Designing Aural Information Architectures

Davide Bolchini
University of Lugano
via G. Buffi, 13
6900 Lugano, Switzerland
+41 (0) 58 666 47 13
davide.bolchini@lu.unisi.ch

Sebastiano Colazzo
Politecnico di Milano
P.zza L. Da Vinci, 32
20133 Milano, Italy
+39 02 2399 9627
colazzo@elet.polimi.it

Paolo Paolini
Politecnico di Milano
P.zza L. Da Vinci, 32
20133 Milano, Italy
+39 02 2399 3520
paolo.paolini@polimi.it

Daniele Vitali
Politecnico di Milano
P.zza L. Da Vinci, 32
20133 Milano, Italy
+39 02 2399 9627
vitali@elet.polimi.it

ABSTRACT

Nowadays websites use, above all, the "visual channel" to communicate content, functionality and navigation/interaction capabilities. Users who, permanently or temporarily, cannot use their eyes to interact with an application (not only visually-impaired but also people who cannot look at the screen while interacting) need a new paradigm of interaction, based on an "aural" access to information. Current technological approaches for developing "aural hypertexts" (such as speech markup languages, or code optimization techniques for screen-readers) provide technical solutions to transform a visual interaction into an aural one, failing to support adequately the overall design process. This paper presents an overview of critical design issues to consider when conceiving aural hypertexts, namely when designing "aural" information architecture, navigation and interaction features. Examples of possible "aural" design solutions are also discussed to support specific requirements. The work is based upon real-life project experience in both designing web applications and developing applications for visually-impaired users.

Categories and Subject Descriptors

H.1.2. **[Models and Principles]:** User/Machine Systems – *Human factors, Human information processing.*
H.4.3. **[Information Systems Applications]:** Communication Applications – *Information browsers.*
H.5.2. **[Information Interfaces and Presentation]:** User Interfaces – *Input devices and strategies, Interaction styles, Standardization.*
H.5.4. **[Information Interfaces and Presentation]:** Hypertext/Hypermedia – *Architectures, Navigation, User issues.*

General Terms

Design, Human Factors, Standardization.

Keywords

Accessibility, Aural Hypertexts, Design, Interaction Primitives.

1. INTRODUCTION

Web applications are usually designed to communicate through a graphics interface. Through the visual channel, a website (from very simple ones to complex applications), conveys various messages simultaneously, basically concerning the content and the interaction/navigation capabilities (buttons and links). Besides these elements, other important clues are communicated through the visual interface: the graphical layout of the elements of the page (perceivable by looking at the screen), the grouping of the links, the distance and the order of the elements, their size and relative proportions, the chromatic code used to characterize links, content, headings, titles, subtitles, buttons, and so on. These design elements are crucial for the quality of design and for the usability of the overall application.

By aural hypertexts we mean hypermedia applications which have to be consumed via the aural channel, i.e. "by listening to them" instead of "looking to them". There are at least two scenarios in which aural hypertexts are increasingly needed:

a) visual impairness: people with visual disabilities (currently relying on "screen readers" software to browse websites) still experience frustration in getting a communication which is far to be effective through the aural channel.

b) mobile contexts: interactive applications supporting mobile scenarios, such as visiting a city, a museum, or car navigator applications should assume that the user cannot (always) look at the screen while interacting. Aural-based interaction design solutions should be devised.

Aural hypertexts may be the aural "counterpart" of existing interactive (web) applications (e.g. the website version accessible to visually-impaired users), or may be a stand-alone application conceived and designed from scratch to meet specific purposes.

This paper focuses on aural hypertexts as counterpart of standard websites, illustrates some crucial design issues for conceiving them – taking into account usability and effectiveness of the communication – and uses specific examples to anchor the research to tangible case studies. In this paper, we will not explicitly deal with "mobile applications" which require considering some additional requirements related to the mixed use of both the aural and visual channel (i.e. synchronization). On the basis of the experience gained in the development of traditional web application through systematic design methods, and by having developed specific aurally-oriented design patterns and solutions in real-life project contexts, issues at various design levels are discussed.

The results of this paper may be used by aural designers to gain an insight about what aspects to carefully bear in mind when designing for the aural channel. On the other hand, ready to use solutions or design recommendations can be taken and applied, independently by the specific technology adopted.

The paper is structured as it follows. Section 2 reports some background for this work, including international accessibility standards (W3C guidelines) and specific implementation languages supporting aural hypertexts. Section 3, 4, 5 and 6 discusses how the linearity of the vocal channel can affect the design activity and presents important design issues to consider when conceiving aural hypertexts, especially when dealing with "aural" information architectures and navigation. Finally, section 7 draws some conclusions and points to current and future research.

