“Usable Accessibility” to the Web for blind Users

Nicoletta Di Blas[1], *Paolo Paolini*[1], *Marco Speroni*[2]

1 HOC (Hypermedia Open Centre), Department of Electronics and Information
Politecnico di Milano, Italy; diblas@elet.polimi.it, paolo.paolini@polimi.it

2 TEC-Lab (Technology Enhanced Communication), Faculty of Communication Sciences
University of Lugano, Switzerland, speronim@lu.unisi.ch

Abstract. Websites were originally conceived for a text-based interface, but very early (from MOSAIC on) they became the epitome of graphic interface applications. A Web page is inherently complex since it simultaneously conveys pieces of content and relationships among them, “links” to other pages, etc. If the core of the communication lies in the content, a lot of additional but often fundamental semantics comes from visual features: i.e. layout, colors, fonts, spatial relationships, positioning within the page, etc. In addition, in modern Web applications, a large portion of the content is visual (e.g. pictures, graphics), or based on visual perception (e.g. tables, diagrams, etc.).
Accessing the Web, for users with disabilities, can be difficult or even very difficult in many senses. In this paper we address the issue of accessibility for a specific type of disability, i.e. blindness. The most useful technology devised for blind users is based upon “ screenreaders”, i.e. software tools that “read” pages aloud. The W3C consortium, through the WAI initiative, has taken the step of providing guidelines, in order to tell developers what they should (or should not) do in order to build “readable” pages.
The main thesis of this paper is that the W3C guidelines only guarantee “technical readability”, i.e. the very fact that screenreaders can work; they do not ensure at all the fact that the a Website is “accessible” by blind users, in the sense that blind users can effectively access it. For this reason we advocate “usable accessibility”, ensuring an effective user experience, as apposed to “technical accessibility”, that is the main concern of W3C guidelines.
In the paper we present some empirical solutions, toward usable accessibility, that we have devised for a specific site (www.munchundberlin.org) and also a more long term approach (WED – WEb as Dialog), based upon linguistic research and on the assumption that a Web experience can be treated as a kind of dialog between a user and a machine, and therefore compared (in terms of quality and effectiveness) to a dialogue between the same user and another human being (the curator of the exhibition, for example).
This research was partially funded by the Swiss National Fund (contract FNRS 105211-102061/1) and by Culture2000 (project HELP), a research program of the European Commission.

1. HOW BLIND PEOPLE ACCESS THE INTERNET

Whilst character-based interfaces offered blind people the extraordinary possibility to make use of their skills in using keyboards and interacting with software tools, graphic interfaces, implying complex pages’ layouts, many visual features and above all the use of the mouse have made their use of the many valuable resources offered by the Web a difficult and cumbersome task.
Developing separate Websites specially dedicated to this category of users is definitely not the right solution: first of all, not all the institutions would be willing to pay double costs to develop and also to keep updated two different Websites; a check of the multilingual versions of many Websites clearly demonstrates that usually the main Website is updated whilst its “foreign clones” are left behind, in terms of graphic, content, services, etc. Moreover, blind users themselves refuse being “ghettoized”, rather claiming that a better design would enhance the efficiency and satisfaction of the Web experience for *any* kind of user (Theofanos & Redish, 2003).
Visually impaired people currently access the Web by using screenreaders, that is, software tools capable of interpreting the HTML code and reading it aloud (with a synthesized voice); interaction is allowed by the use of Braille keyboards. Screenreaders’ worth is clear;

nonetheless, their limits too start being recognized and discussed in literature (see again Theofanos & Redish , 2003). We shall recollect them here in short (Di Blas *et alii*, 2004):

- They read everything, including elements of HTML that are useful for visualization only (and do not convey relevant 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. The reader is invited to read aloud a page of a daily newspaper adopting the same strategy and measuring how long it takes until something relevant is read.
- They fail to convey the overall organization of the page, with the relative priorities of the different parts.
- They interleave the reading of content with the reading of links, with a total confusion for the listener. The listener can get the links' list (in alphabetical order), without the content, but s/he can't get the content without the links! In addition, even the list in alphabetical order is not effective; 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 links' meaning to clarify, or wrong and time-consuming moves in the site (Theofanos & Redish, 2003).
- 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 almost never works.
- Pages' layout and the "graphic's semantics" (that is, fonts' size and color, position on the page) are completely lost: the metallic voice of the screenreader will read one by one all the pieces of information of the page with the same emphasis and tone (the landmarks, the main content, the service links…), as if they all shared the same degree of importance.

The point is that screenreaders are… "screen-readers", that is, they basically read what appears on the screen, with a "book-reading" strategy, as if it were the most plausible equivalent to the "at a glance" comprehension of a sighted user. As we will argue later, the key to the solution is to separate the visual from the audio experience: not all that is written or visualized must be read, not all that is read by the screenreader must be visualized on the screen.

Some of the problems of the screenreader are "technical", in the sense they can be (almost) mechanically checked, while some other problems are more "conceptual", involving design techniques and usability issues.

In the next section we shall discuss, in detail, the last version (version 2.0, in preparation) of the W3C guidelines, and we will argue that they correctly address technical accessibility issues, but they are vague (if not wrong) on design or usability issues.

