When Text is not the Answer:
Supporting Low-Literate Deaf Online

by

Doug McNamara

June 2012

Presented to the

School of Information Arts and Technologies

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

Approved by: ____

Kathryn Summers, Ph.D., IDIA Program Director

Acknowledgements

Thank you to Drs. Kathryn Summers, Nancy Kaplan, Stuart Moulthrop, and the Division of Science, Information Arts and Technologies for the knowledge and opportunity to complete this paper, let alone the patience it took to tolerate my constant changing of plans.

Thank you to my students and co-faculty for being the often-unwitting sounding board as I insisted that the only acceptable content was accessible content.

Lastly, thank you to my wife, Jessieca, for understanding the grumpy behavior of a curmudgeonly writer.

Table of Contents

Introduction

Presently, text is the most common medium of accessible content offered to the Deaf[1] community that otherwise would be presented in an audio format if the recipient were not deaf. However, this fails to address the high incidence of low-level literacy within the Deaf community. As news, education and entertainment content are increasingly presented through online, often interactive, channels addressing multiple input modalities including text, a significant portion of the Deaf community has been left out of the discourse commonly shared within the dominant hearing community.

This paper's scope and purpose is primarily that of a literature review exploring the progress made towards support of Deaf access to online information. The first section is a general discussion of Deaf literacy and learning, describing issues that affect both language acquisition and the possible cognitive disconnects with a caveat of how incomplete the research is.

[1] When speaking of the community, Deaf, with an uppercase "D", is the preferred term. When speaking of the physical condition, a lowercase "d" is used.

Section two discusses issues that affect translation, whether that translation is done manually or automated. While manually translating a spoken language to a second spoken language is reasonably straightforward, assuming that the language communities are at comparable linguistic development, the manual translation—interpretation—between a spoken language and a signed language is not. There is not an accepted written form of any signed language and, as seen in Deaf community forums like AllDeaf.com, some animosity exists towards attempts to create one. More than one poster have referred to video blogs (vblogs) as the written form of ASL despite the inherent impermanence of digital video hosted on the Internet.

Section three explores current English to Sign Language translation software and current research towards supporting English to ASL translation on the Internet. The latter focuses on script-driven plugins.

Section four discusses content development and design considerations including the lack of a realistic understanding of Deaf literacy issues made evident by both current law and W3C recommendations. This section concludes with a discussion of the problems faced in plugin development for an increasingly mobile-device oriented Internet.

Section five is a practical attempt to address the previous issues using non plugin-dependent current standards. While this is not a complete solution, and relies heavily on bandwidth intensive video, it allows for device-agnostic content development and is presented as a stepping-stone towards a fuller realization of text-accessibility for the online Deaf community. This involved the development of an HTML5 webpage using CSS3, JavaScript, and two video files. The video files use broadly supported codecs that do not rely upon software decoding. Initially developed only to test cross-platform operability, the webpage was shown to two non-native signers. The opinions provided by them are just that and are included only for completeness.

1. Issues affecting language acquisition and English literacy amongst the Pre-lingual Deaf

Pre-lingual deaf adults in the United States trend far behind their hearing peers in English literacy, reading just below the fourth-grade level (Traxler, 2000). Suggested reasons for this include a lack of exposure to any language during the critical language-acquisition years, Deaf education solely in the tradition of oralism, which involves methods strongly discouraging exposure to a natural sign language like American Sign Language (ASL) instead concentrating on word formation to the detriment of actual content, and both the use and lack of bilingual/bicultural approaches to education. Much of this conflict is centered on whether or not ASL impedes the development of English literacy skills. The "oralist" tradition is one of immersion in an English-only environment where students are taught to lip read, speak and read only English. Often, Signed English would be used as a transitional manual coding system. "Deaf Culture" treats ASL as the primary language and encourages English literacy development as a second language. Gallaudet currently follows the latter. Neither method has been shown to increase English literacy within the pre-lingually deaf population. While all of the above – and more – have been cited as possible causes for a higher incidence of low-literacy performance in the Deaf community, there appears to be little empirical research to support any of them (Spencer & Marschark, 2010).