2. RELATED WORK

The broad objective of enabling aural access to web applications has been addressed for some years both in the research and in the industrial arena [1]. Many solutions have been developed using different technologies that have lead in turn to many different standards[2][3], being later proposed or enforced by various agencies [4][5].

Among the complex picture of the efforts (at various levels) done to address the development of aural hypertexts, we quote two salient areas: the area of the international accessibility standards (W3C guidelines), and the area of specific implementation languages in support to aural fruition.

2.1 Accessibility guidelines

The World Wide Web Consortium (W3C) developed specific guidelines for web application accessibility. The most important set of recommendation is the Web Content Accessibility Guidelines [3] that explains how to create accessible web content. WCAGs are composed of 14 general principles that aim to avoid the dependency of hypertext interaction on the visual channel. Other support guidelines have been developed, such as Authoring Tools Accessibility Guidelines [6], that explain how to develop authoring tools to be compliant with WCAGs; or the User Agent Accessibility Guidelines [7], that explains the user agents (such as browsers or multimedia players) responsibilities in meeting the needs of user with disabilities.

In general, web pages that are compliant with the W3C guidelines are intended to be interpreted by screen readers (i.e. [8]), aural browsers (i.e. pwWebSpeak [9]) or specific displays for visual impaired users, such as Braille Displays.

The international community (both researchers and practitioners) is recognizing that just following W3C guidelines is not at all enough to offer "accessible" aural websites. W3C guidelines are more oriented to improve the accessible versions of existing websites, rather than supporting in re-thinking the whole design process into an aural perspective.

Consequently, the complexity of existing website pages basically remains also when following the W3C guidelines. Therefore, one of the main frustration still experienced by visually-impaired users using screen readers is that they are forced to listen for long to irrelevant information on the page until being able to catch the relevant content [1].

In a situation where there is a lack of common principles in developing effective audio interfaces [10], a deep rethinking of the design principles for aural hypertexts is strongly needed.

2.2 Domain specific languages

Instead of responding to this lack of "design support", the industries develop implementation solutions to develop audio interfaces.

Two main languages lead the market in the field of voice applications: VoiceXML [11] and SALT [12].

VoiceXML is a XML language broadly used in telephone applications such as call centers or voicemail [13]. It specification has three main goals: simplify the creation of web-based voice response services, enable integration of voice access to existing intranet services and help enable new voice capable devices [14]. Voice applications can take advantage of the same application development and deployment paradigms that are commonly used by visual applications. As visual browsers interpret HTML documents, voice browsers interpret VoiceXML documents (aural hypertext).

In spite of that, VoiceXML does not allow integrating the aural channel with the visual one. Adapting the existing content to be vocally accessible often means to develop a new application from scratch, since good speech user interfaces are significantly different from their visual counterparts [15] and manually maintaining a correspondence between HTML and VoiceXML documents is a time-consuming and error-prone task [16]. Automatic transformation of HTML pages into VoiceXML has been attempted with good results [17][18]. The main attempt to drop the barrier between visual and audio channel is the XHTML+Voice Profile [19]. Web pages developed in XHTML are enriched with VoiceXML, allowing visual and aural access at the same time.

SALT (Speech Application Language Tags) is an XML based markup language that is used in HTML and XHTML pages to add speech output and interaction capabilities to web based applications. It is designed to allow multimodal and telephony-enabled access to information at the same time, and allows various types of user input, like speech, keyboard, mouse, keypad, or stylus. SALT is developed by Microsoft, and plug-ins are available to enable MS Internet Explorer to access SALT enabled applications. The markup language permits a clean integration of speech with web pages, separation of the speech interface from business logic and data, on a various range of devices [20]. The principle of SALT design is that programmers can use web pages as model for the spoken dialog interaction, using the single page to achieve each sub-goal of the task [21]. SALT is younger than VoiceXML and is not fully adopted in the market of Speech Applications.

These languages are crucial for the implementation of aural hypertexts but, of course, they cannot solve the design issues needed to be coped for an effective user experience. In other words, language-oriented solutions allow having web content and functionality "readable" by aural technologies, but they cannot answer to the question: which is the better communication strategy to convey web content to the users?

A first attempt to overcome the above issues has be done by the WED (Web as Dialogue) research project, carried out at the University of Lugano in collaboration with the Politecnico di Milano. This research initiative aimed at gathering theoretical, methodological and practical insight from a simple hypothesis: the interaction between a user and a website can be interpreted as a sort of dialogue. The project produced relevant results both for the linguistic community and for the web engineering ones. For example, a web design model (called IDM: Interactive Dialogue Model) entirely based on dialogue primitives has been defined [22][23][24][25][26].

