

American Sign Language of the Web

Deborah I. Fels
Ryerson University
350 Victoria St.
Toronto, CANADA
dfels@ryerson.ca

Jan Richards
University of Toronto
Toronto, CANADA
jan.richards@utoronto.ca

Jim Hardman
Canadian Hearing
Society
Toronto, CANADA
jhardman@chs.ca

Sima Soudian & Charles Silverman
Ryerson University
350 Victoria St.
Toronto, CANADA
csilverma@ryerson.ca

ABSTRACT

The development of non-western character encodings has empowered linguistic communities all over the world to create their own on-line Webs. However, in the case of sign languages, which convey meaning by gestures moving in time and space, the static and textual nature of the WWW medium has, until now, continued to prevent the development of on-line Webs by signing linguistic communities. The challenge then is to enable web designers to create on-line, linked Webs based on moving gestures and signs without the need to use static image or text-based equivalents. We have developed a mechanism, signlinks, that facilitates the development of such Webs, without requiring any degree of bilingualism with a written language for the user. Signlinks use a special form of hyperlinking within video material to enable web browsing without written language.

Categories and subject descriptors: H.5.2 [Information Interfaces and Presentation]: User Interfaces - interaction styles; I.3.6 [Computer Graphics]: Methodology and Techniques - interaction techniques; H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia - navigation, user issues

General Terms: Design, Human factors

Keywords: hypervideo.

INTRODUCTION

The World Wide Web has evolved over time to be an interactive multimedia environment. However, the dominant form of interface design remains text and static image based. This is especially apparent when the Web is compared with television, a medium that is almost entirely dynamic and visual.

In general, the textual nature of the Web is beneficial because character encoding systems (ASCII, UNICODE, etc.) have been developed for most written languages. These encoding systems are highly efficient and enable automated transformations into other forms (e.g. text-to-speech, text-to-Braille, separating presentation from

content, etc.) that benefit accessibility. In fact, the Web Content Accessibility Guidelines [7] include text equivalents as one of the highest priority recommendations.

Access to WWW content and interactive structures for people with disabilities has been advocated and indeed legislated for some time. However, much access to rich media for deaf and hard of hearing users involves conversion of spoken language and sounds into text. For many deaf users communicating primarily in sign language, print-based material, including closed captions, is experienced as a second language. Creating original ASL content will benefit both users who prefer content in their native language and those who find it difficult to decode the second, or print language of the original content.

Sign languages are linguistically complete, natural language systems that express vocabulary and grammar visually and spatially by series of hand gestures, facial gestures such as eyebrow motion and lip-mouth movements, and body movements that change in time and space [6]. This series of gestures cannot easily be represented by a single written or spoken word, or static image equivalent. American Sign Language (ASL) is the most prevalent sign language used in North America, although it is not the only one. ASL is not a translation of English and in fact differs grammatically in many ways.

Some Web sites do provide sign language content (e.g., [1]) although many of these sites are often language dictionaries or text-based information sites rather than signed web content per se. However, even where sign language videos are used to provide content on these sites, the navigation mechanisms are still often provided exclusively in a second (textual) language (e.g. English), since the static and textual nature of the Web medium makes this the simplest way to provide this functionality. Signing users of these sites must, therefore, continually switch between their language of choice and a second language, which in some cases may present a literacy barrier. It may be fitting to describe most online signed content today as being "on" the Web rather than "of" the Web. This situation is related to the wider problem of providing hyperlinks from within any video. To address this wider problem, Hypervideo systems, such as that described by Shipman et al [5], have been advanced. Hypervideo consists of linking mechanisms contained

within a video itself where links can take users to other portions of the video or to further information about specific elements.

However, while some aspects of hypervideo are relevant to the sign language hyperlinking problem there are some areas of divergence. Hypervideo only provides a link through to a referent section of video from a reference section, and does not support authoring and replaying of a portion of that reference video to establish context. In addition, text is used to describe links - inappropriate to use in a non-text ASL environment.

This paper provides a description of a new interaction technology for providing hyperlinks within video that extends the concepts of hypervideo specifically to sign language-based, text optional web environments.

