

# Sign Language Interpreter Module: Accessible Video Retrieval with Subtitles

Primož Kosec<sup>1</sup>, Matjaž Debevc<sup>1</sup>, and Andreas Holzinger<sup>2</sup>

<sup>1</sup> University of Maribor, Smetanova 17, SI-2000 Maribor, Slovenia  
Faculty of Electrical Engineering and Computer Science

<sup>2</sup> Medical University Graz, A-8036 Graz, Austria  
Institute for Medical Informatics, Statistics & Documentation  
Research Unit HCI4MED

{pkosec,matjaz.debevc}@uni-mb.si,  
andreas.holzinger@meduni-graz.at

**Abstract.** In this paper, we introduce a new approach to the integration of sign language on the Web. Written information is presented by a Sign Language Interpreter Module (SLI Module). The improvement in comparison to state-of-the-art solutions on the Web is that our sign language video has a transparent background and is shown over the existing web page. The end user can activate the video playback by clicking on an interactive icon. The mechanism also provides a simplified approach to enable accessibility requirements of existing web pages. In addition, the subtitles are stored externally in the Timed Text Authoring Format (TTAF), which is a candidate for recommendation by the W3C community. Empirical results from our evaluation study showed that the prototype was well accepted and was pleasant to use.

**Keywords:** Web, Multimedia, Video, Sign Language, Subtitles, Deaf, Hard of hearing, Human-Computer Interaction (HCI), Usability Engineering.

## 1 Introduction

The World Wide Web (the Web) has evolved to become an interactive multimedia environment. However, information for the broader population is mainly provided as written text, static images or static videos. The existing accessibility guidelines for rich material on the Web (Web Content Accessibility Guidelines, International Telecommunication Union and International Organization for Standardization) are unfortunately too general and inappropriate for the specific needs of users with disabilities, such as deaf people who use sign language as their first language and consider the national language as their second language. The areas of difficulty for deaf people include a possible lack of reading and writing abilities. Based on research studies, many have a low level of literacy and a poor understanding of wording. For this reason, the written information on the Web should alternately be presented with videos of sign language interpreters.

In this paper, based on our extensive research on the needs and requirements of the deaf and hard of hearing community [1], we introduce the next prototype of Sign Language Interpreter Module (SLI Module), which considers the following four aspects:

- (a) Providing accessibility feature for deaf and hard of hearing to existing web pages;
- (b) Inclusion of various form of information (video, audio and subtitles);
- (c) Applying subtitling Timed-Text Authoring Format [22]; and
- (d) Cross-platform browser Flash Player plug-in.