A number of fruitful research directions have born from the results of the WED project, including a systematic reflection on the requirements to be considered for support the design of aural hypertexts. The paper starts from these requirements and discusses design issues to be coped for addressing them.

3. DESIGNING AURAL HYPERTEXTS

When dealing with aural hypertexts, an important design aspect to consider is the *linearity* imposed by the aural channel. Listening to content, links, and interface functionality happens strictly in linear sequence, and this completely changes the interaction paradigm with respect to an interaction relying on the visual channel. Whereas a visual-based interaction allows mastering the complexity of a page at a glance, aural hypertexts should consider that the user processes items *in sequence*, thus missing a synopsis of the interaction context. This paradigm shift has deep implications on at least two design levels:

1. *Information Architecture Level:* it is strictly related to the user's need to understand, memorize and use the overall structure of the hypertext.
2. *Page Level:* it is related to the user's need to understand, access and navigate the content of a single page. Looking at a web page, a user may identify at glance the overall structure of the page, and decide what to "consume" and in what order. Dealing with aural hypertexts, "looking at a web page" corresponds to "sequentially listening to it".

4. AURAL INFORMATION ARCHITECTURES

Information architecture is the term usually employed to identify how content is structured within the hypertext, the relationships between various content pieces and how these are accessible through navigational paths. In this section, we will discuss the various issues involved in aural information architecture design, providing some examples of solutions.

4.1 Providing ways to understand the hypertext structure

Information intensive hypertexts contain a large amount of data, and one of the biggest problems for the inexperienced user is to understand what the application is about. One of the most time-consuming tasks is to learn what are the topics covered by the application, what kind of website is it, how the navigation of a website works, where the links are, what their meaning is, and so on.

In traditional visual-based interfaces, designers can rely on the graphic interface of the home page to convey these messages (when they do it). However, in aural hypertexts, designers need to find solutions to address these needs. Aural applications need to communicate to the user (at the beginning and whenever necessary) two kinds of meta-messages concerning the information architecture:

a) An aural *semantic map* of the whole hypertext (answering the question: what is the application about?)

b) An aural *quick glance* of the hypertext (answering the question: how is the application structured?).

Using website maps seems the most appropriate solution to answer both requirements. Whereas website maps are traditionally used by users as a last resort to find a way into the maze of information of the application, "aural" users would use a kind of site map right at the beginning and as a constant reference point throughout the interaction.

However, instead of just "listing" all the possible links available in the application, "aural" site maps should follow specific requirements in order to be effective in supporting the aural fruition:

- Content items – when possible - should be "typed" into meaningful categories which allow identifying the actual topics treated in the application. For example, in the case of a museum website, it is important to communicate to the user that the application talks about "paintings", "prints", "artists", "techniques", "artistic movements", and so on, before presenting specific paintings, authors, techniques.
- Once clarified these kinds of topics, the user should be able to understand and use the strategies employed by the designers to access this content. Therefore, the access structure for each kind of topic can be made explicit to the user. For example, the map should convey that "prints" are organized "by artist", "by historical period", highlighting a selected set of "masterpieces", by an index of "all prints by alphabetical order", and so on.

These basic messages allow the users to get an overview of the content to be found in the application and to find a suitable way to access and explore it.

Aural hypertexts force the presentation of this information to be strictly sequential and require a number of interaction steps. For instance, a possible strategy to present the abovementioned structure of the application is an aural tree-like navigation menu, as the one shown in Figure 1.

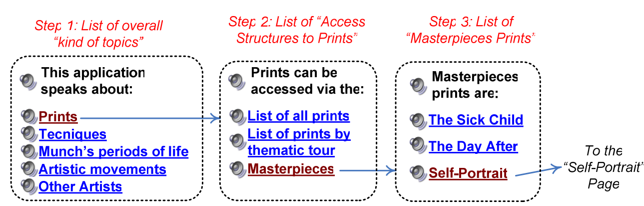

Figure 1. The semantic map: an example of aural presentation

At the first step, the hypertext presents the list of the topics to the user allowing him/her to make a selection. After the selection, the user can decide either to listen the description of its semantic (i.e. the topic "print" provides a description of a print of the exhibition) or to select one of the possible access structures (i.e. the "List of all prints", the "List of prints by thematic tour" and the "Masterpieces"). The last step presents a list of topic instances allowing the user to access the actual content (i.e. "The Sick Child (Sick Girl)", "The Day After", "Self-Portrait", etc.).