WEB PHILOSOPHY

At a high level, the WWW may be considered to be a collection of sub-webs, each created by a particular linguistic community (e.g., “English Web”). Signed content will only be “of” the Web when Signing Webs that are similar in nature to other sub-webs have begun to form. A Signing Web would be a collection of Web pages created by a particular signing community (e.g., “ASL Web”) and bound together by sign language based connections.

While this may be an overly simplistic view of the complex nature of the WWW and individual web page design strategies contained within it, one fundamental premise of a Web is based on the notion of sharing and networking thought linking and connectedness [2]. In creating new Webs or participating in existing ones linking functionality must be preserved at the root of each design. Otherwise, there is no Web or grouping but only a series of unconnected content artifacts.

Hyperlinking on the Web is commonly defined as linking points in documents or between documents on the WWW using a “bold or underlined link” [5]. For the purposes of this paper, this definition is too focused on text-based or static image contexts. However, the Signing Web embraces the spirit of this definition by providing linking functionality. This linking functionality is embedded in video or moving content rather than static content and it is visually represented as a frame within a video rather than an underlined text presentation.

THE SIGNING WEB DESIGN

The Signing Web content and navigational structures are based on moving images such as video or animation. Each page on a Signing Web uses a single main video or animation (320px wide by 240px high), featuring whatever signed content an author wishes to create, whether informational, commercial, personal, artistic, or even whimsical. That video content can be captured and edited using a video editor (our ASL Web editor, SignEd, provides basic video capture and editing).

The navigational structure of the Signing Web content is also based on the moving, gestural signs in the main video so that users can browse and navigate between pages in a Signed Web without any necessity for text. We have termed this mechanism, signlinking.

Signlinking: Sign Language Links

Hyperlink design usually involves associating a string of text or image in a source document with the URL of a target document. The user clicks the mouse on the text string or image, which is typically represented by browser applications in blue with an underline, to load the target page. Signlinking involves an equivalent design concept, but realized through moving images in video and animation rather than text. The meaning of the signlink is embedded in the moving images, rather than in a static representation.

In the signlink system, the author flags links in the video for -specified time interval that they can specify (the default interval is 3 sec.). During this time interval, the presence of a link is indicated to the user by a red rectangular outline in the main video, called the signlink video indicator. (see Figure 1). A rectangle shape was chosen as the link indicator within the video so that it would stand out but not distract from the main video information. The rectangle is properly positioned when it surrounds the current speaker’s upper torso, head and arms.

The indicator is displayed in red. Red was selected, despite the strong association of the color blue with hyperlinking, for two reasons. First, red provides a better contrast with dark backgrounds and, second, the use of red allows us to “piggyback” on the fairly common practice in text-based Web pages of changing the color of links (often to red) under the mouse pointer to emphasize that they are ready for immediate selection.

This lets the user know that a link exists, but this is only part of the role that blue underlined text plays on a textual Web page. When multiple text hyperlinks appear in a page, sighted users can scan over them in a single viewing. The overall “picture” and layout of this multi-link environment combines to enable a *gestalt*-type perception of the linked content on the part of the viewer. Specifically, by paying attention to the number and clustering of instances of blue underlined text and images, users can quickly come to understand whether a page is higher or lower in a site structure, what navigation tools are available and which groups of links might be thematically related. Once users understand a page, they can make better navigational decisions with a single click.

Representing the gestalt of a Web page in a video based signlinking system is a challenge because a video is essentially a serial medium. Video output can move along at a steady pace or be jumped arbitrarily forward or backward in time, but the user’s view of the content remains serial and time-based.

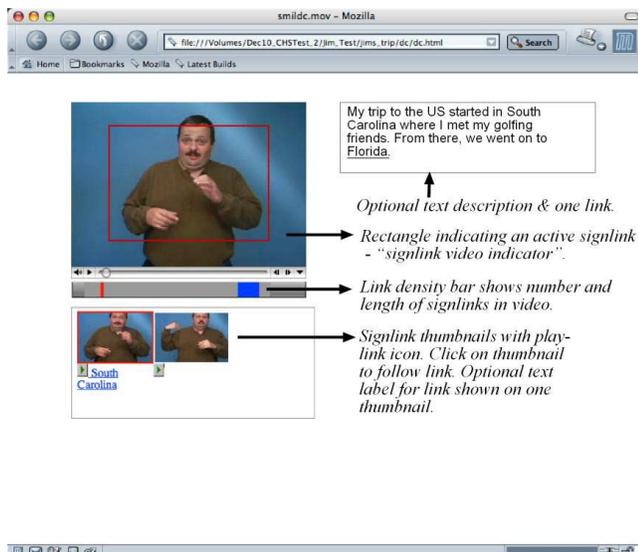


Figure 1: Screen shot of a signing web page.