What is known is that early intervention and language exposure does support language development though at a somewhat below average rate when compared to hearing children (Leigh et al., 2010; Spencer & Marschark, 2010). While pre-lingual deafness occurs in fewer than two births per thousand, it does have a high rate of co-occurrence with one or more additional disabilities at about thirty-nine percent (Spencer & Marschark, 2010). Spencer and Marschark also report that Deaf performance in other academic domains lags behind that of their hearing peers, notably math and science. Deaf may not know when they do not understand written content nor are they as likely to use past learning experiences to foster their own understanding of what they have read (Allen; Marschark, Sapere, et al., 2004; Spencer & Marschark, 2010). This achievement gap grows over time. It is likely that in higher education, deaf students may be lacking in the vocabulary necessary to comprehend technical content areas. Therefore, when creating interactive media content for an audience that includes the Deaf, one must consider not just the language barrier but also the likelihood that the audiences' prior learning may not adequately inform their decision-making or current learning.

The language of the hearing culture is not the language of the Deaf culture. In the United States and Canada, ASL is the preferred primary language for the Deaf community, although it was not until the 1960s that ASL began to see real acceptance as a language (Liddell, 2004). It cannot be overstated that ASL is not English. While the concepts of nouns, verbs and most other parts of "speech" exist within ASL, the syntax is quite different. Verb tense doesn't exist within ASL; the ASL-native relies upon context for temporal setting. ASL does not possess articles, prepositions or grammatical number. On the other hand, ASL possesses a visual organizing system for handling pronouns that exceeds the three n-person pronouns of English. Though concept ("word") order may at times follow the subject – verb – object model used in English (Pinker, 1994), do not confuse ASL with English. ASL syntax tends to follows a topic-comment model much the way Japanese does (Liddell, 2004). To add to the confusion, ASL gloss, the capitalized words found in ASL dictionaries meant for interpreters, looks like English and is spelled the same way as its English counterpart but is not considered to be English (Liddell, 2004).

Regardless of where one might fall on the Language of Thought argument over whether or not we think in our native tongue or a non-spoken[2] language like Mentalese (Pinker, 1994), we can accept that verbal communication between two or more people requires a common language. If there is no common language, then some form of agent must translate. In written communication, this is a reasonable task for a person highly literate in both the source and target languages, given enough time to complete the translation. In face-to-face situations, real-time one-to-one often involves a human translator though there are assistive devices like *iCommunicator* and Signtel *Interpreter*. Both use similar speech-to-text software. While not designed for

[2] In order to avoid excessively long noun phrases, spoken refers to both oral and manual languages but excludes text-based representations of language.

Internet interpretation, both are worth consideration for their innovative use of video. Due to the similarity between both *iCommunicator* and *Interpreter* in their input and output modes, only *Interpreter* is further discussed in the existing software section.

Prosody, the content expressed through tone, pitch, pacing and other non-word elements within a verbal language, is conveyed through non-manual means, using facial expression, gesture and posture in a sign language. This is similar to verbal languages, though phonologically the dominant speaking culture relies upon audible prosody. Verbal languages can be categorized by placement of phonological phrases. Languages like Japanese and ASL place emphasis on the word at the left of a phrase. English places that emphasis on the right. Therefore, there are distinct differences in syntax at the level of both the phrase and the sentence. This impacts both Deaf language acquisition and the design of real-time translation technology. The latter is discussed in the software section of this paper. In "Prosodic Structure And Syntactic Acquisition", Ooyen suggested that pre-lingual exposure to audible prosody within the native language starts the word-order recognition process despite not being able to understand the words (Ooyen, 2003; Nespor, 1999). It follows that exposure to prosody accompanying a given sign language at the pre-lingual stage may be the means in which acquisition of that sign language begins. If Ooyen and Nespor are correct, then early identification of any impairment to language learning is vitally important if the affected child is to trend with its peers. Unfortunately, as late as two decades ago in the United States, that identification was not happening until the 24-month cutoff implied by their conclusion (Spencer & Marschark, 2010). Approximately ninety-six percent of all deaf children are born into homes where the parents are not deaf; therefore the parents are not native signers (Mitchell & Karchmer, 2004). Taken to its logical conclusion, the intersection of these factors leaves the majority of Deafs in the United States at not just a fourth-grade reading level in English, but with the likelihood of significant delayed development in any language, thereby having a strong possibility of lacking fluency in any language, which may cause other, significant issues in cognition (Luft). However, this is another area in which little research has been done with the Deaf community.

