Multimodal Enhancements and Distribution of DAISY-Books

Wolfram Eberius	Alexander Haffner
Dynamic Designer	German National Library
Schönbrunnstrasse 1	Adickesallee 1
01097 Dresden	60322 Frankfurt am Main
Germany	Germany

Abstract

An increasing demand for providing accessible content to widely differentiated target groups calls for a universal eBook standard combined with a flexible user interface. For this purpose, we enhanced the current DAISY standard to serve the blind and visually handicapped as well as deaf, hearing impaired, dyslexic and elderly users. In addition to supporting sign language videos and alternatives in simplified language, we introduced videos with media enrichments as a primary source. Additionally a personalizable web player, based on the Adobe Flex Framework, allows real-time streaming and local playback of original and newly developed multimedia DAISY books. Furthermore, the player supports multimodal input such as gesture and speech recognition to facilitate use by physically handicapped people.

1 Introduction

As the world wide eBook standard for people with print disabilities, DAISY offers wide access to print media in digital form; in particular, DAISY's rich text navigation techniques allow for a flexible way of reading. Currently, DAISY books are mostly distributed on CD's or via download. The accessibility of DAISY books finally depends on accessible software and/or hardware players. The DAISY standard currently supports visually impaired users, but unfortunately no users with other disabilities. There is no comparable standard for users with other disabilities that offers additional modalities such as text synchronized sign language videos. If there were additional standards for each specific group of disabled users, redundancies would emerge in the production and in the delivery processes since the book's text, the metadata and the structure information would be the same.

The concept of Design-for-all calls for systems and applications that are suitable for most of the potential users without any modifications (Bühler&Stephanidis 2004). Our goal was to show that an extended and therefore more universal DAISY standard can satisfy the needs of highly differentiated target groups. To fulfil this goal we have enhanced the DAISY standard and developed a highly flexible player that can be adapted to different user needs. Combining standard and player, we fulfil the requirements of Design-for-All. Furthermore, text currently acts as the base for production of DAISY books, but we believe that it is possible that video finds usage as a primary source as well, given that videos can be made accessible by enriching them with subtitles and captions. We have also

developed a solution for real-time streaming of DAISY books, allowing the synchronised streaming of all included media within DAISY book online delivery.

2 Related work

2.1 Sign language

Reading print information is complicated for hearing impaired users since their first language is a sign language (Fisher&Petrie&Weber 2005). Therefore, it is required to include sign language in DAISY in order to fulfil the needs of Design-for-all. We introduce two related projects.

2.1.1 Link-It

The National Agency for Special Needs Education and Schools in Sweden developed *Link-It*, which includes a format, an authoring tool and a player that synchronizes text and sign language videos (SPSM 2009). Integration of sign language videos in DAISY is desired for future DAISY standards (DAISY 2009).

2.1.2 The eSiGN project / SiGML

The international eSIGN project introduced SiGML, an XML based markup language for sign gestures (E-SIGN 2009). SiGML can be rendered by Java based browser plugins via selectable avatars. Delivering sign language through SiGML lowers costs for production and network traffic in comparison to natural videos.

2.2 Minority Cube

The "Minority Cube" project demonstrates the possibilities of controlling a web application with gestures (Klingemann 2009). A camera connected to the computer captures the user's movements and detects two dimensional gestures to control a cube displayed on the screen. Gesture control can increase the accessibility of a user interface and so support users with physical and linguistic handicaps.

3 Multimodal Enhancements of DAISY-Books

The extension of the DAISY/NISO 2005 Standard was made by integrating new elements and attributes in the given XML syntax. Those elements and attributes were taken from the latest SMIL 3.0 DAISY Profile (W3C 2008).

3.1 Videos as Primary Source

The *<video>* element of the SMIL 3.0 DAISY Profile has been integrated. Videos can be enriched by differentiated captions and additional audio tracks. In order to support captions the *<textstream>* element has also been integrated. Videos can be used in SMIL and NCX files. In order to distinguish audio tracks for different purposes, the new test attributes listed below in Table 1 have been integrated.

Test attributes for audio tracks	Description
systemOverDubDialogue	Audio track for dialogues
systemOverDubSound	Audio track for sound and ambience
systemOverDubMusic	Audio track for music
systemOverDubNarration	Audio track for narration and audio descriptions

Table 1: Test attributes for audio tracks

To reduce authoring efforts, the original audio track of a movie can remain within the movie file but has to be marked up separately as *<audio>* element with the appropriate movie file as source and a *systemLanguage* attribute so that audio tracks in other languages can be used as well. To distinguish captions, the new test attributes listed below in Table 2 have been integrated.

