DAISY, the Best Way to Author Sign Language Publications.

Guillaume Olivrin Golivrin@CSIR.co.za
and Willem Van Der Walt WvdWalt@CSIR.co.za
CSIR Meraka Institute
Meiring Naude Road, Brummeria
0184 Pretoria
South Africa

Abstract

There is no standard electronic publication format for sign languages besides ad hoc video media such as the DVD. Sign language requires the structuring and navigational capabilities of DAISY, adapted to both video modalities and visual languages. This paper investigates how video playback can be introduced in DAISY by building on the DAISY 3 standard and on existing open-source software. The paper presents the process of creating and authoring a video-based sign language DAISY publication to serve as test cases. Additionally, the Ability Based Technical Intervention prototype demonstrates how content can be further adapted to users' abilities, learning style and literacy levels to make all-inclusive video-based publications.

1 Introduction and Background

The user needs and functional requirements for a video-based sign language DAISY publication are first captured. Secondly, a selection of existing open-source DAISY players are reviewed that provide a foundation on which to build sign language and video support. And thirdly, the creation process for sign language content and the various technical solutions for authoring structured video content is explored.

1.1 Daisy Books in Sign Languages

According to a European survey (Pyfers 1998, 6), the creation process for sign language books, from producers to end-users, lacks best practices, standards and international cooperation. The situation has not improved much since then and Deaf associations still express their need and willingness to author digital books, and even sign language DAISY publications for the Deaf, as reported in "Daisy for All 2003-2008" (Maharaj 2008 / Kawamura 2008, 28). Internationally, there is a growing need to archive sign language digital records : digital heritage, teaching material, research and documentation of sign languages, inclusive public information on AIDS and disaster preparedness.

The DAISY publication standard can be extended to provide such a unifying creative framework, one which would become a dedicated means of expression for sign language users across the world. The objective of DAISY For All is to provide an inclusive publishing standard with technical functionalities and best practices that enhances usability and accessibility (Maharaj 2008, 21).

The functional requirements for the publication of sign language as videos are:

- Meta-data: to determine whether a video does feature sign language and which sign language it is (according to the ISO 639-3 language codes, e.g. 'sfs' for South African Sign Language and 'fsl' for French Sign Language.
- Structure: to provide intuitive sign language-based organisation of a document or book. This structure includes time and space based hyper-linking for cross-referencing within and between videos, and to provide concordance to external resources such as text, timed-text or captions.
- Distribution Best Practices: to make video content easier to exchange and adapt.
- Accessibility: to cater for varying fluency in sign language (for example the playback speed), various levels of literacy (captioning and text simplification) and supplementary communication strategies (for example lip reading, or cued English).

1.2   A Review of Open-Source DAISY players

Table 1 below provides a comparison of currently available open-source Daisy players under Linux. A transcript of the table is available online (VanDerWalt, 2009).

Player	Amis	Daisy-delite	DBR	Emerson	Idair	Listen-up
Author	DAISY Team	Samuel Ytterbrink	Francisco Martnez	Marcus Gylling	Johan Knol	K. Reiser, J. Zhang
Latest Release	05/05/2009	15/02/2008	01/10/2008	29/05/2009	18/12/2003	17/09/2003
Version	3.0	0.1.3b	1.1.2	0.6.2	0.8.3	1
Maintained	Yes	No	Partial	Yes	No	No
Implementation	C++	Python	Python	Java	C++	C
Libraries	Ambulant, SAPI, DirectX	CoreAudio, PyMad	GTK, Gstreamer	Javax.sound	LibMad, TCL-TK	osalp, Ncurse,libxml2
Endorsed	Yes	No	No	Partial	No	No
DAISY 2.02	Yes	Partial	Yes	Yes	Yes	Partial
DAISY 3	Yes	No	No	Yes	No	No
Full-text / Full-audio	Yes	No	No	Partial	Yes	Yes
Audio-only	Yes	Yes	Yes	Yes	Yes	Yes
Text-only	Yes	No	No	No	Yes	Yes
Interface	Self-voicing, DirectX	Text-based, GUI	GTK	Eclipse SWT	Text-based, TCL/TK	Text-based Only
Portability	Difficult[3]	Good[4]	Good	Good	Good	Possible[5]
Compilation	Difficult[6]	Easy	Easy	Difficult[8]	Difficult[9]	Difficult
Linux	No	Yes	Yes	Yes	Yes	Yes
Audio formats	WAV, MP3	MP3 only	MP3 only	MP3	MP3, WAV	WAV, MP3
Robust Parsing	Yes	Partial[7]	No	Yes	Yes	No
Navigation	Good	Partial	Partial	Partial[7]	Good	Good
License	LGPL	GPL	GPL	EPL	GPL	GPL

Table 1: Comparison of open-source DAISY players under Linux.

