999), in our case. This methodology already ensures a number of desirable features; such as understandable overall organization, simplicity (compatibly with the intrinsic complexity of the application), orientation, consistency, etc.

The research effort, called **WED** (WEB as Dialogue), stems from the assumption that a Web experience can be compared to a sort of dialogue between a user and a machine. The conversational turns of the machine are represented by displaying pages: they represent, at the same time, content replying to previous user requests (see below) and offer new conversational 'painting'. If the user, for example, asked for *Adorazione dei Magi* in <http://www.nga.gov> (see Figure 6), the reply would be a page showing the desired content, and at the same time the offer of new possibilities of conversations (e.g., among others, you may see the full size picture, the details, the bibliography, ...). The conversational turns of the user are represented by the explicit selection of one of the offers (e.g. I would like to see the details').



Figure 6: NGA Web site (<http://www.nga.gov>) Web page describing the painting Adorazione dei Magi

The highlighting of analogies and differences helps to improve the design technique, shifting the Web experience from the purely visual to the audio channel (see Di Blas & Paolini, 2003a, and Di Blas *et al.*, 2003).

The WED research group is composed of Web designers, linguists, usability experts and communication scientists. The method of research consists of parallel work by different sub-groups, sharing and putting together their results. Communication experts and usability experts record human-machine dialogues (i.e. sessions of use of a Web site), using video cameras and the thinking aloud method. The dialogue is then transcribed and interpreted by the linguists using existing dialogue models. (There are too many references to list them

all: try <http://www.usilu.net:90/~wed.>) Analogies and differences are put in evidence and discussed with the Web designers, who then try to adapt the existing design technique in order to make the interaction with the Web site more natural, either by adding new features or by re-interpreting existing features in dialogic terms. For example, the navigation among nodes concerning the same (in W2000 terms) *entity* (a painting of NGA, for example) can be seen as an investigation of the different aspects of a topic of discussion (Di Blas & Paolini, 2003a).

WED in Practice

The first exploitation of WED has been a Web site for an exhibition of prints of Edward Munch. It took place at the Staatliche Museen of Berlin in Spring, 2003.



Figure 7: Web site of the Edvard Munch Exhibition in Berlin (<http://www.Munchundberling.org>) Home Page

Many special features were implemented in the Munch site to optimize it for visually impaired users; lack of space prevents us from discussing them all in detail. Therefore, only a few are considered here. But visit the site and try to 'listen to it', using the screenreader according to the instructions given in the site itself (<http://www.Munchundberlin.org>). We must warn the you not to compare the experience to a traditional 'visual' visit to the site, but instead to the use of a screenreader in an average Web site.

Content is divided into usable chunks

The W2000 design methodology has been used as a basis for efficiently organizing all the different pieces of information into *nodes*, planned in a systematic, consistent and (hopefully) usable manner.

Page schemas

A sighted user can (more or less easily) detect whether the desired page has been reached (otherwise 'back' must be used again) and can quickly (depending on the designer's ability) locate, within the page, the navigation link that will start a new exploration.

A blind user, however, will have to listen (at least for a while) to the screenreader reading the page contents, in order to understand first, whether this is the right page, and second, precisely where the desired link is, on that page.

The command **My history in the site**, implemented in the Munch site, tries to overcome the problem: it offers a list of the semantic steps the user has made thus far in the site, thus facilitating a quick re-selection of previously visited content. It corresponds, in a human dialogue, to a conversational turn like this: (user) 'you said you knew something interesting about Botticelli, can you tell me, please?' instead of: (user) 'could you please repeat your last 4/5 conversational turns? I thought you mentioned Botticelli some minutes ago'.

The WED research, though still ongoing, has already allowed a significant step forward towards the optimization of Web sites. Regarding the Munch Web site, we have received very positive feedback from visually impaired German users, such as the following:

The first impression of the site is very positive. The pages are clearly structured. **All the links have detailed titles, which allow an informative and nice internet session.** With my favourite screenreader JAWS (version 4.51.212) **I needed about 1.5 minutes to get a general overview for all further action.** This seems to me an acceptable time, considering that this form of documentation of such an exhibition is quite unusual at the moment.

Heading with ALT-TAGs - great idea! The engineers did an excellent job by not using the headings only as structural elements, but adding Alt-Tags to them. This creates a **very fast and effective structure of orientation**, which above all will be very convenient for users who surf only occasionally.

To interest sighted people in listening to Web sites?

The desire of the engineers to interest sighted people in listening to the Web site is very ambitious, but seems quite hard to achieve as the visitors of the exhibition are living in a visually oriented reality. I tried a few times to convince sighted people to have a try at a computer game without the screen and only presented with sound - but only a few people did so; they were rather an exception. But I like to be proved wrong.' (*mail from Mr. Martin Kirchner, May 18th, 2003; emphasis added by the authors*).

Conclusions

The first thing to be said is that WED is an ongoing research project: 'we feel that we are moving in the right direction, not that we have achieved a fully satisfactory result. We still do not have good solutions for problems that we have understood (e.g. long lists of items are not easily manageable orally), and we have a lot of problems not yet well understood (e.g. the link names or the flexibility of reading strategies). Our approach is to work at the same time on theoretical ground (studying what linguists can tell us) and on empirical ground (designing improved solutions, implementing them and testing them with actual

users).