Note that each interaction step corresponds to an "aural page" and requires the user to make a selection. This consideration leads us to introduce additional issues to take into account: what is the better way to organize lists of items (especially when they are long) for the aural interaction? How can the navigation within the information architecture be properly managed?

4.2 Designing long lists of items

Access paths to content (a key element of the information architecture) are based on lists of items to be selected from. By iteratively selecting from a list of items, the user will eventually end up to reach some content. Long list of items, however, represents one of the hardest problems for aural interaction. When a list is "listened to", the user has to listen to a synthesized voice reading each element of the list, thus being not able to capture the meaning of the list at a quick glance. It is thus clear that a list of 20 prints becomes quickly unusable for an aural user: s/he will likely stop reading the list very soon.

One strategy to make lists more usable is to provide a sort of *executive summary* of any list of items (especially for the long ones) answering questions such as: "What is the list about? What is coming ahead?". In this way, before starting reading the long list item by item, the user can get an idea of the overall list and decide whether it is worth exploring or not. Moreover, this list summary may provide precious introductory content at the top of the list, such as a brief anticipation of the topics described, and the criterion by which the list is organized. For example, for introducing a large set of masterpiece from the collection, the introductory text may say: "The curator selected the following 20 masterpieces of prints from the collection, listed by their artistic relevance and value".

A second – complementary – strategy to design suitable lists for the aural navigation is to partition long lists in smaller, meaningful chunks. Long lists can be grouped in sub-lists according to some *semantic criterion*, in order to let the user manage shorter lists. For example, instead of presenting a unique list of 30 prints, there could be a list of few themes around which the prints are organized ("portraits", "landscapes", "everyday life", etc). The user should first select a theme, and then access to a list of prints concerning the selected theme. Similarly, more syntactical criteria can be used to organized lists. For example, if designers want to provide the list of all prints (e.g. 100 or more), they can be grouped by alphabetical letter to make this long list shorter and more easily manageable.

4.3 Designing "backward" navigation

As a matter of fact, large part of the user interaction in complex hypertexts consists in going back to previously visited pages. Backward navigation is usually employed for different purposes:

a. *to recover orientation*: when orientation is lost, the user feels more comfortable to get back to a familiar page (e.g. the homepage), where s/he can more easily spot access points to the content;

b. *to retrieve a previously visited content*: the user remembers to have visited a content but does not recall its position in the hypertext. Going backward page by page should assure to eventually retrieve that content;

c. *to retrieve a previously visited list of items*: after having selected an item from a list, the user usually "goes back" to the list to select another item.

The "go back" possibilities may be offered by the hypertext itself, or by the browsing technology ("back" button"). However, such functions only provide a syntactical mechanism, similar to the physical browsing of a book (going back "page by page"). When dealing with aural hypertexts, browsing back page by page can be very frustrating for the user, since s/he is forced to (linearly) listen to the content of all intermediate pages being accessed (which are not all meaningful for his/her goal). Obviously, this is not a problem for the visual interaction because the user quickly visually scans the accessed page as s/he goes back. However, during aural consumption, the user has to "listen" to some content of each page to understand whether or not it is the one looked for. If the web page is not optimized for this aural scenario, the user may take some time to get to some "identifying" content of the page.

A design strategy to address this problem is to provide a *semantic back* mechanism, which should emphasize the history of the visited topics (the actual content), rather than the sequence of all the physical pages visited. Let us consider, for example, the history path shown in Figure 2. The user navigated through a number of pages of different kinds: from the page describing the life of Raffaello, to the list of his masterpieces, to the description of masterpiece 1, to the list of related works of art (e.g. sharing a similar theme), to the description of masterpiece 2, then to the homepage, and so on. Let us assume now that the user wants to go back to see the life of Raffaello. The page history would require 7 navigational steps to reach that page. Assuming that the user has to listen to a significant part of each page to understand what the page is about – e.g. at least 10 seconds/page –, this back navigation will require approximately 70 seconds to reach its goal.

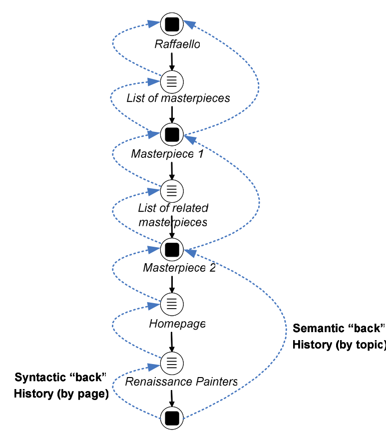

Figure 2. Page history vs. Topic history

If designers offer a backward strategy based on a history gathering the actual topics visited and not considering the "list pages", the backward navigation would be much more agile and usable. The topic history would need only 3 navigational steps to reach the target page, with a reduction of access time greater than 50%.