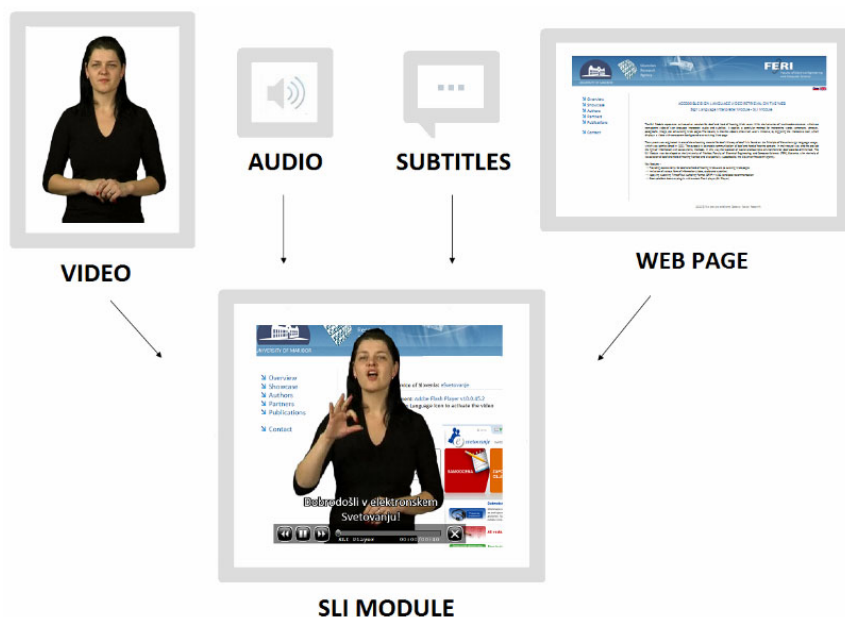
## 2 Background and Related Work

The existing solutions on the Web, for example SMILE [2], ShowSounds [3], Signing Web [4], ATBG [5], SignOn [6], History of the Deaf [7] and Signing Savvy for American Sign Language [8], use various Flash player plug-ins, which are, in most cases, embedded at a predefined web page location. In this way, web designers have to acknowledge the sign language video location while designing web pages. Existing informative websites, which do not consider accessibility recommendations, frequently contain text, images and photos which expand throughout the whole page, leaving no space for the additional sign language video. The previous work, within other research and projects, suggests that for deaf people who use sign language as their first language, a similar idea has not yet been implemented. Because of its simplicity and appropriateness of usage the solution could be integrated into "WCAG Guidelines" for the needs of deaf and hard of hearing people. Currently, the Web Content Accessibility Guidelines definition states that *"using clear and simple language also benefits people whose first language differs from your own, particularly those people who communicate primarily with sign language"* [9], and does not yet determine that the transparent video should be used, which would solve the problem of rapid and easy integration of sign language videos. The main question arising here is how to provide accessible web pages to deaf and hard of hearing web users with sign language videos and subtitles in an effective way, without redesigning entire website from scratch or drastically changing the layout?

## 3 The Sign Language Interpreter Module

The initial idea was to provide the deaf and hard of hearing with a new approach for easy and rapid access to information, tailored to their needs, while not discriminating against other users (non-signers). Past research has shown that deaf signer users, who use sign language as a first and desired language, are often helpless and become confused when searching for information on web pages [7][10]. The deaf and hard-of-hearing population is already familiar with a powerful media channel, the television. In comparison to the Web conceptuality, it is more dynamically driven and visual. For deaf and hard of hearing television viewers, the original video and audio is streamed in the background while the sign language interpreter is layered over with additional subtitling. Our proposal for providing accessibility to web pages is the Sign Language Interpreter Module (SLI Module) [11], which presents an addition in WCAG 2.0

(Guideline 1.4 Distinguishable: "Make it easier for users to see and hear content including separating foreground from background" [9]). In this segment, one can find instructions and guidelines for colour, audio, contrast, resize and images of text. However, it does not provide instructions and guidelines for the sign language video, which occurs mainly in Guideline 1.2 Time-based Media: "Provide alternatives for time based media, which is designed to access to time-based and synchronized media on the Web". The SLI Module uses a custom Flash player (SLI Player), which acts as a wrapper of various modalities: a video of a sign language interpreter, audio translations and written translations as subtitles (Fig.1). The video is shown over the web page with a transparent background and video playback controls (a button for hiding/terminating the video is also provided).



**Fig. 1.** Inclusions of SLI Module: video, audio, subtitles integrated into a web page

### 3.1 Video Preparation Phase

The starting phase is the video and audio preparation. The deaf and hard of hearing who use sign language require the best possible video quality in order to focus on the details, such as finger movements for the sign language and lip movements for lip-reading. Therefore, the quality of the video can only be assured by using a high quality digital video camera. The video of the sign language interpreter should be recorded in front of a green background. There are some regulations for interpreter's clothing (for example, black clothes) which provide the contrast between the person standing and the background. The lighting of the recorded subject is also of great importance. The video recorded on the computer must be uncompressed and in original size. For multimedia container we have used the uncompressed Audio Video

Interleave (AVI), with a resolution of 750 x 567 and 25 frames per second (fps), with 48 kHz audio and 32-bit sampling. The video was then imported into the modelling software, where the Color Key effect was used to remove the green background and soften the edges between the object and the background. This procedure and the use of the Shockwave Flash (SWF) format has resulted in the high quality of the video, presenting the moves of the interpreter clearly enough for the users to see the facial mimicry and to focus on the fingers without a blurry image.