2. THE W3C GUIDELINES: A CRITICAL OVERVIEW

The W3C consortium made public a first set of guidelines in May 1999. The second version of these guidelines is currently under preparation (www.w3.org/TR/2003/WD-WCAG20-20030624). It consists of 4 major guidelines prescribing that an application should be *perceivable*, *operable*, *understandable* and *robust*. For each of the four guidelines, *checkpoints* (18 in total) are defined. For each checkpoint (that are considered normative) *definitions*, *benefits* and *examples* (non normative) are provided. Checkpoints are classified either as "core" or "extended": to conform to WCAG 2.0, the Required Success Criteria of

Core Checkpoints must be satisfied; the "extended" ones are additional checkpoints that may be reported in addition to Core conformance.
Let us comment in details the guidelines, as defined by the W3C: we should remind readers that while W3C address all kinds of disabilities, we will comment them taking into consideration blind users only.

Guideline 1: PERCEIVABLE. Make Content Perceivable by Any User

1.1 [CORE] All non-text content that can be expressed in words has a text equivalent of the function or information that the non-text content was intended to convey.

This is a concern about content: the idea is that graphic and visual content should have a text equivalent. Still, what equivalence means is very difficult to define (see *figure 1*): which words are equivalent to a painting, an image or a map? Should the text convey the look, the semantics, the emotion, or what else? It is obvious that mechanically satisfying the guideline will not ensure "real" accessibility.

Figure 1: from the Museum of Modern Art Website www.moma.org - is it a text equivalent to the picture's meaning?

1.2 [CORE] Synchronized media equivalents are provided for time-dependent presentations.
Time dependent presentations, with audio synchronized to changing images, for example, are clearly a major problem for blind users.

1.3 [CORE] Both [information/substance] and structure are separable from presentation.
This is an important guideline, the potential meaning of which is much deeper than the W3C guidelines seem to imply. We should remind the reader that the key problem lies with HTML where presentation is intermingled with content. In addition, the guidelines focus on presentation details (which are important) and substantially neglect the problem of presentation strategy (which is even more important than details). Furthermore they overlook the fact that for "reading aloud" a page a presentation strategy is necessary: an "oral strategy" very different from the one based on visualization (as it is the one commonly used for Web pages).

1.4 [CORE] All characters and words in the content can be unambiguously decoded.
This a technical requirement, necessary and, in a sense, obvious.

1.5 [EXTENDED] Structure has been made perceivable to more people through presentation(s), positioning, and labels.

This is a very ambiguous, and in a sense, incorrect guideline. It is (practically) impossible and (above all) useless to attempt to describe with words the "look" of a Web page. The reader may try this simple experiment: try to read the page of a daily newspaper to someone else. Very likely the reader will try to read aloud the semantics (e.g. "the most important news is… the second news is…") rather than trying to describe the visual aspects of the page. So the key point is to take a different point of view: a Web page holds a deep semantics, that is translated into a visual presentation. In order to make a page readable the best option is to start again from the semantics, not from the visual presentation.

1.6 [EXTENDED] Foreground content is easily differentiable from background for both auditory and visual default presentations.
In this checkpoint we spot again what we think is a major problem of the W3C guidelines: they focus on the symptoms neglecting the causes. The visual communication provided by a Web page is a mixture of background (same for each page) and foreground (different for each page): the overall semantics of the page, conveyed by background and foreground, must be translated into an "oral" communication.

Guideline 2: OPERABLE. Ensure that Interface Elements in the Content are Operable by Any User
2.1 [CORE] All functionality is operable at a minimum through a keyboard or a keyboard interface.
This is a necessary and obvious requirement, very important for users with operational disabilities.

2.2 [CORE] Users can control any time limits on their reading, interaction, or responses unless control is not possible due to nature of real time events or competition.
This is an important and necessary requirement. Our observation is that the corresponding implementation can be very difficult!

2.3 [CORE] User can avoid experiencing screen flicker.
We do not question the checkpoint, but it seems to be rather specific and too detailed: it could have been combined with other ones.

2.4 [EXTENDED] Structure and/or alternate navigation mechanisms have been added to facilitate orientation and movement in content.
This is a requirement concerning interactive content: every interaction provided by visualization and pointing mechanisms (e.g. the mouse) should be also made possible with different mechanisms. Important requirement, but difficult to implement; also we should work (in the research community) not at the mechanical reproduction of a visual interaction for a blind user, but to an "equivalent" solution. In other words, if normal sighted users get some "message" from a visual interaction, we should try to deliver (with different means) the "same message" to blind users, rather than trying to reproduce the interaction.

2.5 [EXTENDED] Methods are provided to minimize error and provide graceful recovery.
This is an obvious, but quite vague guideline. It is a feature desirable for all kind of users, although users with disabilities need to be especially "protected".

Guideline 3: UNDERSTANDABLE. Make content and controls understandable to as many users as possible
3.1 [CORE] Language of content can be programmatically determined.

Changes of languages are more easily understood with visualization (also for visual clues as, for example, use of different fonts) than by listening. We have experimented how difficult it is to listen to a sudden change in the language being used. Beside technical details, we think that change of languages should be banned, unless if forced by a quotation.

3.2 [EXTENDED] The definition of abbreviations and acronyms can be unambiguously determined.
Again we have realized that while looking to acronyms is "usable", listening to them makes very hard life for a user, if he can't look at the page. We think that acronyms should always have an alternative text, just like for images.

3.3 [EXTENDED] Content is written to be no more complex than is necessary and/or supplement with simpler forms of the content
This is a simplistic guideline. The problem of tuning content to the "profile" of the user is a standard one, and it has nothing to do with disabilities: a good application should always provide content of the proper level for all the different members of the intended audience.

3.4 [EXTENDED] Layout and behavior of content is consistent or predictable, but not identical
Again this is a true, but simplistic, checkpoint. 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 (required success criteria for checkpoint 3.4) is useless. It would certainly be more important to tell the designer how to shape content and navigation patterns in a consistent manner.