Therefore, in case that the user wants to retrieve a previously visited content, providing a "topic history" (rather than the "page history") allows the user going back "topic by topic" and not being forced to visit intermediate non-content pages.

This design strategy more closely corresponds to a typical situation happening in human dialogues, where a partner asks to the other partner to recall a topic previously talked about (e.g. "let us get back to talk about Raffaello"). In human dialogues, it is natural to point directly to that content and start talking again about Raffaello. On the other hand, rewinding all the turns of the dialogue until Raffaello pops us would be completely unnatural, artificial and frustrating. Similary, the topic history gets closer to this more natural dialogic situation, in which a topic is more directly pointed to, skipping the "turns" of the dialogue (i.e. the pages) which do not contain any potentially relevant content.

Along this design direction, a more advanced strategy can be pursued. Instead of having the user get back "topic by topic" to the previously visited content, an explicit group of previously visited topic can be accessible. For example, the user could access at any time to the set of already visited artists. In this way, s/he can access directly to any of them through a simple index. In dialogic terms, this strategy can be named as "anaphoric group of topics", meaning by "anaphora" a generic mechanism to get back to previously visited topics during the dialogue.

Figure 3 shows an example of anaphoric group of topics, where the user can directly access the set of prints already visited.

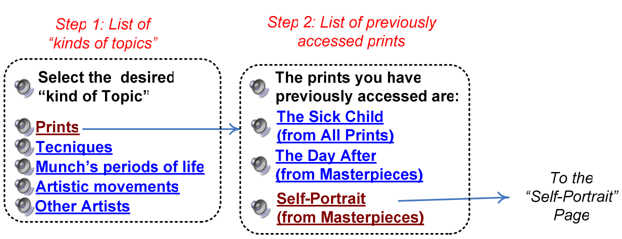

Figure 3. An example of anaphoric group of topics.

The anaphoric group of topics allows the user to reach a previously accessed topic with two navigational steps. The first step presents the list of all "kinds of topics" to the user allowing him/her to select one of them. The second step presents him a list of all the instances of topics s/he has visited (ordered by the time s/he visited each topic).

4.4 Designing "Upward" Navigation

With respect to the three scenarios of section 4.3, semantic back mechanisms address scenario "b" (to retrieve a previously visited content) only. We need to address, now, also scenarios "a" (orientation) and "c" (list of items).

Whereas the "topic history" allows accessing to previously visited content, a complementary strategy would be to offer the user the possibility of going back previously visited "lists".

Decoupling the history of the topics from the history of the lists enables to support more effectively two different scenarios. As we have discussed, the topic history answers the need of getting back to talk about an already visited content; a "list" history would allow the user to easily "go up" (see Figure 4) in the hierarchical structure of the information architecture level-by-level, passing through the visited lists of item (thus skipping the content pages). This strategy would be useful to recover orientation (being in control of groups of content) and, most of all, to select new items from the lists. In this case, no equivalent counterpart can be found in human dialogues, whereas lists of items are almost never explicit and put forward. However, the direct navigation between lists is of crucial importance to move around the information architecture. In fact, the typical sequence of action, for a standard web site, is to start from a list of items-links (index), select an item-link, and browse around. If another item, from the original list, is desired, the user has to click (one or more times, depending on the length of the browsing) on the back button, until the original list is recovered.

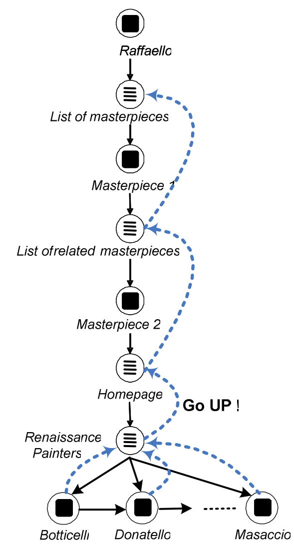

Figure 4. Going Up

At each list retrieved, the summary of the list will be provided to the user allowing him/her to either start listening to the list itself or go up again. In the last case, a meaningful way to partition the list could be among last visited and new items. Portioning the list in this way will allow the user to skip the items s/he is not interested in, reducing the overall access time to the list itself.

5. AURAL PAGES

In the above sections, we referred to the concept of "interaction (or navigational) step" to denote two different aspects in the interaction process between the user and the aural hypertext. The first one consisting in presenting the information to the user, the second one consisting in allowing him/her to make a choice. In standard website, each interaction step usually corresponds to a page, which is composed of a limited number of visual regions (or sections) where page elements are placed: the user can look at the content on the page and select links.