Test attributes for captions	Description
systemCaptionsDialogue	Captions for dialogues
systemCaptionsDialogueExtended	Captions for dialogues with additional rendering information, e.g. text colour & text position
systemCaptionsSoundMusic	Captions for sound and music
systemCaptionsNarration	Captions for narration

Table 2: Test attributes for captions

Below, Listing 1 outlines an example for including video in DAISY.

```
<par>
 <!-- video -->
  <video src="movie.mpg" />
 <!-- audio -->
  <audio src="movie.mpg" systemLanguage="en" />
  <audio src="movieGerman.mp3" systemLanguage="ger" />
 <!—additional audio tracks -->
  <switch systemOverdubDialogue="On">
   <audio src="movieDialoguesEnglish.mp3" systemLanguage="en" />
   <audio src="movieDialoguesGerman.mp3" systemLanguage="ger" />
  </switch>
  <switch systemOverdubNarration="On">
   <audio src="movieNarrationEng.mp3" systemLanguage="en" />
   <audio src="movieNarrationGer.mp3" systemLanguage="ger" />
  </switch>
 <!-- captions -->
  <switch systemCaptionsDialogue="On">
   <textstream src="movieDialogueEn.tt" systemLanguage="en" />
   <textstream src="movieDialogueGer.tt" systemLanguage="ger" />
  </switch>
  <switch systemCaptionsNarration="On">
   <textstream src="movieNarrationEn.tt" systemLanguage="en" />
   <textstream src="movieNarrationGer.tt" systemLanguage="ger" />
  </switch>
</par>
```

Listing 1: Video in DAISY

3.2 Sign language videos and SiGML

Natural sign language videos are included analogously to videos in 3.1. SiGML gets included with the *<text>* element, the new test attributes listed below in Table 3 have been integrated. Sign language videos and SiGML can be used in SMIL and NCX files. Listing 1 outlines an example for including video in DAISY.

Test attributes for sign language	Description
systemSignLanguageVideo	Natural sign language video
systemSignLanguageSiGML	SiGML mark up
systemSignLanguageDialect	Sign language dialect

Table 3: Test attributes for Sign language

```
<par id="chapter1">
  <text src="dtbook.xml#dtb1" />
  <video src="signLanguage.mpg" clipBegin="0:00:00" clipEnd="0:00:07"
      systemSingLanguageVideo="On" />
  <text src="signLanguage.sigml#dtb1" systemSignLanguageSiGML="On" />
</par>
```

Listing 2: Sign language videos and SiGML in DAISY

3.3 Simplified language

Simplified language supports people with cognitive disorders, such as dyslexic users, as well as children or the elderly. Content in simplified language can be delivered through audio and text and can be used in SMIL and NCX files. We therefore need a new test attribute for simplified language integration (Table 4). Below in Listing 3 an example for including simplified language in DAISY is found.

Test attribute for simplified language	Description
systemSimpleLanguage	Simplified language

Table 4: Test attribute for simplified language

```
<par id="chapter1">
  <text src="dtbook.xml#dtb1"/>
  <audio src="audio.mp3" clipBegin="0:00:00" clipEnd="0:00:04" />
  <switch systemSimpleLanguage="On">
    <text src="dtbookSimple.xml#dtb1"/>
    <audio src="audioSimple.mp3" clipBegin="0:00:00" clipEnd="0:00:11" />
  </switch>
</par>
```

Listing 3: Simplified language in DAISY

3.4 General notes

A DAISY book can now as well reference audio and video material from external sources.
Concerning names for attributes, the author proposes a general name for test attributes that are only used for enabling/disabling. This attribute accepts one or many of an enumeration of allowed values (e.g. test="systemOverDubMusic" instead of systemOverDubMusic="On").

4 The DAISY Web Player

The DAISY Web Player is a feature rich DAISY Player built on the Adobe Flex Framework. The Player runs in any environment possessing a browser with the Adobe Flash Player Plugin and is accessible to assistive technologies such as Jaws screen reader and was successfully tested with visually impaired and blind users. The developed DAISY Web Player features include:

- support for DAISY 3 Books
- support for DAISY Videos, sign language, simplified language , SiGML
- integrated self-voicing (for environments without assistive technologies), self-signing
- real-time streaming via RTMP (Adobe Flash Media Server, Red 5)
- synchronization of multiple streams
- highly adaptable user interface
- customization by user profiles
- multilingualism
- short cuts
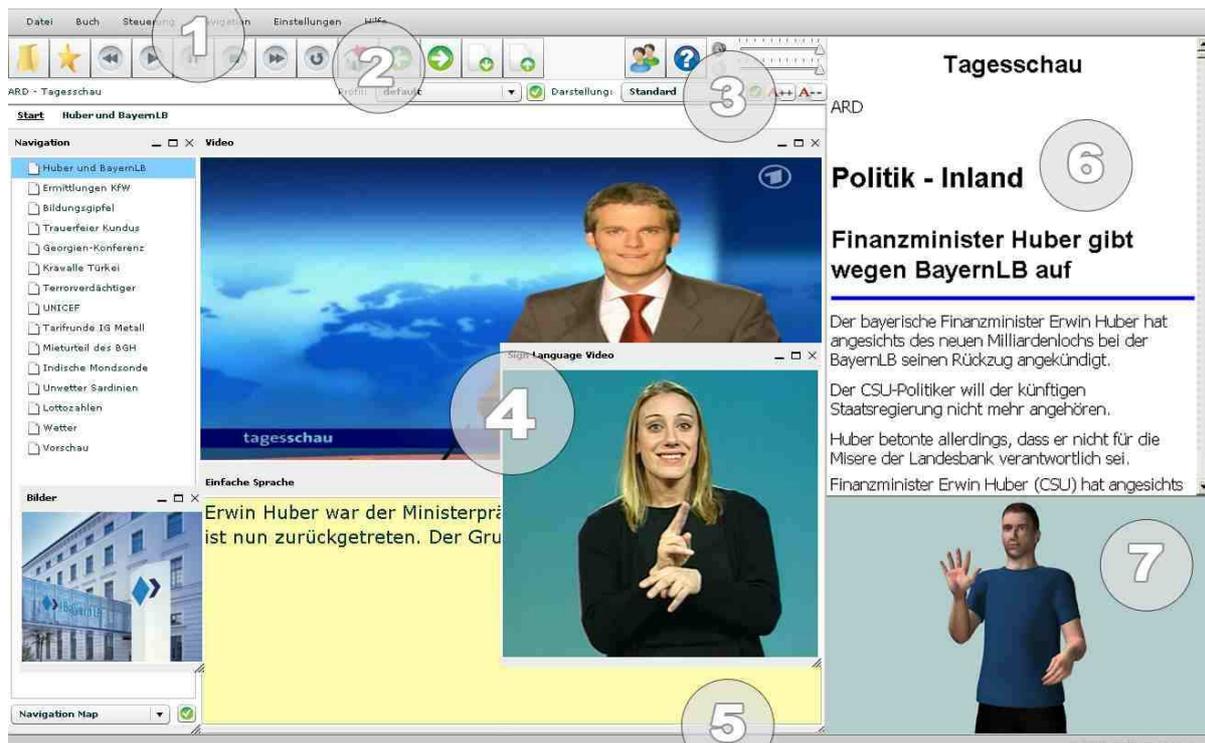- alternate navigation concepts: breadcrumbs and reading history

Figure 1: Screenshot of the DAISY Web Player

For demonstration purposes the screenshot in Figure 1 displays all possible widgets. The interface consists of the main menu [1], button bar [2], settings bar [3], main stage [4], status line [5], text [6, within an HTML environment] and the SiGML avatar [7, as Java Plugin within an HTML environment].

Feedback and current developments

The DAISY Web Player was awarded with the AKEP Junior Award 2009. AKEP is the working group for digital publishing at the German association of book publishing houses and sellers. The current and future development of DAISY Web Player is continued by the BGB Company *Dynamic Designer* in Dresden, Germany (Dynamic Designer 2009).

5 Conclusions and Outlook

The DAISY Web Player and the enhancements of the DAISY standard are to be seen as a proof of concept for DAISY being able to support additional target groups besides blind and visually impaired users. DAISY has the potential to become a universal standard for accessible eBooks in the context of Design-for-all. Newly supported target groups are deaf and hearing impaired users as well as people with cognitive disorders, along with mainstream users including children and the elderly.
An accessible web-based player was built that supports streaming, synchronization and output of multiple media streams as well as multimodal input. The next major goals for the player are speech and gesture control and the support for DAISY 4 as soon as it is published. The DAISY Web Player project is supported by the DAISY Consortium.

References

Bühler, Christian, & Stephanidis, Constantine (2004): European Co-operation Activities Promoting Design for All in Information Society Technologies. In Proceedings of the 9th International Conference on Computers Helping People with Special Needs (ICCHP 2004), Paris, France, 7-9 July, 80-87.

Fisher, Wendy & Petrie, Helen & Weber, Gerhard (2005): Personalization, interaction, and navigation in rich multimedia documents for print-disabled users. In IBM Systems Journal archive, Volume 44, Issue 3, 629 - 635.

SPSM (2009): http://www.spsm.se/Startpage/Educational-materials/Deaf-and-Hard-of-Hearing/Products/Link-it/Link-it-Use (Page view: 28/08/2009)

DAISY (2009): http://www.daisy.org/z3986/requirements/detail.php?i=81 (Page view: 28/08/2009)

E-SIGN (2009): http://www.visicast.cmp.uea.ac.uk/eSIGN
(Page view: 28/08/2009)

Klingemann, Mario (2009):
http://incubator.quasimondo.com/flash/minority_cube.php
(Page view: 28/08/2009)

W3C (2008): SMIL 3.0 DAISY Profile. http://www.w3.org/TR/2008/CR-SMIL3-20080115/smil-daisy-profile.html

Eberius, Wolfram (2008): Multimodale Erweiterung und Distribution von Digital Talking Books. Diploma thesis, Technical University of Dresden, Germany.

Dynamic Designer (2009): http://daisy.dynamic-designer.de
(Page view: 28/08/2009)