Contributing to fluency issues are the variety of sign languages used and who uses them in the US. Manually Coded English (MCE) systems, including Cued Speech, Seeing Essential English, Signed English (mentioned above) and others generally follow English grammar and syntax though these artificial "languages" may have borrowed signs from ASL. However, while some of these systems possess morphemes representing articles, tense and other grammatical structures not found in ASL, these structures are often dropped, resulting in a pidgin comprised of ungrammatical ASL and the MCE used by the non-ASL participant. Signers also code-switch, moving from ASL to an MCE as needed. Unless all involved in the conversation are aware of the switch and have some fluency in both ASL and the MCE in use, this may increase incidences of misunderstanding (Spencer & Marschark, 2010).

Fingerspelling is common to ASL and MCE. In ASL, fingerspelling performs an essential service in quoting English, specifying proper nouns and appears to be the initial step in the lexicalization of “borrowed words” (Liddell). In Science Technology Engineering and Math (STEM) courses, interpreters and the ASL-using Deaf often invent signs for immediate use in situations where fingerspelling would be onerous (personal communication with a deaf student and their interpreter). Over a semester, numerous signs are created to cover that course’s lectures only to be discarded at the end of the lecture or course.

To combat this, websites such as www.shodor.org and aslstem.cs.washington.edu offer ASL signs for use in science, technology, engineering and mathematics (STEM). However, adoption of these terms appears low (Cavender, Otero, Bighamy, & Ladner, 2010). Upon visiting both websites, it is immediately apparent that neither receives much traffic. The latter site uses Facebook but only possesses 111 members (at time of writing) and few comments or entries not posted by a moderator. Considering the lack of traffic on the ASL-based STEM website and the reinventing the wheel method of signs being created in isolated situations, one can surmise that technical ASL vocabulary grows at a slower pace than that of verbal languages. However, like so much else in Deaf research, there appears to be no empirical research suggesting that sign languages might possess a smaller vocabulary than verbal languages.

Languages borrow words/concepts from other languages. Over time, those borrowed words, if accepted and used by the speakers of the borrowing language, are lexicalized in to the language. Within English, this conventionalization may take the form of Anglicizing the loaned word. Fingerspelled words may be added to the lexicon through a process that appears to be truncation of the interior letter shapes, thus creating a new sign, or it may remain a stable fingerspelled word. Both verbal and signed languages create new words and concepts through compounding. In ASL, this compounding can take the form of combining two signs into a new sign or affixing fingerspelling to an existing sign (Liddell, 2004; Brentari & Padden, 2001).

2. Issues affecting translation

Notation Systems

Sign languages do not have a "natural" written form, though a number of writing systems exist, such as HamNoSys, Si5s, and Sutton SignWriting; the latter is based upon the work of Stokoe (Hanke, 2004). Stokoe's notation shows hand-shape, position of the hand in relation to the body and motion, but does not address non-manual prosody such as facial expressions and body position. Systems such as Si5s, a fairly new system that appears to not be based upon Stokoe's work, attempt to address this but do not have much acceptance. This should not be surprising as ASL is a young language – about 200-years old – with a considerable amount of isolation between groups of ASL users.