Guideline 4: ROBUST. Use Web technologies that maximize the ability of the content to work with current and future accessibility technologies and user agents
4.1 [CORE] Technologies are used according to specification.
The use of "unofficial" features of technologies must always be avoided, not just for users with special needs.

4.2 [EXTENDED] Technologies that are relied upon by the content are declared and widely available.
Availability of the technologies required for using the application is again desirable for all kinds of users, not just for the ones with special needs.

4.3 [EXTENDED] Technologies used for presentation and user interface support accessibility or alternate versions of the content are provided that do support accessibility.
This is a dangerous guideline: if the goal is understandable, we should also realize that current technologies for accessibility (e.g. current screenreaders for blind users) are not fully satisfactory. Technologies for accessibility still need a great impulse, and further research needs to be pursued.
Freezing the solution to the technologies available today is very dangerous.

Let us finally summarize our comments about the W3C guidelines:

- *Guideline 1: PERCEIVABLE. Make Content Perceivable by Any User*
  Some detailed guidelines are absolutely correct. But there is something confusing (if not wrong) about the presentation: apparently the guidelines fail to understand that the

semantics of the page should be the starting point, not the way the page itself is being visualized.

- *Guideline 2: OPERABLE. Ensure that Interface Elements in the Content are Operable by Any User*
  We do agree with most of the recommendations, which in general are more important for users with operational disabilities, with respect to users with visual disabilities.
- *Guideline 3: UNDERSTANDABLE. Make content and controls understandable to as many users as possible*
  This is the weakest part of the guidelines, vague and not usable, with the exception of the references to languages and acronyms, which are clear. There is a total lack of references to design principles and to semantics that should be the most important factor in guidelines concerning understandability.
- *Guideline 4: ROBUST. Use Web technologies that maximize the ability of the content to work with current and future accessibility technologies and user agents*
  These guidelines are concerned with issues so general, that the specific concern for users with special needs is unclear.

In the next section we describe a long-term research approach: WED - "WEb as Dialogue", in the frame of which accessibility is addressed in an original way.

3. A LINGUISTIC APPROACH TO ACCESSIBILITY: WED

3.1 "Dialoguing" with the Web

Experts and scholars from different areas compose the WED research team: linguists, usability experts, communication scientists, Web designers and engineers. Usability experts record (by means of a video camera and the thinking aloud method) sessions of use of "information intensive" Websites (such as Museum Websites); linguists and communication experts interpret them in the light of existing dialogue models and linguistic principles, highlighting their special characteristics, and the analogies and differences with respect to comparable natural dialogues. Web designers use the "understanding" of both types of dialogues in order to adjust design methodologies and in order to build interactive applications, based on the "oral channel", rather than on visual support.
The WED "dialogic" approach stems from an observation and a basic assumption:

- *Observation*: visually impaired users can't look at a screen; therefore the interaction must switch from the visual to the oral channel.
- *Assumption*: the interaction between a human being and a Website can be interpreted in terms of a dialogue (although a very peculiar one!)

The observation is self-evident; let's therefore move directly to the assumption. A Web session consists of the user getting pages, as a consequence of his/her "clicks" on the same pages. We can therefore consider a Web session as a sort of dialogue between a human being and a machine: the machine's conversational turns consist in offers of content/ interaction; the user's turns consist in the selection of an offer (by clicking or performing some equivalent action).
In order to clarify this point, we shall present now a short example of "dialogue" between a user and the Website of the San Francisco Museum of Modern Art (www.sfomoma.org); the transcription is quite "free", being its purpose just to show that a sort of dialogue *does* take place, and to make evident the basic features of the dialogue.

In the following "W" stands for Website and "U" for user. The pair "Tx", "end of Tx" is used in order to identify portions of content.

W: Welcome to the San Francisco Museum of Modern Art. [*T1*] You can choose between the following links: visit, info, calendar, education, membership, shop, exhibitions, collections, e-space [*end of T1*]. As an alternative, you can use the search engine. On evidence we have an exhibition of Romare Bearden, February 07, 2004 May 16 2004. Do you want the exhibition info? Or the interactive educational program (*connected to this exhibition*)? The news are: celebrate the art of Romare Bearden with Sunday jazz brunches, Examine Jay DeFeo's The Rose in February 28 Symposium, Meet architecture and design curator Joe Rosa at book signing on March 18, Learn about SFMOMA art auction 2004. Copyright © 2004 San Francisco Museum of Modern Art. The site is best viewed with Explorer 5.x or better. Do you want to know about the Privacy Policy and the Terms of use?

U: I want to explore the collections.

W: Overview of the collections. [*T2*] You can choose between: recent acquisitions, painting and sculpture, architecture and design, media arts and photography [*end of T2*]. As an alternative, you can use the search engine. Do you want to search our collections? We are developing an online database of works in our collections that will be available here in summer 2004. In the meantime, thumbnails of many SFMOMA works are available online through AMICO, the Art Museum Image Consortium. The pictures you're seeing are: 1. Brice Marden *Cold Mountain 6 (Bridge),* 1989-1991; 2. Robert Rauschenberg *Collection* (formerly Untitled), 1953-1954; 3. T.L. Pflueger *Bird's-Eye-View Renderings of the Proposed Airport at China Basin,* ca. 1928; 4. Dan Graham *Opposing Mirrors and Video Monitors on Time Delay,* 1974; 5. John Gutmann *Memory,* 1939.
[T1 is repeated]

U: I want painting and sculpture.