Page regions may be, for instance, a "content section", a "main navigation section", a "log in section", a "local navigation section", "service links section", "footer section", etc (see Figure 5). The semantic and the importance of each section are based on a **visualization strategy**. For example, the content section is usually placed in the middle of the page, the navigation bars at the top, service links are at the bottom, contextual links on the left side, and so on.

Analogously, in aural hypertexts, each interaction step will correspond to an "aural page". However, the characteristics of the vocal channel require a more detailed design: the designer should be able to communicate the same meta-information conveyed by the visualization strategy.

Figure 5. Example of visual page sections

Principles for an "aural presentation strategy" need to be defined. As for the visualization strategy, the aural presentation strategy aims at specifying which sections compose an aural page, which is their semantic and their importance within the page itself, how the user can access and interact with them.

5.1 Designing aural templates

Web designers know that using page templates helps keeping consistency across pages, thus facilitating the achievement of regularity in the user experience. A website – according to its size and complexity – may thus employ a few number of templates (4, 5 or 10), each corresponding to a specific page purpose (page for content, page for a list, page for a highlight, page for service subscription, and so on).

We argue that, as designers strive to create visual page templates, **aural page templates should be also conceived and offered to the user**. In order to do that, we introduce a model describing aural template (see the Aural Access Model in Figure 6).

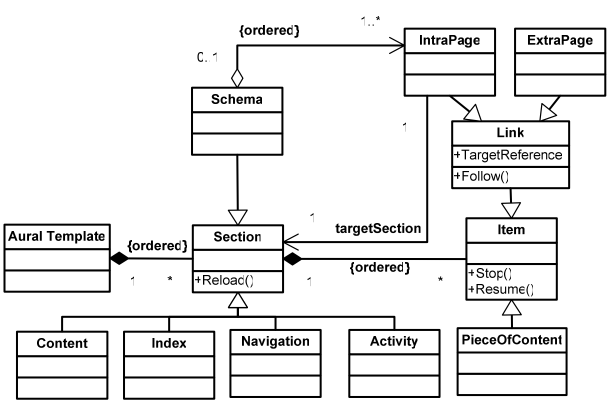

Figure 6. The Aural Access Model

In a nutshell, the above model provides a hierarchical description of aural templates: an "**Aural Template**" is composed of one or more "**Sections**" which, in their turn, can group a set of "**Items**". In its turn, an Item can be specialized into "**Link**" or "**Piece Of Content**".

Of course, an aural template is read linearly. Therefore, on the basis of the visual templates (i.e. capturing its semantics), designers are "forced" to reflect upon meaningful templates for the aural interaction. Sections should be defined, properly named and ordered for the aural channel, in such a way that the user is able to understand and recognize recurrent page structures across the site.

Note that different section types have different features and play different roles within the aural page. A proposal for classifying the aural page sections follows, but the "aural" designer can easily extend it:

- *Content section*: it contains the proper content that should be conveyed to the end user;
- *Index section*: it introduces a group of topics, enables to access the topics member of the groups and may explain the meaning of the group itself (i.e. the list of books belonging to the same category and the list of categories itself).
- *Navigation section*: it groups a set of overall links that performs the same function (i.e. the navigational toolbar in websites).
- *Activity section*: it allows the user to perform some operations (i.e. submitting a form).
- *Scheme Section*: it will be later introduced and explained.

With respect to the webpage shown in Figure 5, the corresponding aural template should be also composed of 5 sections (see Table 1). Using aural templates (as well as visual ones) brings a further positive effect on the user experience. As the user session evolves, the user learns the structure of the aural pages by understanding the templates and, over time, the user is able to recognize the purpose of the page before accessing its actual content.

Table 1. An example of Aural Template

Reading Order	Section Name	Section Type	Description
1	Book Description	Content Section	Actual information about the book
2	Section Links	Index Section	Links to the great offers
3	Login	Activity Section	To login into the private area
4	Landmark Links	Navigation Section	Links to the relevant areas of the website
5	Service Links	Navigation Section	Links to service information

Once defined a template, designers need to decide how they want the user to interact with such template, thus specifying an effective ***aural reading strategy***.

5.2 Defining the reading strategy

When looking at a website page (for the first time) – independently from the specific tasks s/he is trying to accomplish – the user usually performs the following actions:

1. Try to quickly scan the page to catch the overall template (what the page is about and which elements play which role).
2. Focus on the main message of the page (its content).
3. Focus on other elements to check their correspondence with what s/he is looking for (e.g. links to move to other pages).
4. (Eventually) make a selection (refocus on content and process it again and more in depth, or move to another page).