Widely used and rooted in Stokoe, the Hamburg Notation System (HamNoSys) is a graphical system for representing hand shapes without reference to any particular sign language (Hanke, 2004). Earlier versions HamNoSys fail to record the non-manual components of sign languages (Hanke, 2004). The current version, version four, indirectly addresses prosody by suggesting a multi-tiered notation system using some other representation of prosody due to technical limitations within font descriptions and the number of characters used to represent a single sign (Hanke, 2001). Though Si5s is superior to HamNoSys in fully representing prosody, Si5s is directly linked to ASL, thus making linguistic comparison between signed languages difficult. The wide acceptance of HamNoSys coupled with its language-agnostic properties makes it attractive for categorizing all but prosody. The latter could be included using a separate system. Hanke recommends the Edinburgh Non-Manual Coding System (ENCS) as it follows conventions established by Stokoe (Hanke, 2001).

However, linguistic notation systems like HanNoSys do not show the meaning of a sign performance as they are not intended for common usage outside of linguistics. Despite its acceptance within the linguistics community, HanNoSys's complexity may be a detriment when translating from English to ASL outside of that community unless that system is used as a transitional method that is not visible to the end user. In pure machine translation, this would be possible, if not preferred. However, a secondary system would still be required, greatly complicating algorithm development for processing. Using a system with closer ties to ASL may simplify translation. Of those systems that do represent ASL without resorting to a secondary system to carry prosody, which HamoSys must do, Sutton SignWriting and Arnold's Si5s can be reproduced mechanically.

Sutton SignWriting, created in 1974, possesses 36,600 graphemes, less than half of the number of characters found in Chinese, easily fitting within the UTF-8 upper limit of approximately 1.12 million characters thus making it a good, if somewhat verbose candidate for computer use (Sutton, V & Slevinski, S.). SignWriting is not sequential, but spatial, in keeping with the three-dimensional quality of ASL. This means that SignWriting requires a plugin to display correctly in a browser or would need to be reproduced as either a PDF or an image. The latter can be made accessible to the deaf-blind by using CSS to include the English text though changing its visibility or display thus providing Braille haptics with the ASCII text needed by the haptic to convert from the visual modality of print to the tactile modality of Braille.

However, there is a SignWriting fingerspelling font, Sutton US, which would allow for recoding from English alphanumeric characters into ASL fingerspelling without requiring the content provider to have any knowledge of ASL. While this method works for direct translation between languages that follow a similar word order, therefore appearing to offer some support, it may be no better than reading English as the English grammar, spelling, and structure would be preserved. That preservation would include fingerspellings of articles like "the" that have no ASL equivalent. What it may offer is an intermediary step in learning written English, therefore it should not be discarded out of hand and may be used as a tertiary method for content display.

Word-for-word translations between written languages often fail to convey context and idiom as the source language may be culturally quite different (Banjar, 2008). In order to capture the contextual meaning of the content presented in the source language and transfer that to the target language, a translator needs both literal and cultural fluency in both the subject and target languages allowing for “cultural transplantation” (Ghadi, 2010). If there is to be useable machine translation between English and ASL, it appears that cultural fluency would be required, as many idioms do not transfer between the source and target languages. To do this at the machine level, a body of language performance would need to be collected and annotated in such a manner that allows access to the performance through algorithms capable of contextual comparison and selection. This body of collected language is known as a corpus.

Corpora Development

In linguistics, a corpus is a large body of annotated text gathered and archived for research purposes. For comparison and study, a corpus should be designed around how well the text represents the actual language and not the actual content of the text. If designed – limited – around the content, then the corpus will reflect the attitude of the designer towards the content. This creates a non-representative corpus and should be avoided (Sinclair, 2005).