W: [*T2 is repeated*]. Do you want to use the search engine? Collections: painting and sculpture. The Department of Painting and Sculpture remains true to the Museum's founding vision of presenting substantive and adventurous artistic programming through the acquisition and exhibition of work that reflects the full range of twentieth-century art. You can choose between the following works of art: Jeff Koons, *Michael Jackson and Bubbles,* 1988; Robert Rauschenberg, *Collection* (formerly *Untitled*), 1953-54; Richard Diebenkorn, *Ocean Park #54,* 1972, Chuck Close, *Robert* 1997, Anselm Kiefer, *Osiris und Isis* (Osiris and Isis), 1985-87.
[*T1 is repeated*]

U: I choose Richard Diebenkorn's Ocean Park.

W: You're exploring the collections: painting and sculpture, and you're seeing the work by Richard Diebenkorn's Ocean Park #54, 1972.
Richard Diebenkorn's work took yet another wide swing in 1967 when he returned to non-objective painting and began a series of monumental abstractions which blended linear geometric structure with fields of luminous color. [*more text follows*]. Copyright © 2004 San Francisco Museum of Modern Art. Do you want to go back?

U: yes, I want to go back

W: [*T2 is repeated*]. Do you want to use the search engine? Collections: painting and sculpture. The Department of Painting and Sculpture remains true to the Museum's founding vision of presenting substantive and adventurous artistic programming through the acquisition and exhibition of work that reflects the full range of twentieth-century art. You can choose between the following works of art: Jeff Koons, *Michael Jackson and Bubbles,* 1988; Robert Rauschenberg, *Collection* (formerly *Untitled*), 1953-54; Richard Diebenkorn, *Ocean Park #54*, 1972, Chuck Close, *Robert* 1997, Anselm Kiefer, *Osiris und Isis* (Osiris and Isis), 1985-87
[T1 is repeated]
Comment: the user wanted to go back to "photography" but s/he couldn't do so directly.

U: I choose photography.

W: …

One of the aims of our research is to understand the essential feature of dialogues similar to the one above; although we are far from having completed our basic investigation, we have already understood some of the relevant features of a man-machine (Web) dialogue. The Web dialogue, with respect to human-human dialogue, is:

1. asymmetrical
2. rich (on the machine's side)
3. non-totally relevant (or "impertinent")
4. redundant
5. syntactic rather than semantic (on the user's side)

1. asymmetrical: the Website clearly holds the so-called "semantic dominance", that is, the power of putting forth content, topics, arguments etc. (Linell & Luckmann, 1991). The user can only make a choice, but s/he cannot put forth a new topic.
2. rich: thanks to its visual nature, a page can deliver a lot of content of different kinds. In particular, we find long menus of choices very unlikely to be found in oral (human-human) dialogues, for they would be cognitively difficult to handle.
3. non-relevant (impertinent): not all the parts of the site's conversational turns are relevant with respect to the user's selection, as we can see in the above short example (U: I want painting and sculpture. W: [*T2 is repeated*]. Do you want to use the search engine?)
4. redundant: some pieces of content are repeated many times; again this is very unlikely to occur in a human-human oral dialogue.
5. syntactic rather than semantic: usability experts, using log files, have detected that one of the most common moves in a Website consists in using the button "back" of the browser (Brincks, 2002). The reason is, quite often, that navigation has been conceived in such a (poor) way, that going back (e.g. to a list of items) is the only meaningful way to proceed. The move is therefore "syntactic" ("take me one step back") rather than semantic ("take me to content X"). Linguists call "phoric elements" all those parts in the text referring to other parts of the same text: "anaphora" in particular is the name for referring backward, to something already said. We can therefore conclude that "syntactic anaphora" is very common on the Web, and one of the major sources of ineffectiveness (for all users in general, and for blind users in particular).

3.2 Linguistics and Web design: a fruitful "WEDding"

As previously stated, even if a Website complies with the W3C guidelines, it can offer a very ineffective dialogue, not usable in practice by a blind user. The real challenge is to find design strategies and solutions (in terms of content structure and navigation capabilities) able to consider and solve the needs of visually impaired users.
In the last ten years there has been a growing interest in conceiving structured and comprehensive methodologies to hypermedia and the Web. In fact, in order to cope with the different problems the Web designer has to deal with, design issues must be solved in a systematic and modular way. The assessed design models and methodologies that have been developed (such as HDM, W2000, and OOHDM, UHDM, RMM, WSDM, WebML and many others) try to grant that each design activity addresses different concerns at the proper stage and at the proper level of abstraction (Schwabe D. *et al*.: 1998; Paolini P. *et al*.: 1999). Currently, the most comprehensive design models for Web applications - see for example UWA (UWA Consortium: 2001) and AWARE (Bolchini: 2003b) - consider the following elements:

- *Requirements*: the general objectives of the Web application and the strategy employed to communicate effectively with the users taking into account the goal of the developers and the motivations of different possible users.
- *Information Structure:* the way content pieces are organized within the applications.
- *Navigation architecture:* the strategies by which information pieces can be accessed and traversed using links.
- *Operations:* the set of functionalities (e.g. put an item in the shopping bag) and transactions (e.g. complete a commercial order) that the user might want to perform on the Website.
- *Interface and Presentation:* the orchestration of the visual communication strategies aiming at presenting effectively to the user the content, the navigation possibilities and the operations available.