Captions : 1. AmisCore and Linux (Van Der Walt and DeMeglio 2008). 2. Text-based interface could instantiate separate graphical interface to render videos. 3. Core lib from Windows OS. 4. Core libs from Mac OS. 5. Audio from Linux OS. 6. Ambulant. 7. Bug in parsing/navigating levels. 8. Eclipse Plugins. 9. Old libraries requires legacy compiler.

The open-source route was adopted to select a candidate DAISY player because of the inherent freedoms it offers for running, studying, redistributing and

improving a programme (Stallman 2008). In order to validate our method of creating a new kind of daisy book —a sign language DAISY book— a practical implementation of the book and a player to play it are required. The choice of player was not based on the player being feature-rich or bug-free, but rather on how easy the player is to modify to fulfil the new requirements.

After looking at various players, DBR 1.01 (Martnez 2008) was chosen for the project for the following factors:

- it is written in Python which is ideal for prototyping.
- it uses Gstreamer, a library which is well-suited for rendering audio and video,
- it runs under Linux, a platform with inadequate Daisy support.

The way DBR has been modified for sign language playback has resulted in the desirable side-effect of playing a wider range of audio formats, hence providing better support for existing DAISY books.

Emerson (Gylling 2009) was the second best candidate for implementing a sign language DAISY player. Precisely which new audio-video engine library to integrate in Emerson code wasn't immediately obvious : the Java media framework might not have had suitable media decoding support whereas the GStreamer-Java bindings were not yet well-documented.

In order to test the new player's implementation (including specific video and sign language requirements) a compliant full-text, full-audio DAISY book was created and then extended to allow for the integration of video and sign language.

2 Sign Language DAISY Publication

Video introduces new possibilities in DAISY publications at the cost of increased complexity in the standard and in its players' implementations. A simple approach to making video DAISY books is presented, which requires minimal changes in DAISY players and which is within the standard. Implementation issues and content adaptation needs are further discussed that will influence the design of future DAISY standards.

2.1 Creation of Sign Language Content

To create a full-text/full-audio and full-text/full-video DAISY test-book, the original content of "Laws of the Game 2008/2009" (FIFA, 2008) served as a starting point:

1. A full-text DAISY 3 book was first authored and validated,
2. The latest Obi (1.1) was then used to record and export a full-audio DAISY 2.02 book. The full-text was lost but then reinserted within the DAISY 2.02 book. The generated audio files were injected within the DAISY 3 book,
3. South African Sign Language was filmed sentence by sentence to match the full-text book,
4. New versions of DAISY 3 and 2.02 books were finally created by substituting video references in place of audio resources.

The reason why the original test-book was retrograded to DAISY 2.02 was to make it easy to record audio files using Obi (under Windows) and to make a first book instance that worked with DBR. For experimental purposes it is useful to have a set of test-books that cover the various technical implementations of video within the DAISY 3 and upcoming DAISY 4 standards.

Our test-book includes three chapters which are full-text, full-audio and full-video. Each sign language video clip corresponds to one sentence in the original text-book, making it easy to cross-reference text and video for concordance at sentence (<sent>) level.

2.2 Authoring the Book and Implementing Playback

In order to make minimal changes to both the book standard and the player implementation, <audio> tags were reused to reference video resources. The ability to pass either video or audio resources without discrimination is made possible by the playback library Gstreamer. Gstreamer's 'playbin' can automatically select the appropriate playback mechanism for any video and audio resource. For a better handling of DAISY books however, it is advisable to extend the standard with a dedicated <video> tag which would notify future DAISY player implementations that it needs both video decoding and a graphical canvas to display the video.

Once DBR's Gstreamer decoding pipeline had been modified, it was able to accept any audio format as well as video. In DBR's original architecture, the rendering engine is initiated only once, before the graphical user interface. In order to handle video, the graphical interface had to be instantiated before the rendering engine to pass a handle to the video rendering canvas. It is however preferable to fire up new instances of video canvas together with dedicated decoding pipelines as required at the opening of the book or 'just in time' at playback time. This last technique proved useful in the AbTi prototype where multiple videos and audio resources need concurrent rendering but have varying start and end timings.