Computational linguistics concerns itself with machine translation. A common reason to build a corpus is for translation. Unlike language-to-language dictionaries, the study of a corpus gives insight into a language's grammar and syntax. Direct attempts at translation rely upon *string-substitution,* using rather straightforward arithmetic programming to substitute words and phrases through comparison (Huenerfauth, 2004). This would be the language dictionary approach. However, languages are far too complex to be mined in this fashion. Idioms that would not be taken literally in the source language (SL) are translated as if they were literal content for the target language (TL); words from the SL that do not have a direct analogue in the TL would not be translated into the TL; and the grammar of the SL would be preserved, turning the resulting translation into a pidgin.

As mentioned earlier, grammar, syntax and phonology can differ widely between languages, thus requiring an approach resembling emergence or evolution following a set of rules or algorithms that match the rules of both the source and target languages. Designing such an interlingual system, one that can do syntaxical analysis, requires parallel corpora, a corpus translated into at least one other language (Huenerfauth, 2004; 2008).

English follows a subject-verb-object (SOV) phrase arrangement. Therefore, we would expect multiple rules that would parse an English utterance like "Bob threw the ball" into its respective phrases. ASL may follow the same order, or its order may be topic-comment (OSV). So, a simple English sentence like: "Bob fed the cat." could be glossed in SOV:

B-O-B FEED CAT

Or in OSV:

CAT B-O-B FEED

Both are equally valid with the latter appearing to be in the passive voice, which may be the past tense rendering. Therefore, it is possible that fluent ASL would have an algorithm that references word order to determine tense yet the lack of verb tense still makes this ambiguous. Including time would improve the algorithm and it appears that ASL does just that at the beginning, following a time-topic-comment order. However, as with other languages, fluent ASL users may drop phrases, relying upon the other person in the conversation to place the sentence into context (Liddell). In situations where the signer or speaker drops words and phrases, a human interpreter would have enough knowledge of the source and target languages to be able to translate through context with little trouble. (Sinclair, 2005). In order for machine translation to work at this level of sophistication, the corpus would need internal limitation thus violating Sinclair's admonition, as mentioned at the start of this section. However, Huenerfauth claimed that there is not a natural-source ASL corpus (Huenerfauth, 2004). This appears to ignore the existence of ASL video blogs (vblogs), which could be collected, mined, and annotated within a research milieu. However, as these vblogs are published, no matter how informally, they enjoy copyright protection and could not be used within the development of product or service without permission. Given the level of hostility shown towards non-deaf researchers on forums such as AllDeaf.com, it is possible that obtaining permission would be difficult. Also, there appears to be little in the way of parallel formal corpora, which are often a byproduct of government support for multiple languages (Huenerfauth, 2004). Common parallel corpora would include Federal Income Tax forms available in both English and Spanish and the state of Louisiana's website, which is available in English, French and Spanish.

3. Existing Software

Authored translation

One translation software package, *Performing Hands* used a custom-created *Shockwave* Xtra that took input from Signing Gesture Markup Language (SiGML) content to control an avatar built in *Macromedia Director*, now an ignored Adobe property (SITE). Unfortunately, updates to any ASL animation would require finding the authoring software, and would-be users would then probably find that the plugin is not supported on many mobile devices. This also rings true with solutions developed to run on Sun Microsystem's Java Virtual Machine. Third-party plugins are essentially banned from Apple's iOS devices, and Adobe abandoned mobile Flash development in November of 2011, essentially removing it from Android and Windows-based mobile devices. As mobile devices continue to supplant the full-featured laptop, the use of a third-party plugin is contraindicated.

An alternative to plugin use would be an XML-based language that controlled an animation written using an open-source language. *eSign* and Visicast used SiGML within very constrained domains, i.e., employment postings and weather forecast broadcasting, respectively, to translate without requiring the content-creator to know sign language (Zwitserlood). While both are self-contained software, these are noteworthy in that they used British Sign Language (BSL) instead of Signed English.