According to communication theories, the designer of a Web application has a very powerful role: s/he sets the boundaries for communication and creates a stock of signs that the users may activate (potential "dialogues").
These considerations are the ground for modeling and designing the interaction between a Web application and its user as a particular kind of dialogue. The designer of a Website tries to imagine all the possible interesting conversations for the user and provide navigation mechanisms in order to make them possible. The designer thus plays a crucial role in the dialogue process because the range of possible interactions available to the user is actually defined by his intentions, expressed through the content, the navigation and interaction capabilities offered by the Web application.
Since a user experience is a dialogue, a Website can be considered a form of "dialogue generator", i.e. a device capable of supporting several different conversations with different types of users.
In the above overall scenario come the questions: "How should we design a Web application in order to generate successful dialogues with its users? How should we structure the information? How should we design the navigational capabilities? How should we consider the needs and limits of a particular category of users, such as visually impaired people? The WED research effort tries to answer these questions by observing traditional human-human dialogues, in order to grasp recurrent dialogic strategies used in a traditional oral interaction and interpret them in Web design terms. Indeed, there are some synergies between dialogic and Web design theories: in dialogic theories we have concepts and models that help

planning what to say (the so called *inventio* in ancient rhetoric, that is, the collection of all the ideas and pieces of information), which structure should be given to the content (the so called *dispositio*, the ordering of the elements according to the overall text's strategy), how we want to tell it (*elocutio*, the adequate wording of the meanings to be conveyed by the message) and how present it (*actio*, the actual performance) with respect to the audience considered (Cantoni & Paolini, 2001; Di Blas & Paolini 2001). Concepts used in Web design techniques such as W2000 (UWA Consortium, 2001) are very similar in the purpose: there are concepts for describing what we want to say (the so called *hyperbase design*), how to reach the information (the so called *access structure design*), how we want to tell it and which order we want to give to the different elements (the *navigation design*) and how we want to present it to the audience (the *publishing design*).

Web design techniques make a sort of *separation of concern*, for better understanding a Web application from different viewpoints and levels. These methodologies help designers in planning and effectively shaping a Website in all its complex communication elements, with a clear view of all the interactive mechanisms that stay behind it.

The WED Project found in W2000 methodology a very interesting ground for research, since many of the concepts used are easily applicable and comparable with dialogic theories. Let us consider, for example, the navigational dimension of a Website. In W2000 there are three possible navigational contexts: *structural navigation*, in which the user explores the pages belonging to the same "topic" (e.g. the pages corresponding to the same "painting"); the *semantic navigation*, in which the user navigates from one topic to a semantically related one (i.e. from a page describing a painting to the pages describing the "author"); *collection navigation*, in which the user explores a group of topics (e.g. all the paintings of a certain period).

Each navigational context can be described in a dialogic perspective. In *figure 2*, for example, an example of modeling of a page of the Oscar Awards Website (www.oscar.com) is shown, using W2000 (UWA 2001) notation. The design of the pages is straightforward: there is a page where a list of the winners is presented and the user navigates from the list to any of the winner actors.

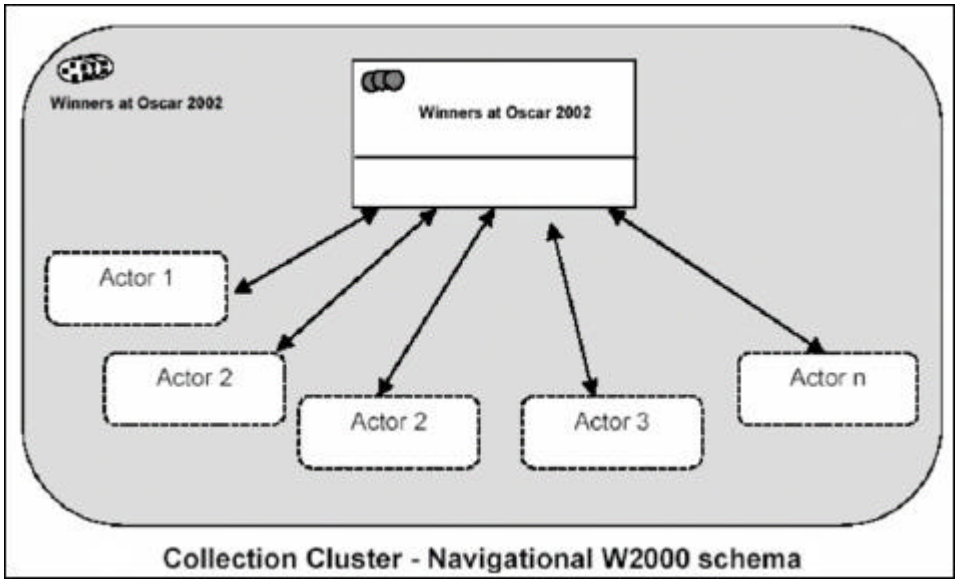

Figure 2 – Modeling of the 2002 Oscar winners Web page – www.oscar.com

From a dialogic perspective, here we are describing the introduction to a set of possible topics of conversation. In a traditional conversation, it would correspond to the question: " Which winner do you want to talk about?"
By considering this simple example, a lot of deeper design questions arise: "when the user selects a particular winner in the list and gets access to the details, how can she/he select a new one?" "Should she/he go back to the list, or should the Web application "guide" her/him and suggest what might come next?" In dialogic terms, if we consider partner A the application and partner B the user, the first solution could correspond to give the partner B the possibility to make a question to partner A like: "Could you repeat the names of the other winners you mentioned before?". The second solution could correspond to the offer of partner A: "If you want, I can tell you about another winner: winner x". Both from a dialogic and a hypermedia perspective, these strategies are very different and the result is a different interaction between the two partners.
Other crucial question is: what is the best way to organize the list, in order to make it effective? What criteria should be followed? Should items be listed in alphabetical order? Or should the names of the movies be used? How many items should be stored in the list? Should the list be split into sub-lists?
A designer should provide different answers according to the type of "channel" being considered: a purely oral dialogue requires different and specially tailored solutions with respect to a "visual" dialogue. If we can rely upon visual aids, we can display a list of 20 items, whilst if we are using the oral channel only that very same list becomes unusable: we would never list 20 names of awards winners in a normal conversation, asking our partner "which one do you want to talk about?"
The lesson is that purely oral dialogues use different strategies than visually supported ones; this is the reason why we do not think that accessibility can be reduced to technical issues. If the dialogue strategy is involved, than we must consider design and usability issues, in order to find viable solutions.

