Identifying Adaptation Dimensions in Digital Talking Books
Carlos Duarte, Luís Carriço
Informatics Department, Faculty of Sciences, University of Lisbon
Campo Grande, Edifício C5
1749-016 Lisboa, Portugal
+351 21 7500{519, 603}
{cad, lmc}@di.fc.ul.pt

ABSTRACT
In this paper we identify adaptation enabling variables, and components that can be adapted, in a setting specific to Digital Talking Books. We then propose an evolution to our Digital Talking Book builder framework, to enable the production of adaptive books.

Categories and Subject Descriptors
H.5.2 [Information Interfaces and Presentation]: User Interfaces – Evaluation/methodology, Input devices and strategies, Voice I/O; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – Audio input/output; I.7.2 [Document and Text Processing]: Document Preparation – Multi/mixed media.

General Terms
Design, Human Factors.

Keywords

1. INTRODUCTION
Digital Talking Books (DTBs) are expected to be the next generation of information technology for the blind and print-disabled. By combining written text with audio narration, they represent a significant evolution from analog audiocassettes. DTBs' digital format, not only improves the reproduction quality, but also supports a number of interaction capabilities that would be cumbersome or even impossible to put to practice with audiotapes, like word searching and annotation entering. Besides the main target audience of blind and print-disabled population, other can benefit from a platform with these properties. The book’s multimodal interaction capacities broadens its usage scope to situations where visual attention is required elsewhere, either continuously (e.g. surveillance, driving), or temporarily (e.g. distraction).

We have developed an automatic DTB production platform [3], which is capable of flexibly generating different user interfaces for talking books. DTBs are built from digital copies of the source text and audio narration, through a two phase process. The first phase, which deals with content only aspects, aligns textual and audio content, standardizes it, and adds indexing information. The second phase deals with the interaction and presentation aspects, generating interfaces that can use visual and audio input and output [5].

It seems reasonable to argue that it would be very hard for a visually impaired or blind person to customize such a DTB interface to her preferences and possibilities. If we consider the multitude of possible configurations by combining different playback platforms, user characteristics, changing environmental conditions and book characteristics, we can imagine that even for a non visually impaired user the customization process might be a cumbersome task. When including in the framework the possibility of enriching the books with complementary, media such as audio clips and images [2], the user might be facing a complex task. All these reasons motivate us to further develop our framework, and enable the production of adaptive DTBs.

In the following section we present a brief overview of related work, covering DTB standards, categories and features, voice browsers and how they relate to DTB interfaces, and adaptive interfaces. This is followed by a discussion on adaptive DTBs, with considerations on what governs the adaptation, what can be adapted and we propose to evolve the DTB production platform. Finally we conclude and refer planned future work.

2. RELATED WORK
2.1 Digital Talking Books
DTBs are intended to allow easier access to books for the blind and print disabled communities. DTB standards have been developed by different organizations with the involvement of members of those communities. The Daisy Consortium and the NISO (National Information Standards Organization) have cooperated in the development of the most important DTB specification, the ANSI/NISO z39.86 [1].

The standards development led to a possible categorization of DTBs according to the functionalities they make available, which also reflect their inherent complexity: full audio with title only; full audio and navigation control; full audio, navigation control and partial text; full audio and full text; full text and partial audio;
and full text and no audio. The features that a DTB should support include: basic navigation capabilities, fast forward and reverse, reading at variable speeds, navigation through table of contents of a navigation control file, reading notes, cross-reference access, bookmarks, searching and others.

The standard, wisely, does not propose specific interaction solutions. However, the problems originating from such navigation features, along with the non-visual nature of the environment require complex solutions.

2.2 Speech Interfaces
The voice browser interface shares with the DTB interface some common problems: (1) the audio format is a temporal medium, with the capability of presenting only one word at a time, while visually a page can render images, tables and text simultaneously; (2) issuing voice commands and processing audio output consumes short-term and working memories, conflicting with planning and problem solving tasks. Visual information is processed by separate cognitive systems [4]; and (3) the unavoidable recognition errors.

However, multimodal system’s research has made clear that speech input is advantageous in several circumstances, identifying that “the best tasks for speech input were tasks in which the user has to issue brief commands using a small vocabulary” [4][9]. This means that the characteristics of DTBs are advantageous for a speech interface. Some of the recommendations available to the construction of voice browsers can be adopted to DTB design: links should be easily spoken text; links should be short (a few words); avoid links with similar sounds; and develop alternatives to numbered link, as these cause cognitive overload.

2.3 Adaptive Interfaces
An adaptive interface has been defined as “a software artifact that improves its ability to interact with a user by constructing a user model based on partial experience with that user” [6]. However, we may argue that relying only on the user model, as the basis for adaptation may be insufficient. There are situations where the drive for adaptation originates from external events and environmental changes.

Adaptive interface systems cover many areas, with educational applications and on-line information systems being the most popular. We can relate the creation of adaptive DTBs to both of these areas, very clearly when producing a DTB with educational purposes, but also from the information systems area when considering how to adapt based on location and behavior in physical spaces [7][8].