In post-production of the video preparation phase, the result appears as video with a transparent background (non-opaque). The video is then converted to a Flash .FLV format. The audio of a qualified person reading the transcript was recorded additionally.

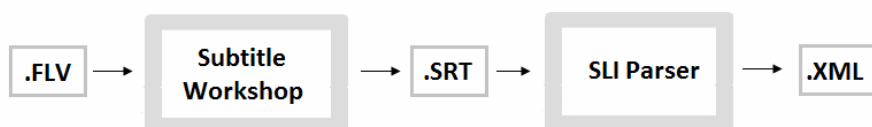
### 3.2 Custom Flash Player as a Wrapper

Adobe reported that the Flash Player plug-in was installed on 99% of Internet-enabled desktop computers worldwide [12]. There are not many Flash players on the market that support TTAF. JW FLV Player [13], which is a free solution for non-commercial purposes and the ccPlayer [14] were initially investigated. Some of the main reasons why we developed a custom Flash player included implementation-flexibility (layering approach and customization of user interface) and proprietary-independence, which were crucial points for further improvement of the graphical user interface based on evaluation studies with deaf and hard of hearing participants. ccPlayer uses a component called CCforFlash for displaying captions with Flash video and audio content. These captions are stored in external files formatted in the TTAF. The SLI Player uses the Adobe Flash FLVPlaybackCaptioning component (version 2.5) for handling subtitles. The component supports Timed-Text (TT) tags for captioning XML files. The style of font and timing are defined by predefined tags. The limitation is the text outline attribute (*tt:textOutline*), which is used to “*specify a style property that defines a text outline effect to apply to glyphs that are selected for glyph areas generated by content flowed into a region*”. To date, this functionality has been enlisted as a non-supported attribute by Adobe Flash Live Docs [15]. The same effect of text outline can be accomplished through the programming language ActionScript 3.0, which is based on ECMAScript, the international standardized programming language for scripting. SLI Player uses this approach and correctly renders the subtitles' appearance; fonts are displayed as white, surrounded by black outline (like on television).

### 3.2 Authoring Timed-Text Subtitles

There are not yet many free tools for authoring subtitles in the TTAF. The present online solutions did not meet our requirements for authoring subtitles. For example, YouTube Captions [16], Overstream [17] and DotSub [18] provide subtitling solutions but with the restriction that the videos are hosted on their web servers. One of the online authoring tools that does not need hosted videos is the SubtitleHorse [19], which is a Flash-based subtitle editor, generating subtitles in various formats, including TT. There are also a few offline applications. Media Access Generator (MAGpie) [20] is a free closed-captioning tool, which is designed to produce and synchronize information for QuickTime™ player. Since the latest version 2.5 of MAGpie does not

support FLV video types, and we found the usage was unfriendly and unnecessarily complicated (no automated process), we investigated another free authoring tool. The Subtitle Workshop [21] from UruWorks is an application for synchronizing subtitles with the video and generating subtitles in all kind of formats. It provides playback FLV video types. The authoring process is quite simple. Although the TTAF export feature is enabled, it does not work in the present version (4.0 beta). For this reason, we have implemented the SLI Parser, which is an automated PHP script for converting .SRT subtitling formats to .XML TTAF. The SLI Parser adds also styling-tags for the outline effect, described in Section 3.2. The workflow of producing the TTAF subtitles is presented in Fig. 2.



**Fig. 2.** Workflow of TTAF. The .FLV video is imported into Subtitle Workshop where the synchronization process with the video playback is done. The output file is the .SRT format and contains time-stamps with text. Next, the SLI Parser converts this file to compliant TTAF.

## 4 The Study

### 4.1 Participants

Thirty-one deaf and hard of hearing students (age range 15-21 with an average age of 17.8 years) from the Ljubljana School for the Deaf and hard of hearing participated in our study to evaluate the prototype. Seven of the thirty-one pupils were in elementary school and twenty of them were in secondary school. There were 23 males and 8 females. 11 were deaf, 14 were hard of hearing, 5 had a speech disorder and one with autism. 14 of them rated their sign language proficiency as excellent, 12 as good and 5 as poor. 18 participants used the Internet daily, 12 occasionally and one never. A user and task analysis was conducted, where data were collected using simultaneous verbal translations with sign language interpreters, note taking and pre/post study questionnaires.