4. PUTTING WED TO TEST: THE MUNCH'S EXHIBITION WEBSITE

WED is a long-term research program; we did try, nevertheless, to immediately experience with real-life problems and practical solutions. We have exploited the preliminary results of WED for the design of a site supporting an exhibition of prints by Edgar Munch, that took place (in the late spring and summer of 2003) at the Staatliche Museen of Berlin. The www.munchundberlin.org Website complies with (almost all) the guidelines of W3C.

Fig. 3 – Website of the Edvard Munch's Exhibition in Berlin (www.munchundberling.org) – Home Page

In terms of design and implementation strategies, we have developed a few empirical "tricks":

1. *Content is divided into sections (usable "chunks") and organized into a "usable design"*
   The W2000 design methodology was used for carrying on a consistent organization of the whole site, in order to make it easily understandable by any user.
   This can be considered as a practical implementation of the W3C guideline 3, concerning understandability.
2. *Page schemas*
   If we were to read a newspaper to a blind user, we would never start from top left reading in details information, titles, texts, advertisements, captions, etc., but we would offer our interlocutor a sort of synoptic view of the basic pieces of information, highlighting the most relevant ones, waiting for her/him to decide what to choose. We have therefore found a way to force the screenreader to read a "page schema" first, that is, a short summary of the basic sections of the page. Therefore the user can directly access the section s/he's interested in. The page schema is a *purely oral* feature: it reflects the page's conceptual organization but it is not visualized as text on the screen.
3. *Reading order of the page's content*
   Especially while navigating a site, the user's selection of links are mainly "semantic" that is, explicit requests of content. In a natural dialogue if our partner tells us "do you want to know about Botticelli?" and we answer "yes, please, go on", we expect her/him to speak about Botticelli and not about let's say the copyright of the book s/he's taking the information from, the other content s/he could speak about or the services s/he could offer us. The very same thing should happen when "dialoguing" with a site: if we choose the link labeled "Botticelli" this clearly means we want to access content regarding this painter, although the page we reach may host many other additional pieces of information, such as the landmarks' list, the service links, etc. While a sighted user easily skips all the information s/he's not interested in, blind people have to listen either to the whole page or to try to directly access the links' list, but with the limitations highlighted above. In the Munch's site (see *figure 4*), the problem is overcome thanks to the reading strategy of the screenreader, programmed

to read the main content of the page first (immediately after reading the "page schema"). This relates to several guidelines, including 1.5 and 1.6.

Figure 4 – Munch's exhibition Website - Example of template describing a painting

4. *Consistency across pages*
   Pages' structure, that is, images, texts and links' positions, remains almost the same in the whole site, enhancing the user's orientation (see *figure 4*). All the pages are designed according to only two basic templates. This is a combination of what is supposed to be the goals of guidelines 1.5, 2 and 3.
5. *Links' names*
   The screenreaders allow immediate access to a list of the page's links; in order to still "mirror" the page's organization, we changed the links' names, specifying the first and last links of a section. The screenreader therefore will say "First link of [name of the section]". In this way, the information about the basic sections of the page is not lost.
6. *Prints' descriptions*
   The descriptions attached to a work of art usually do not bother about sketching its basic features and elements, for they are meant for a sighted user who can look at it and therefore knows, let's say, how many people are portrayed, how they are dressed, what they're doing, etc. In the Munch's site, the descriptions of the prints were prepared having in mind a user who can't see the print: all the elements of the print itself are introduced and then commented.

We have also tried to improve the "anaphoric strategy", i.e. how the user can go back to previously visited pages. Somehow they correspond to guidelines 2 and 3: improving operability and understandability of the application.
The following are some among the most relevant mechanisms that we have introduced:

7. *Visited & missing*
   Sometimes a user wonders: "what have I already seen? What's missing?" The "visited & missing" command allows answering to these questions: all the pieces of information already seen (while navigating) or missing (not yet visited) are shown in a very efficient synoptic view, allowing the user (blind or not) to efficiently enhance his/her visit.
8. *Just visited*
   This command allows going back following a semantic pattern, rather than a syntactic one: the user can directly revise the main topics of what s/he has seen so far. This

avoids tedious repeated uses of "back" of the browser (very difficult for everybody, and for blind users in particular).

9. *Decisions*
   This command allows the user to quickly revise the situations where s/he has taken major navigation decisions (i.e. selection into indexes). This feature allows a faster and simpler navigation across the application.
10. *My history in the site*
    One of the most cumbersome navigation moves (not only for blind, but also for sighted users) and still one of the most frequent ones is the button "back" of the browser, most of the times used not to visit the same content again but to resume navigation from a previously visited node.
    A sighted user can (more or less easily) detect whether s/he has reached the wished page (otherwise "back" must be used again) and can quickly (depending on the designer ability) locate, within the page, the navigation link that will start a new exploration.
    A blind user, instead, will have to listen (at least for a while) to the screenreader reading the page's content, in order to understand first whether s/he's reached the right page, second where the precise link s/he was looking for is on that page.
    The command "My history in the site", implemented in the Munch's site, tries to overcome the problem: it offers a list of the "semantic" steps the user has made that far in the site, thus facilitating a quick re-selection of a 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 think you mentioned Botticelli some minutes ago".