We can report that the modified video DBR can render a full-video DAISY 2.02 sign language book with no sound (unless the soundtrack is within the video stream) and no text. Here is a list of further technical considerations for the addition of video in DAISY publications:
- Implementing clip-begin="npt=4s" clip-end="npt=5s" for videos,
- Audio/Video-Text concordance : linking forward and back using IDs in SMIL and in the DTBook, a requirement discussed in (Larsson 2009)
- One SMIL file for audio and video or separate SMIL files,
- SMIL 1, 2 and 3 DAISY Profiles,
- Extending DAISY with video and sign language modules,
- Extending SMIL 3.0 DAISY Profile by adding modules, having skip-content attributes and new XML namespaces such as SignLanguage and Video,
- Managing concurrent requests for playback mechanisms,
- Hard-coded player graphical interface versus book-instance based graphical interface (which depends on the SMIL layout declarations and on the video and text resources requested),
- Navigational interfaces architecture which provides full accessibility support for the DAISY player text-only interface, audio-only interface and new video and sign language based graphical user interfaces.

2.3 Digital Signing Book and Adaptive Content

The addition of video multiplies the number of possible combinations of modalities available in DAISY. It becomes necessary to stipulate the roles and relations of video with other resources:
- Does the video contain audio and text streams?
- Is the video an illustration or example for the text?
- Is the video an alternative representation of the text?
- Is the video another language altogether, i.e. sign language?
- Is the video captioned and if so, what is the relation between the timed text and the full-text in the body of the DAISY book?

In his online DAISY player implementation, Wolfram Eberius (Eberius 2008) makes use of SMIL system tests to discern a generic video from a sign language video.

To give a concrete case of sign language DAISY publication, such a publication would be typically produced as a full-video DAIY „book" where sign language is the original and primary content. This is similar to full-audio DAISY books which are created first hand without necessarily a text-book as premises. In our prototype test-book, sign language was introduced in the DAISY publication as a derivative of and an alternative to the original text. This however will not be the case in real-world sign language production and publication scenario. In a typical filming setup, sign language video will come first, followed by optional text-subtitling and audio captionning.
With DAISY, it becomes possible to have both videos and timed-text captions in the SMIL profile. The place of the postproduction edited text would then be in the DTB publication body. All the referencing and identification mechanisms in XML can be used to link, cross-reference and synchronize the various media.

Another particularly important aspect of sign language specific content is the integration of sign language animation and notations' scripts within DAISY. Sign language animation scripts and notations make it possible to create a Digital Signing Book (or sign language publication) with a signing avatar, much like it is possible to render a full-text DAISY book with a speech synthesizer. The inclusion of these scripts and notations in DAISY can be modular as it is already the case for math notation and SVG scripts in DAISY. These notations are essential to provide a search functionality for gestures and sign occurance within a book. This functionality is similar to searching text in a full-text/full-audio DAISY book in order to locate the corresponding speech in an audio clip. However sign language, unlike audio, is not appropriately encoded by text transcripts (i.e. Gloss) and is better described using notations for gestures.
Since most gesture animation scripts and sign language notation systems have an XML representation, DAISY is the best standard candidate to realise the integration of sign languages in a common publication format.
Another prototype, the Ability Based Technical Intervention (AbTi) project (Coetzee 2009 and Olivrin 2008) provides a comparable framework for storing and presenting content according to user needs. Amongst users of the AbTi system are the Deaf who communicate in South African Sign Language and those who use other communication strategies such as Simple English, Symbol Writing and captioning. This project demonstrated that the various semantic roles of content and the relations between contents, need to be explicitly defined to provide more inclusive and adaptive playback mechanisms for users. AbTi investigated how content can be adapted and presented differently for users with

varying abilities, learning styles and literacies. The AbTi „Integrator" is very much like an experimental DAISY player whose playback mechanism is implemented using a finite state machine, Python and GStreamer. The role of content is clearly defined according to 8 categories (introduction, summary, theory, examples, questions, exercises, activities, links) which can then be mixed to provide the best learning experience for each user-specific abilities and preferences. There are many content derivatives in an AbTi content package. Text, for instance, can assume many forms: translated, simplified (Simple English), English in sign language syntactic order, timed-text such as captions and even pictographic-text which includes image-symbols to replace word-concepts.

Similar adaptive functions for contents can be achieved within DAISY using existing navigations functionalities (IDs, smilrefs, Spine, Guide, Tour). These various navigation strategies for DAISY could be generated automatically by a user profiling tool such as AbTi (Coetzee 2009 and Olivrin 2008).

3 Conclusion and Future Work

The DAISY standard already includes all the prerequisites to handle visual and timed media with SMIL. Further modular extension of DAISY using existing sign language and gestural markup is only a formality. This addition greatly extend the impact of DAISY publications to those who are word-illiterate and to those who communicate in sign language and for whom there is no digital publication standard.