In an attempt to bring in some of the Web page gestalt to the Signing Web, two views of the signlinks are extracted from the video material. The first view is a set of thumbnail images arranged in a row below the video, and the second is a link density bar that shows the number and relative duration of signlinks in the video.

For each signlinked time interval, there is a thumbnail image that is a full or partial frame captured from within the linked video interval. Each image is given focus (and red highlighting) when the corresponding signlink occurs in the video.

One important design dilemma occurred with how to indicate play and link controls with the thumbnail since “clicking” on a link normally takes the user to the referent URL, while “clicking” on a video thumbnail normally plays the video from the pictured frame. In order to allow users to re-check the context of a link or follow the link, we needed to allow both actions with the same thumbnail.

Our design decision was based on a more common action with Web content - selecting a link to visit it. Selecting the thumbnail image loads the linked URL, whereas selecting the smaller “play link” button below the thumbnail plays the signlink time interval in the main video. This serves to clarify the meaning of the potentially ambiguous static thumbnail image of a sign language speaker.

The list of thumbnails, with their play link (time interval) buttons, is always available, allowing a user to get a sense of all of the possible links contained within a page. Also, a user who has previously visited a page can easily find and select the thumbnail image of remembered signlinks. This provides the user with a partial gestalt of the Web page.

The link density bar provides the user with a graphical depiction of the number, length and distribution of all the signlinks in the video. Clicking on a link indicator in the density bar gives focus to the corresponding thumbnail and

plays the corresponding signlink interval in the video. In addition, when the video is played and the red signlink video indicator rectangle appears the corresponding link depicted in the link density bar changes from blue to red.

Text Equivalentts

In order to support bilingual applications, users who do not sign, and signers who may be losing their sight, three optional text features are available. The first is an *optional* text label that can be added next to the link icon, below each thumbnail. The label is a hyperlink with the same URL as the signlink.

The second text feature is an *optional* text content area. How this text is used is left up to the author, but some possibilities include: a list of searchable keywords, a short description, a full alternate text version, or form controls, if user input is required. The text can include hyperlinks, but using our SignEd software these are limited to target documents that are also linked with a signlink, so that no hyperlink can be made in text that is not also available as a signlink.

EVALUATING A SIGNING WEB – LESSONS LEARNED

Method

Nine deaf ASL speaking youth (age range 18 – 30) participated in a study to evaluate the learnability and usability of a model Signing Web. The test content consisted of a series of nine videos that detailed one person’s trip through the mid-eastern United States, linked together using signlinks. Before viewing the content, the participants were each shown a pre-recorded help video in ASL that explained the general purpose of the Signing Web, and the role of each control in the interface. All videos had standard video controls, stop, play, rewind and fast forward. Data were collected using the Gestural Talk Aloud Protocol [4] with simultaneous verbal translation, notetaking and pre/post study questionnaires. ASL communication was recorded with one video camera while the participant’s screen activities and any verbal translation of ASL were captured with a second video camera.

After watching the introductory video, participants were asked to browse the content through a series of guided exercises (e.g., try a specific link). Once participants became familiar with the interface, they were asked to answer five specific questions about the content. This exercise served to encourage a more complete navigation of the test content. Participants spent approximately 1 hour performing all of the study tasks.

Results and discussion

The results reported here relate to participant responses on various aspects of the Sign Web design, and are summaries of the notes taken during the study and post-study questionnaires. The detailed video data are not reported.

Some striking results with important implications for design decisions appeared. Five participants found the Signing

Web difficult to learn and three found it easy. One was neutral. All participants were able to grasp the concept of signlinking and navigate the various signlinks within the allocated time. They also commented that the ASL Web was an innovative and enjoyable experience.