The Website for Munch's exhibition has been online for a few months; nonetheless, we have received a variety of positive comments, from users whom we did not know, that we have no space to analyze in this paper.
One specific comment, addressing the overall design, was 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 favorite 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 before all will be very convenient for users who surf only occasionally..."
> (*mail from mister Martin Kirchner, May 18th, 2003; bold fonts by the authors*).

5. CONCLUSIONS AND FUTURE WORKS

Accessibility is a complex matter involving people (users with disabilities) who have already strong disadvantages; therefore it must be dealt with precaution.
We propose, in this paper, to define and treat accessibility as a branch of usability: if we say that an application must be usable by all users, then users with disabilities must be included too. If this is deemed to be too difficult, then we, as designers and developers of Web applications, must clearly and carefully define which "user profiles" we mean to take into consideration and/or which ones not.

PROPOSITION 1: each Web application should clearly state which categories of users with disabilities have been taken into consideration (or not) as a target for accessibility.

The guidelines being proposed (version 1.0) or being considered (version 2.0) by W3C, within the WAI initiative, deal with a specific class of problems affecting usability: problems stemming from a bad use of technology. They do not address at all the question whether the application is actually usable (at least in a decent manner, and possibly in an optimal manner) by users with disabilities. As a matter of fact, it is known that it is possible to build applications fully compliant with WAI guidelines, and still unusable by users with disabilities. This situation is "politically" very dangerous since in several European countries (and Italy is a clear example of this) pieces of legislation are being approved, which make accessibility mandatory. If this is a positive aspect on its own, it becomes very negative when we consider that many bureaucracies (and the Italian one is an example) will take the easy way: "accessibility = WAI guidelines". What is likely to follow is easy to guess: "certifiers" (automatic or not) will certify that an application is compliant with the guidelines; the bureaucracy will be satisfied, certifying that the application is compliant with the law; the application developer will be satisfied since accessibility has been "achieved" at a reasonable cost; disabled users will be denied effective access to the Web application as before, since it won't be actually usable for them.
We think that researchers and practitioners working on this sensitive area of accessibility must fight in order to avoid this development.

PROPOSITION 2: a "manifesto" (signed by researchers and practitioners) should be used to make clear to all the politicians and bureaucrats, world wide, that satisfying the WAI guidelines does not mean at all that an application is accessible. Therefore it also should be made clear that "accessibility certification" can't be dealt with in a superficial manner.

If we equate accessibility to a usability problem, we acknowledge that ensuring it and checking it is more difficult, less automatic, and more debatable with respect to the expectations of those who rely on technical guidelines as the "solution". We can contribute to accessibility in two possible ways (with analogy to what has been done for usability): improving the way we check accessibility and developing tips and guidelines (based on best practices) in order to help developers in achieving it (at some level difficult to assess formally).

PROPOSITION 3: researchers and practitioners should work out a common way to carry on tests for assessing whether an application is accessible, to what degree and for whom; these "usable accessibility" tests must be based on a shared set of check lists and "assume" compliance with WAI guidelines (that can be checked separately).

PROPOSITION 4: researchers and practitioners should work out a set of "best practices" for designing and implementing really accessible applications (not nominally accessible); these best practices are a necessary complement to WAI guidelines.

If the above propositions are quite general, we can now draw the conclusions on our specific research, which is confined to a specific kind of disability, blindness, and to a specific technology, screenreaders and alike.
Within these limits we describe our current achievements and the broad lines of our future research:

- *Scientific background:* we believe that it is wrong to start from a Web page, conceived for being looked at, trying to make it readable. It is better to start from a step before: the "semantic of what is being said". This content must be delivered through an "oral channel", as opposed to the page, which is based upon a "visual" or "multimedia" channel.
  In addition we have found out (from empirical evidence and from linguistic literature) that the overall "dialogue strategy" is different, if the oral channel is being used, with respect to a channel with visual support.
  We are working, as far as basic research is concerned, in two promising directions: an empirical work of comparison between human-human oral dialogues, and human-Web (visually supported) dialogues. The development of a "dialogue model" (based upon semantics, rather than on syntax or rhetorical schemas), capable of capturing the essence of both types of dialogues, is what we are working at, together with a group of linguistic researchers.
- *General design*: we have already revised our previous design methodology (W2000), coming up with IDM - Interactive Dialogue Model. IDM is a tool (set of concepts and notation) to design an interactive application in a "conceptual", manner, i.e. independently from the specific channel that will be used for delivery. In a second stage the application will be "transformed" according to the need of the specific channel (oral, visual, ...).
  Although when the Munch's Website was developed, IDM was not fully defined, its basic principles were already there and were actually used; the result has been a very usable design, i.e. a structure of the application where consistency and self-evidence were emphasized. The benefit for the user is that s/he can easily understand the structure of the application, and how to move around.
  We are currently working on this notion of "usable design", that in our opinion lays at the very heart of true accessibility: if the user can understand the design and the motivations behind it, s/he will find the application more "natural".
- *Presentation strategy*: an oral presentation is radically different from a visual-supported presentation. We therefore came to the conclusion that it is useless to start from the page (from its "look") and to try making it accessible.
  Our goal is therefore to develop guidelines for an effective "reading strategy", based upon the intended semantics and the "raw content" of a page, rather than upon its look.
  A reading strategy can be considered at different levels of granularity: a section, a page, a group of pages, etc. We have found some simple rules, that we have already applied for Munch's Website, but much more research is needed.
- *Anaphoric strategy* (i.e. how to "go back"): one substantial contribution of linguistics to our research has been the recognition that a number of problems for accessibility stems from the practice, for the Web, of forcing the user to "go back" to already visited pages. This practice is ineffective for "normal users" and devastating for blind users (who must go through the whole page before getting to the point of interest). We have also understood the analogies between "going back" and the practice of "syntactic anaphora", as defined in linguistics (Di Blas 2003b).
  Our research works in three directions: trying to deploy navigation strategies that minimize the practice of "going back"; trying to improve the mechanisms implementing syntactic anaphora; trying also to implement mechanisms of semantic anaphora.
  The first two directions were somehow already considered for the Munch's Website, but we need to improve, in a number of ways, the solutions devised there. Considering