The prototype sign language DAISY publication, the modified open-source DAISY player and the experimental AbTi prototype presented in this paper show that video playback with DAISY is accessible and inclusive. The challenge of designing graphical user-interfaces and rendering pipelines for multiple audio and video streams will benefit current DAISY player implementations by making them more robust:

- DAISY 3 already allows for a wide range of media and codec which very few players support. Therefore existing players need to have their rendering pipeline revised to provide wider and more robust support for audio and eventually, video.
- Current DAISY player architectures are lacking in the design of multimodal interfaces. A more generic and inclusive design for a DAISY player must allow for command-line text-only DAISY player interfaces as well as prompt-based audio-interface, images and video in graphic user interfaces and combinations of all these modalities. User interfaces must not necessarily be hard-coded to be deterministic: adaptive user interfaces are instantiated just-in-time or according to a DAISY book's directives.

From the experience of building these prototypes we uncovered some of the requirements and design choices that are needed in the DAISY next generation standard (DAISY 2009a and DAISY 2009b). Robust authoring software for new DAISY standards, including DAISY 3 new modular extensions, is clearly a necessity to maintain best practices and ease of creation.

With regard to sign language DAISY publications, various implementations are possible and the authors intend to release test sets for the upcoming DAISY 4 standard to evaluate new DAISY player implementations. Future work will look at implementing a sign language module and name space to extend DAISY with sign language notations and gesture animation scripts, along with the implementation of signing avatars within a DAISY player.

Acknowledgements

References

Coetzee, Louis / Viviers, Ilse (2008): Accessibility Observations of Visually Impaired Users using the South African National Accessibility Portal. In: IST-Africa 2008.

Coetzee, Louis / Viviers, Ilse / Barnard, Etienne (2009): Model Based Estimation for Multi-Modal User Interface Component Selection, pages 11-16 Proceedings of the Twentieth Annual Symposium of the Pattern Recognition Association of South Africa: CSIR Meraka Institute.

DAISY (2009a): Video in DAISY in order to play back synchronised text and Sign Language. DAISY: R81.
http://www.daisy.org/z3986/requirements/detail.shtml?i=81

DAISY (2009b): Support video as content type. DAISY: R41.
http://www.daisy.org/z3986/requirements/detail.shtml?i=41

Eberius, Wolfram (2008): Multimodale Erwiterung Und Distribution Von Digital Talking Books. Germany: Technische universität Dresden.

Fédération Internationale de Football Association (2008): Laws of the Game 2008/2009. Switzerland: FIFA.

Gylling, Markus (2009): Emerson - a cross-platform Epub and DAISY reader. DAISY: Finland. http://code.google.com/p/emerson-reader/

Kawamura, Hiroshi (2008): DAISY for All Project 2003-2008. Tokyo: Daisy Consortium.

Knol, Joha (2003): Idair. Iduna: Netherlands. http://idair.sourceforge.net/

Larsson, Patric (2009): DAISY As Text Availability Tool for Sign Language - Suggested Video Functions in Daisy. Sueden: Swedish Institution for Special Needs Education.

Maharaj, Shakila (2008): Daisy For All South Africa Report. Johannesburg: Disability Management Services.

Martnez, Francisco Javier Dorad (2008): DBR Daisy Book Reader. http://sourceforge.net/projects/dbr/

Olivrin, Guillaume (2008): AbTi Ninth Report: Adjusting literacy with language transformations and alternative representations. South Africa: CSIR Meraka Institute.

Pyfers, Liesbeth / Robinson, Joanne / Schmaling, Constanze (1998): Signing Books for the Deaf in EU-countries: State of the Art. EU:TAP-DE.

Reiser, Kirk / Zhang, Jinghao (2003): Listen-Up. University of Western Ontario: Canada. http://www.linux-speakup.org/listenup.html

Stallman, Richard (2008): The Free Software Definition. Free Software Foundation. http://www.fsf.org/licensing/essays/free-sw.html#documentContent

Van Der Walt, Willem / DeMeglio, Marisa (2008): [Amis-devl] Code compilation. https://lists.sourceforge.net/lists/listinfo/amis-devl: mailing-list.

Van Der Walt, Willem (2009): Comparison of Open-Source DAISY players under Linux. South Africa: CSIR Meraka Institute. http://www.napsa.org.za/portal/public/content/view/viewContent.jsf?selectedContentId=281664

Ytterbrink, Samuel (2008): DAISY Delight. Luleå Tekniska Universitet: Sweden. http://sourceforge.net/projects/daisydelight