A general goal that we have in mind is to eventually obtain Web sites efficiently usable by all, but at the same time optimal (in a measure yet to be understood) for visually impaired users. This strategy is based on three motivations:

- Making different Web sites for the visually impaired is not economically or organizationally sustainable.
- People with disabilities do not like to be confined in cultural ghettos; they do not want to have separate Web sites.
- We have experienced in practice that a good solution for special users is in fact a good solution for any user. An efficient way to convey long lists of items, for example, would be useful for every user, not just for the visually impaired ones.

On the technical side, we think that making Web sites optimal is not just a matter of good design: we must move from screenreaders to *page readers*, i.e. SW tools that can spell out the *meaning* of the page. Even better would be *Web readers*, i.e. SW tools that can engage the listener in a globally meaningful conversation.

On the utilitarian side, we are working at incorporating preliminary results into Cultural Heritage interactive applications accessible over several channels: the traditional Web sites, but also devices with very small screens. The challenge is to make the applications mainly oral, rather than mainly visual, as they are on the Web today. We hope that, in 12 months, we will have good real-life examples to discuss.

Acknowledgments

We wish to acknowledge the work of all the people that contribute to this exciting still on-going research. We therefore warmly thank our friends and colleagues Davide Bolchini, Sabrina Lurati, Andrea Rocci and a team of brilliant students from Politecnico di Milano (Daniele Gobetti, Marco Marini, Fulvio Prisinzano) and from USI (too many to list them all) who work passionately with us. We also thank Eddo Rigotti and Peter Schulz (faculty members at USI) for their invaluable suggestions and - last but not least - Benedetto Benedetti (from Scuola Normale di Pisa), who coordinated the HELP project for which the Munch site was developed, and Andreas Bienert (from the Staatliche Museen of Berlin), who, together with his wonderful staff, made the experimentation possible.

References

Di Blas, N. (2003). Deixis & Formulations - When the Web speaks of himself. *SCOMS - Studies in COMMunication Sciences* (i.c.s.).

Di Blas, N. & P. Paolini, (2003a). Do we speak with computers? How linguistics meddle with the Web. In E. Rigotti, A. Giacalone-Ramat & A. Rocci (Eds.). *Linguistics & new professions*. Milano: Franco Angeli, 221-233

Di Blas, N. & P. Paolini, (2003b). 'There And Back Again': What Happens to Phoric Elements in a 'Web Dialogue'. *Journal of Document Design*, 4(3), 194-206

Di Blas, N., P. Paolini, M. Speroni, A. Bienert (2003). Listen to a Web site: accessibility

(beyond current standards) and a market opportunity. ICHIM 2003 proceedings, Paris

Garzotto F., L. Mainetti, P. Paolini (1996). Navigation in Hypermedia Applications: Modeling and Semantics, in *Journal of Organizational Computing and Electronic Commerce*, 6 (3).

Garzotto F., P. Paolini, and D. Schwabe (1995). HDM - A Model-Based Approach to Hypertext Application Design. *ACM Transactions on Information Systems*, Vol. 11, No. 1, January 1995.

Nevile, L. & C. McCathieNevile (2002). The Virtual Ramp to the Equivalent Experience in the Virtual Museum. *Accessibility to Museums on the Web*. In Bearman, D. & J. Trant (Eds.). *Museums and the Web 2002. Selected Papers from an International Conference*. Pittsburgh: Archives & Museum Informatics, 93-99. Available: <http://www.archimuse.com/mw2002/papers/nevile/nevile.html>

Paolini P., F. Garzotto (1999). Design patterns for WWW hypermedia: problems and proposals. In *Proc. ACM HT '99. Workshop: Hypermedia Development: Design patterns in Hypermedia*, Darmstadt, 1999.

TEC-Lab - Technology Enhanced Communication Laboratory (2002). *Web as Dialogue: Interpreting Navigational Artifacts as Dialogic Structures*, internal report.

Theofanos, M.F. & J. Redish (2003). Guidelines for Accessible and Usable Web Sites: Observing Users Who Work With Screenreaders. *Interactions*, vol. X.6, 36-51

UWA Consortium (2001a), *D7: Hypermedia and Operation Design*. Deliverable Project, <http://www.uwaproject.org>

UWA Consortium (2001b), *D6: Requirements Elicitation: Model, Notation, and Tool Architecture*. Deliverable Project, <http://www.uwaproject.org>

On-line resources:

Web Accessibility Initiative <http://www.w3.org/WAI/>

Device Independence Activity <http://www.w3.org/2001/di/>

Multimodal Interaction Activity <http://www.w3.org/2002/mmi/>

Voice Browser Activity - Voice enabling the Web! <http://www.w3.org/Voice/>

Read more: [Museums and the Web 2004 : Papers : Di Blas et al, Enhancing Accessibility ...](http://www.museumsandtheweb.com/mw2004/papers/diblas/diblas.html#ixzz2BwxYN4gB)
<http://www.museumsandtheweb.com/mw2004/papers/diblas/diblas.html#ixzz2BwxYN4gB>
Under Creative Commons License: [Attribution Non-Commercial No Derivatives](http://creativecommons.org/licenses/by-nc-nd/2.0/)