This is a simplified view of the complex visual processing of a page; however, even with iterations and variations, it shows that the user has a reading strategy to approach a page.

How to effectively support this flow of actions, as the visual channel is missing? The simple sequential reading of the elements of the page (currently supported by screen readers) fails to do that. As described above, in fact, reading a page cannot be reduced to simply reading every item (one by one) in a predefined order.

Following the logics of the visual reading strategy defined above, the following requirements should be supplied to offer an effective aural reading strategy:

1. **Allow the user quickly grasping how the page is organized by communicating its structure.** In order to do that, the *scheme section* is introduced into the template definition of aural pages. It allows the user to both easily identify the section composing the page and quickly access the content s/he is interested. For example, when accessing the aural page corresponding to the webpage shown in Figure 5 the schema section would say: "*Book: River Café Two Easy. This page is composed of 5 sections. You can either select one of them or wait while the first section starts. Sections are: Book Description, Navigation Toolbar, Service Links, …*".
2. **Read first the key message of the page, and then the other sections.** Presentation order of the sections composing the aural page should depend on their "semantic" importance within the page itself. Usually, the presentation order should start from the scheme section, followed by the main content, and then by the information most related to the content. The navigation toolbar and the service links should be presented at the end of the page (see Table 1).
3. **Define the access modality of each section.** It consists in specifying if the content belonging to a section has to be read either the first time the use will access the section, every time or never. Some possible access modalities are:
   - *Always*: the section is read every time the user accesses the page (i.e. for content sections).
   - *Only the first*: the section is read only the first time the user accesses the page. The next times, the reading of the section will require an explicit user's request (i.e. for the scheme section).
   - *On demand*: only the header of the section is read. The overall content of the section will be read only after an explicit user's request (i.e. for navigation sections).

   Offering different access modalities for sections allows speeding up the reading time of the aural page and implies a strong shift of paradigm: the user has not to choose which sections to skip but she explicitly selects which sections to read.
4. **Allow the user to access directly a section of interest at any time.** It consists in breaking up the linearity of the fruition imposed by the aural channel by defining different navigational paths among sections.

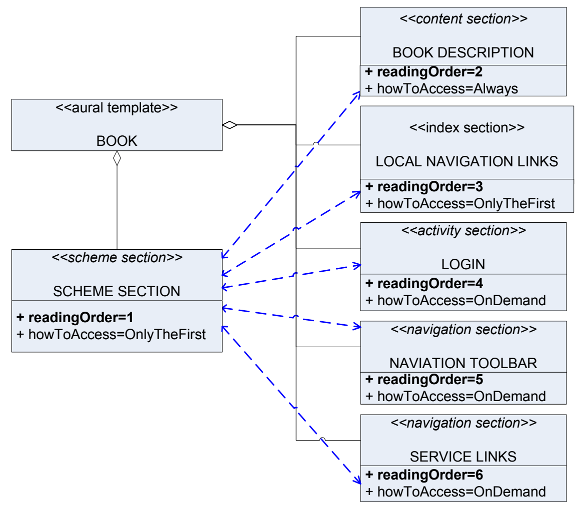

Figure 7. An example of Reading Strategy

Figure 7 shows a complete example of Reading Strategy. When accessing the aural page, the Schema Section is read first, communicating the structure of the page but also acting as an index for the other sections. Moreover, the Schema Section can be always directly accessed from the other sections. This will allow the user to access a section of interest at any time with two navigational steps: 1) from the current section to the schema section, 2) from the schema section to the section of interest.

6. CONCLUSION AND FUTURE WORK

We have analyzed important design issues to be considered when conceiving aural hypertexts. These issues have been discussed in the light of the requirements for an effective aural user experience and by illustrating samples of design solutions from real projects.

Most of the design solutions discussed have been implemented in a specific fashion for the website promoting the exhibition of Edvard Munchs prints at the State Museum in Berlin (www.munchundberlin.org). Although we did not perform formal evaluation or extensive testing, the website received extremely positive feedbacks from the visually-impaired users who visited it, in particular for its effective information architecture, specifically tailored to the aural fruition.

Table 2 summarized the design concerns discussed in this paper.

Table 2. Design Issues - Synoptic table

Design level	Design issue
AURAL INFORMATION ARCHITECTURE	Providing ways to understand the hypertext structure
	Designing "backward" navigation
	Designing long lists of items
AURAL PAGES	Designing aural templates
	Defining the reading strategy

Current and future research is focusing on the development and refinement of the design solutions proposed in this work, and by validating their effectiveness on a larger basis of empirical testing.