Seven of nine participants confused the play and link controls for the thumbnails. Some people thought that clicking on the thumbnail image would cue the part of the video that looked like that the image, whereas it actually follows the link. Some thought that clicking on the video would take them to the linked page. In addition, many participants remained uncertain and confused even after explanations were provided and they had experience with these controls. Two of nine participants suggested improving/providing a way of indicating which links had been visited.

As the signlink is a critical component to our design, changes to the design are required. Our second iteration of the signlink design is to maintain the “click on video to play” but revise the link icon and structure.

In the new formulation, clicking on the thumbnail image will play the link (time interval) in the main video. A new and highly visible “visit link” button will be placed under each thumbnail. The visit link buttons will be blue until the associated link becomes current, at which time the appropriate button will turn red, in order to mirror the behavior of the link indicators in the density bar.

A second result crucial to our design was confusion with the red rectangle appearing within in the main video window to indicate a link. Seven of nine individuals found this feature distracting while trying to attend to the video, or were confused as to its purpose (why different size?, why appearing at specific times?), or found it to have a distracting colour. Also, some participants remained confused even after additional explanations of the rectangle’s purpose. Only one participant stated that the rectangle was sufficient and understandable in indicating that a signlink was present at that point in the video. This same individual stated that there should be an English label below the signlink.

A second iteration of the signlink indicator in the main video will replace the rectangular outline with a small, clickable signlink icon in one corner of the main video window with a consistent size and look. This icon appears the same as the icon used for the “visit link” buttons below each thumbnail.

A final result relates to authoring content for Signing Webs, specifically the quality and size of the content and help video material. Participants wanted a larger video format, slower signing particularly finger spelling because it would be easier to understand, and use of more conventional signs. While the size of the video could be under user control,

signing speed and conventionality of the signs are generally a function of the original video content design. Design and best practices guidelines for producing easy to read and understand video material must be provided to new authors so that they can avoid these pitfalls. This is similar to design guidelines for font, sizing, spacing and layout in text-based Webs.

CONCLUSION

We have begun to examine the creation of Web environments that are based on interaction alternatives to text or static graphics. In the case presented in this paper, the alternative is video. Having links within video provides an opportunity for Web users and creators to use non-written or gestural languages as a foundation for their Web pages and hence to form video-linked Webs in and of themselves. Video and video interactivity becomes “of” the Web instead of just “on” the Web.

The signing web structure and content will also help to “level the playing field” between sign languages such as ASL and spoken/written languages such as English, which have been the exclusive means of linguistic representation on the web until now. For the first time, signers are empowered to develop content in that first language.

ACKNOWLEDGMENTS

Funding for this project is provided by the Applied Research in Interactive Media program of CANARIE, grant number CP13.

REFERENCES

1. ASL Quest. <http://www.aslquest.com/mainatt.htm>. retrieved Oct. 1, 2003.
2. Berners-Lee, T. (1996). *The World Wide Web: Past, Present and Future*. <http://www.w3.org/People/Berners-Lee/1996/ppf.html>. retrieved Sept 30, 2003.
3. Poulin, A. & Treasury Board of Canada (2002). *Hyperlinks – Legal Issues*. www.cio-dpi.gc.ca/ip-pi/forum/4/1/2002/09-19/links-liens/links-liens_e.pdf. retrieved Sept. 30, 2003.
4. Roberts V. & Fels, D. (2002). *Methods for inclusion: employing think aloud protocol with individuals who are deaf*. *8th Annual Conference Proceedings/ICCHP 2002*: Springer, New York. pp. 284-291.
5. Shipman, F., Girgensohn, A., Wilcox, L. (2003). *Combining Spatial and Navigational Structure in Hyper-Hitchcock Hypervideo Editor*. *Multimedia 2003*. pp. 124-125.
6. Stokeo, W.C. (2001). *The Study and use of sign language*. *Sign Language Studies* 1.4. 369-406.
7. W3C. (2003). *Web Accessibility Initiative*. <http://www.w3.org/WAI/>. retrieved Oct. 3, 2003.