semantic anaphora, i.e. moving back "to content" rather than to pages, is new and we need to break some new ground.

- *Screenreaders*: screenreaders have a basic limitation: i.e. they are not conceived for implementing an explicit reading strategy. For the Munch's Website we had to recur to a number of "tricks" in order to force the screenreader to implement what we had in mind. The strategic solution, that we are aiming at, is different: the reading strategy should be explicitly defined (at least at page level); the reading strategy should be "represented" in some ways (e.g. through "reading tags" or "reading instructions") in the page; a new generation of tools, "page-readers" should be used to implement the reading strategy.

The overall conclusion is that accessibility is scientifically challenging but also a socially relevant issue involving disadvantaged users. We must therefore be very careful with our promises and our results: mistakes and exaggerations are worse than usual.
Last, but not least, we, as a scientific and technical community, must warn decision makers that easy and ready-made solutions are not there, yet.

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 Gobbetti, Marco Marini, Fulvio Prisinzano) and from USI (too many to list them all) who passionately work 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's site was developed, and Andreas Bienert (from the Staatliche Museen of Berlin), who, together with his wonderful staff, made the experimentation possible.

REFERENCES

[Bolchini 2003a] Bolchini D., Mylopoulos J., *From Task-Oriented to Goal-Oriented Web Requirements Analysis*, In Proc. of the Conference Web-Wise Sustaining Digital Resources 2003.

[Bolchini 2003b] Bolchini D, Paolini P, Randazzo G., *Adding Hypermedia Requirements to Goal-Driven Analysis*, In Proc. of the Conference Requirements Engineering 2003.

[Brinck 2002] Brinck, T., Gergle, D., Wood, S.D, *Usability for the Web: Designing Websites that Work*. San Francisco, San Diego, New York, Boston, London, Sydney, Tokyo: Morgan-Kaufmann, 2002.

[Cantoni 2001] Cantoni, L. & Paolini, P. *Hypermedia Analysis: Some Insights from Semiotics and Ancient Rhetoric*, Studies in Communication Sciences vol. 1, 2001, pp 33-53.

[Di Blas 2001] Di Blas, N. & Paolini, P. *La Retorica e il Web*, in D'Alessandro, P. (ed.), Internet e la filosofia, LED, Milano, 2001, pp 39-49.

[Di Blas 2003] Di Blas, N. *Deixis & Formulations – When the Web speaks of himself*. SCOMS – Studies in COMmunication Sciences, 2003.

[Di Blas 2003a] Di Blas, N. & Paolini, P. *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, 2003, pp 221-233.

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

[Di Blas 2003] Di Blas, N., Paolini, P., M. Speroni, A. Bienert, *Listen to a Website: accessibility (beyond current standards) and a market opportunity*, ICHIM 2003 proceedings, Paris, 2003.

[Di Blas 2004] Di Blas, N., Paolini, P., Speroni M., Capodieci, A., *Enhancing accessibility for visually impaired users: the Munch's exhibition*. In Bearman D. & Trant J. (eds), Museums and the Web 2004. Selected Papers from an International Conference, Archives & Museum Informatics, Arlington, Washington, U.S.A., 2004.

[Linell 1991] Linell, P., & Luckmann, T. *Asymmetries in Dialogue: Some Conceptual Preliminaries.* In I. Markova & K. Foppa (Eds.), Asymmetries in Dialogue Harvester Wheatsheaf: Hemel Hempstead, 1991, pp. 1-20.

[Nevile 2002] Nevile, L. & McCathieNevile, C., *The Virtual Ramp to the Equivalent Experience in the Virtual Museum. Accessibility to Museums on the Web*. In D. Bearman & J. Trant (Eds.). Museums and the Web 2002. Selected Papers from an International Conference. Pittsburgh: Archives & Museum Informatics, 2002, pp 93-99.

[Paolini 1999] Paolini P., Garzotto F., *Design patterns for WWW hypermedia: problems and proposals*, in Proc. ACM HT '99, Workshop on "Hypermedia Development: Design patterns in Hypermedia", Darmstadt, 1999.

[Schwabe 1998] Schwabe D., Rossi G., *An Object Oriented Approach to Web-Based Application Design, Theory and Practice of Object Systems*, J. Wiley, 1998.

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

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

[UWA 2001] UWA Consortium, *D7: Hypermedia and Operation Design*, Deliverable Project, 2001, http://www.uwaproject.org.

Online resources:

http://www.w3.org/WAI/: Web Accessibility Initiative
http://www.w3.org/2001/di/: Device Independence Activity
http://www.w3.org/2002/mmi/: Multimodal Interaction Activity
http://www.w3.org/Voice/:"Voice Browser Activity - Voice enabling the Web!