### 4.2 Experimental Design

The aims of the experiment were to examine the end-users feedback of the prototype. It was also used to evaluate satisfaction, acceptance and ease of use among deaf and hard of hearing users. The test subjects were gathered in a controlled environment where the prototype was applied - a mock-up of a real website scenario was designed. Participants were shown a demonstration on the purposes and the various usages and functionalities before the experiment. The test subjects were requested to examine a video (5 minutes long), and as reinforcement training, they were allowed to explore the system alone for an additional 10 minutes. At the end of the session, there was a questionnaire about video preferences. Participants were asked to rank the quality of

their experiences using a 5-point Likert scale ranging from »strongly disagree« to »strongly agree«. Areas of evaluation were quality of video and subtitles, media navigation features and parameters for general impressions, satisfaction, ease of use and comprehension. Finally, there was an open discussion with the help of a sign language interpreter. The purpose of an open debate was to encourage participants to reveal their positive and negative effects or missing features of the prototype.

### 4.3 Results and Discussion

The results reported here relate to participant responses on various aspects of the SLI Module, and are summaries of the notes taken during the study, post-study questionnaires and a group debate. Percentage results of post-test questionnaires are shown in Table 1. 77.3% of the participants only watched the hand movement, however; this was influenced by the high percentage of ‘undecided’ lip-readers (32.2%), who do not generally use lip-reading in everyday life. Subtitles were big enough 77.8% (22 pixels were used) and readable for 77% participants. The video playback navigation controls, were useful for 77.4% of the subjects, while the size was a bit questionable. Those who disagreed (12.9%) suggested smaller buttons. Participants were highly satisfied (80.6 %) with the prototype, 77.4% think the usage is simple. One of the most significant indicators of the evaluation results was the comprehension, which was ranked highly. Due to different modalities (visual, audible or textual) users can choose their modality according to their needs (in some cases all three but in most sign language with subtitles).

**Table 1.** Post-questionnaire results (%)

		Strongly Disagree or Disagree	Undecided	Agree or Strongly Agree
Sign Language Interpreter	Lip-reading	19.4	32.2	48.3
	Hands-movement	3.2	19.3	77.3
Subtitles	Size	9.7	12.9	77.8
	Readability	6.5	16.1	77.0
Video playback controls	Size of buttons	12.9	19.3	67.7
	Usefulness	12.9	9.7	77.4
Usability	Satisfaction	6.4	12.9	80.6
	Ease of use	6.5	16.1	77.4
	Comprehension	3.2	12.9	83.8

The group debate confirmed some important implications for the further development of the prototype. Participants claimed that the video was too fast, and that they would need to slowdown the video, hence it is appropriate to implement a feature for slowing the video playback of the sign language interpreter. Few participants wanted to move the video around the screen; one suggested that the video could be shown in the right-bottom corner similar to on television. Almost all participants were the same

mind: video subtitles and speech must be synchronized without any latency. Therefore, authoring subtitles by a sign language interpreter was suggested by interpreters who were also teachers for deaf and hard of hearing students.

## 5 Conclusion and Future Outlook

In this paper, we have provided a technical overview of the Sign Language Interpreter Module. The system presents a combination of multimodal information such as video, audio and subtitles, and offers the option of prioritizing the sign language on the Web for deaf and hard of hearing users. The evaluation findings of the second prototype again showed highly motivated test subjects and their great desire to use this form of sign language presentation. Some of the prototype's features, for example slowing the video playback, moving the video around the screen and the redesign of the video playback controls, have to be taken into account for further development. We also found that, although the synchronization latency of video and subtitles was minimal, it was perceived as a distractive factor. Therefore, authoring subtitles has to be done by a qualified person (such as a sign language interpreter). We think that the system is feasible as it essentially changes the way of retrieving information on the Web. The SLI Module is gaining wide acceptance within the deaf and hard of hearing community as it assists them to obtain written information on the Web in various modalities. This would increase their literacy in their second language and enable them to integrate more easily into the social majority, while at the same time preserving their identity.