Furthermore, we are developing both a set of interaction primitives defining the "interaction language" between the user and the aural hypertext and a software infrastructure allowing deploying the aural hypertext itself. Moreover, thanks to the ORA project [27], a re-interpretation of existing design models and methods in an aural perspective is on going. The expected result is an extension of IDM (Interactive Dialogue Model) specifically aimed at supporting the design of aural hypertexts.

7. REFERENCES

[1] Ramakrishnan, I., Stent, A. and Yang, G. Hearsay: Enabling Audio Browsing in Hypertext Content. In *Proceedings of the 13th international conference on World Wide Web 2004*, ACM Press: New York, NY, USA p80-89.

[2] CSS2 Accessibility Features. Available: http://www.w3.org/TR/CSS-access, last access 2006-03-06.

[3] W3C. Web Content Accessibility Guidelines. Available: http://www.w3.org/TR/WAI-WEBCONTENT/, last access 2006-03-04.

[4] IT Accessibility & Workforce Division (ITAW). Available: http://www.section508.gov, last access 2006-03-06.

[5] Signore, O. and Leporini, B. Web accessibility: principles, international context and Italian regulations. In *Proceedings of CMG Italia, Conferenza annuale*, Pisa, 2004.

[6] W3C. Authoring Tools Accessibility Guidelines. Available: http://www.w3.org/TR/WAI-AUTOOLS/, last access 2006-03-06.

[7] W3C. User Agent Accessibility Guidelines. Available: http://www.w3.org/TR/WAI-USERAGENT/, last access 2006-03-07.

[8] Jaws for Windows, Freedom Scientific- Available: http://www.freedomscientific.com, last access 2006-02-27.

[9] De Witt, J. and Hakkinen, M. Surfing the Web with pwWebSpeak. In *Proceedings of Tech. and Persons with Disabilities Conference*. CSUN, 1998.

[10] James, F. Presenting HTML Structure in Audio: User Satisfaction with Audio Hypertext. In *Proceedings of the International Conference on Auditory Display (ICAD)*, pages. 97- 103, November 1997.

[11] W3C. VoiceXML. Available: http://www.w3.org/TR/voicexml20/, last access 2006-03-08.

[12] Speech Application Language Tags. Available: http://saltforum.org, last access 2006-03-02.

[13] *VoiceXML Forum*. Available: http://www.voicexml.org/, last access 2006-03-06.

[14] Li, H., andYuan, B. Research on voice browsing. In *Proceedings of IEEE TENCON*, 2002.

[15] Signer, B., Norrie, M. C., Geissbuehler, P. and D. Heiniger. Aural Interfaces to Databases based on VoiceXML. In *Proceedings of VDB6, 6th IFIP Workshop on Visual Database Systems*, Brisbane, Australia, May 2002.

[16] Kong, J. Browsing Web Through Audio. In *Proceedings of the 2004 IEEE Symposium on Visual Languages and Human Centric Computing*, 2004.

[17] *NetECHO*. Available: http://www.internetspeech.com/, last access 2006-03-06.

[18] Brown, M., Glinski, S. and Schmult, B., *Web page analysis or voice browsing*, First International Workshop on Web Document Analysis (WDA2001), Seattle, Washington, USA, September 8, 2001.

[19] W3C. *XHTML+Voice Profile 1.0*. Available: http://www.w3.org/TR/xhtml+voice/, last access 2006-03-08.

[20] SALT Forum. Available: http://www.saltforum.org/, last access 2006-03-08.

[21] Wang, K. SALT: a spoken language interface for Web based multimodal dialog systems. In *Proceedings of ICSLP-2002*, Denver, CO, 2002.

[22] Bolchini, D. and Paolini, P. Interactive Dialogue Model: a Design Technique for Multi-Channel Applications. *IEEE Transactions on Multimedia*, to appear on June 2006 issue.

[23] Capodieci, A., Di Blas, N., Paolini, P., Speroni, M. and Bolchini, D. Reading Web Pages to Visually Impaired People. In *Proceedings of EVA 2004*, Florence, 2004.

[24] Di Blas, N., Paolini, P. and Speroni, M. Usable Accessibility to the Web for blind Users. In *Proceedings of 8th ERCIM Workshop "User Interfaces for All"*, Vienna, Austria, 28-29 June 2004.

[25] Di Blas, N. and 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.

[26] Colazzo, S., Di Blas, N., Gobbetti, D. and Paolini, P. Accessible Web Sites Can be Modelled as Aural Dialogues. In *Proceedings of UAHCI 2005: 3rd International Conference on Universal Access in Human-Computer Interaction*. July 22-27, 2005, Las Vegas, Nevada, USA.

[27] The ORA Project. Available: http://ora.redirectme.net, last access 2006-03-15.