3. ADAPTIVE DTB
Evaluation [3][5] of DTB interfaces generated by our framework has produced commentaries from the test subjects, showing the desire for a personalized version of the interface. Some examples follow: “I would like the start of a new chapter to be displayed in the text”, “I would like a sound to signal the start of a new chapter”, “I would like to have the full content of the annotation shown by default”, “The annotation frame could be replaced by a link that when followed would show an annotations window” and “The annotation frame could be used to show images related to the text”. Each of the observations was made by a different subject.

If we consider the blind and print disabled population, which is the primary target audience of DTBs, we can expect users with very different characteristics than the ones of the tests subjects. The identification of the aspects to customize, and the customization process, should be a harder task for the target audience, thus being an extra motivation for the development of an adaptive interface.

Besides the mentioned “personal preferences”, other aspects should be contemplated when considering the creation of an adaptive book. General characteristics include the capabilities of the playback machine (which input/output devices are available), the playback environment (noisy backgrounds), the user’s current activities and others that can be found in common adaptive applications.

Taking into account all these variables, and their relevance during the processes of creation and playback of the book, we identified some of the elements of the presentation of the book that can be adapted, as well as some of the variables responsible for initiating the adaptation procedure.

The adaptation initiating variables can be divided into two groups: user related and environment related. Examples of user related variables are: characteristics, knowledge, preferences, interaction history and current activity. For environment related variables we have: interaction devices available, access to media repository, background noise.

The components of the book that can be adapted may be separated into three dimensions: interaction, content and presentation. Examples of interaction components are: the input/output modalities available that can be enabled or disabled, and used in cooperation or individually. Content characteristics that may be adapted are: enhancement of the book with the introduction of sound, images, and other available media, translation of the text, and hiding or revealing of parts of the book. For presentation components, examples are: size and color of the font used, rearranging of the elements of the book on the screen, type of audio signals used, and synchronization units.

We can also examine DTB related aspects of the articulation between some of these variables and components.

- The visual impairment level of the user has a clear impact on the presentation of the book. From the size of the font used, to not using visual presentation at all for blind users.
- The user preferences, capacities and past interactions influence the presentation aspects of the book, by setting the level of enhancement used.
- The interaction history can impact the synchronization units. If there is a history of jumping to the same (or near) word in the book, the synchronization units should become gradually finer. In this way, the first jump would lead to the word’s paragraph, then to the word’s sentence, and finally to the word itself.
- The output media used can impact the synchronization units. Absence of a visual display leads to the use of less detailed synchronization units. After searching a word in a sound
only output environment, a jump to the searched word would make it more difficult for the reader to understand the “new” context, when compared with a multimodal environment, in which the narration jump can be accompanied by the visual presentation of the surrounding text.

The book characteristics will also play a part in the adaptation process. For example, when processing a novel, it would not make much sense to hide parts of the text, but when processing an educational book it could make sense to hide some of the content already known by the reader.

Considering all the variables and dimensions of adaptation, an extended user model would be needed to allow an adaptation not only to the user characteristics and preferences, but also to the environment conditions.

In order to produce adaptive versions of the DTBs we propose to evolve the DTB production framework. The adaptation dimensions are related to adaptation modules that are responsible for introducing the adaptation capabilities in the generated DTB. For example, if we wish to produce a DTB that arranges the placement of the components on screen, we have to include the corresponding module in the production process.

This way we are able to create different DTB versions, by including more or less adaptation modules. This will also allow the production of books with a set of pre-determined capabilities, by having a pre-selected number of adaptation modules, or more general books, which have all the possible adaptation capabilities. This can be represented by an adaptation slider, which would control the adaptation setting of the book.

The slider set to one of the extremes will create a book customized to a class of users and/or environment characteristics, as if using a stereotype user model, by selecting the appropriate adaptation modules from the framework. This will allow the creation of several customized versions of the same book, but without all the adaptation capabilities. This is suited for situations where there has been a previous identification of environmental constraints or user characteristics.

The slider set to the other extreme would create a general book, but with greater ability to adapt to its reader, similar to the use of an overlay model. This would create only one version of the book, but a version capable of adapting (or being adapted). This is appropriate for situations where there is no prior information about the playback device, environment and reader group.

Settings in between the two extremes would represent a transition between the stereotype and overlay user models, allowing for the creation of books where some of the components would support adaptation while others would not.

4. CONCLUSIONS AND FUTURE WORK
In this paper we have proposed an evolution to our DTB production framework. The framework currently produces DTBs from digital copies of the text and audio narration, which can be presented in a standard web browser. The proposed evolution provides adaptation capabilities to the generated DTBs.

We have identified the variables governing the adaptation, as well as the interface components that can be adapted. Several ways of how the adaptation variables and components are related have been presented. We envision the creation of books tailored to a predetermined group of users and usage situations, and books created for a general group, but with greater adaptation capabilities.

Currently we have a customizable (but not adaptive) module for enhancing the book. This will allow us to conduct a Wizard of Oz experiment, to do a first evaluation of the proposed adaptation capabilities of the DTB.

5. REFERENCES