**Acknowledgments.** The project was partially supported by the European commission within the framework of the Leonardo da Vinci program, project DEAFVOC 2. It is also partially supported by the Slovenian Research Agency in the framework of the Science to Youth program, financial support for young researchers. Special thanks go to Živa Peljhan from the Ljubljana School for the Deaf and all participants involved in the study, as well as to Milan Rotovnik from University of Maribor.

## References

1. Debevc, M., Kosec, P., Holzinger, A.: Improving Multimodal Web Accessibility for Deaf People: Sign Language Interpreter Module (SLIM). *Multimedia Tools and Applications* (2010) (in print)
2. Kronreif, G., Dotter, F., Bergmeister, E., Krammer, K., Hilzensauer, M., Okorn, I., Skant, A., Orter, R., Rezzonico, S., Barreto, B.: SMILE: demonstration of a cognitively oriented solution to the improvement of written language competence of deaf people. In: 7th International Conference on Computers Helping People with Special Needs (ICCHP), Vienna, pp. 17–21 (2000)
3. Vanderheiden, G.C.: Full Visual Annotation of Auditorially Presented Information for Users Who Are Deaf: ShowSounds. In: *Proce. of the RESNA International Conference*, Toronto (1992)
4. Fels, D.I., Richards, J., Hardman, J.L., Daniel, G.: Sign Language Web Pages. *American Annals of the Deaf* 151, 423–433 (2006)

5. Straetz, K., Kaibel, A., Raithel, V., Specht, M., Grote, K., Kramer, F.: An e-Learning Environment for Deaf Adults. In: Proc. of the 8th ERCIM Workshop: User Interfaces for All, Vienna (2004)
6. Hilzensauer, M.: Information Technology for Deaf People. LNCS, vol. 19, pp. 183–206. Springer, Heidelberg (2006)
7. Debevc, M., Peljhan, Z.: The role of video technology in on-line lectures for the deaf. *Disability and Rehabilitation* 26, 10548–10598 (2004)
8. Signing Savvy for American Sign Language,  
<http://www.signingsavvy.com/index.php>
9. Web Content Accessibility Guidelines 2.0, <http://www.w3.org/TR/WCAG20/>
10. Fajardo, I., Cañas, J.J., Salmerón, L., Abascal, J.: Improving deaf users' accessibility in hypertext information retrieval: are graphical interfaces useful for them? *Behaviour & Information Technology* 25, 455–467 (2006)
11. Sign Language Interpreter Module, <http://slimodule.com/>
12. Adobe Flash Player Statistics,  
[http://www.adobe.com/products/player\\_census/flashplayer/](http://www.adobe.com/products/player_census/flashplayer/)
13. JW FLV Player, <http://www.longtailvideo.com/>
14. ccPlayerhttp, [http://ncam.wgbh.org/invent\\_build/web\\_multimedia/tools-guidelines/ccplayer](http://ncam.wgbh.org/invent_build/web_multimedia/tools-guidelines/ccplayer)
15. Adobe Live Docs, Flash CS3 Documentation, Timed Text Tags,  
[http://livedocs.adobe.com/flash/9.0/main/wwhelp/wwhimpl/common/html/wwhelp.htm?context=LiveDocs\\_Parts&file=00000611.html](http://livedocs.adobe.com/flash/9.0/main/wwhelp/wwhimpl/common/html/wwhelp.htm?context=LiveDocs_Parts&file=00000611.html)
16. YouTube Captions, <http://captiontube.appspot.com/>
17. OverStream, <http://www.overstream.net/>
18. DotSub, <http://dotsub.com/>
19. SubtitleHorse, <http://subtitle-horse.org/>
20. MAGpie, <http://ncam.wgbh.org/webaccess/magpie/index.html>
21. SubtitleWorkshop, <http://www.urusoft.net/products.php?cat=sw>
22. Timed-Text Authoring Format (TTAF), <http://www.w3.org/TR/ttaf1-dfxp/>