The present gold standard is Vcom3D's *Sign Smith Studio*. This tool, used to develop the video in my demonstration site, can encode text into ASL signs. Those English words that do not have an ASL sign are fingerspelled. There are two channels for facial expression, thus capturing prosody. The authoring suite also possesses machine-speech capabilities to allow non-signing hearing users access to the content being signed. This tool is not designed for real-time translation and requires fluency in both ASL and written English to effectively author content.

Real-time Translation

Signtel Inc.'s *Interpreter* uses video of an interpreter and has a robust vocabulary of 30,000 words but only translates from English to Signed English – not ASL. While translations from

English to Signed English have articles and prepositions removed, the word order remains in English and the words are replaced with ASL signs. *Interpreter* can be forced to fingerspell. This makes translation a straightforward process, basically little more than string comparison until the application encounters a word not in its dictionary, forcing it to fingerspell the word. Common English idioms are translated into useful information and not word-for-word as those idioms may be absent from the Deaf community (Signtel Inc.).

Interpreter is designed with face-to-face translation in mind but does have the ability to translate text files. Despite claims to work with "Internet text," the application does not interact with webpages in any real fashion. The user must insert the text into the application through either drag-and-drop or the key combination of Ctrl + Insert.

While neither an online interpreter nor a truly ASL-based tool, *Interpreter* is noteworthy for its use of video segments to deliver the signed content. Signtel claims that the video segments are seamlessly stitched together (Signtel Inc.). While this may be true for the performance of the individual signs, there are considerable jumps between them as well as an "unreal" appearance when the signer is at rest that is somewhat disturbing. This disturbing "feeling" is known as the "Uncanny Valley" and is a phenomenon more likely to be encountered in realistic-appearing virtual avatars than in video (Tinwell*, et al*). Tinwel et al., found that the lack of upper-facial expression in digital avatars repulsed viewers. However, the stilted sign performance and oddly dead-appearing facial expressions of the live actor approaches the Uncanny Valley from the other side.

iCommunicator has a similar design, allowing for voice, text document and direct text input with video as the output (*www.icommunicator.com*). However, both are meant for real-time face-to-face translation. They rely on installed dictionaries, therefore, it is understandable that they return Signed English and not grammatically correct ASL. Neither translates from sign to text or voice. As discussed in the next subsection, accurately capturing and encoding ASL performances for translation is difficult.

Web-based Machine Translation

At present, the most promising attempts to provide real-time written-to-signed translation use avatars. While *iCommunicator* is capable of stitching together human signing, and claims to read an HTML document without relying on copy-and-paste, it has a limited vocabulary, one that cannot be updated, and is limited to the PC (*iCommunicator)*.

Developed during the early 2000s and then apparently commercialized through Sys Consulting, *eSign* used an XML-based language to deliver a prosody-capable avatar to a deaf user. This markup language, Signing Gesture Markup Language (SiGML), resembles HamNoSys but is far less complete in its representation of possible signs. While their aim appeared to be the development of a plugin that produced a natural signed language, and not a signed pidgin like Signed English, this project acknowledged that machine translation was not sufficiently mature for the task of translating between languages of different structure. However, their research during the development stage strongly supports the need for correct prosody in aiding sign comprehension. Without the correct prosody, comprehension of both single sign and utterance, the latter is a complete performance of a single thought, was approximately forty percent. With the correct prosody, recognition of the single signs and utterances increased to eighty-five percent. However, this level of accuracy requires a fluent signer to build the animation (Kennaway, 2007).

Both *eSign* and *iCommunicator*, as well as others built in a similar manner, must have a human translator. Even without the non-trivial cost of recording the original sign language performance, these tools put the burden of translation on the content developer. Each content change requires a new translation. While all appear to allow random access to developed content, the author's experience with Sign Smith Studio suggests that it is difficult to produce comprehensible performances without losing synchronization of prosody, mouthing, text captioning and voice-over without spending a considerable amount of time. Content developers required to update page content multiple times throughout the business day, i.e. news sites, would find this task to be well beyond reasonable accommodation. As mentioned in a later section, this level of accommodation is not required; so, it would not be done. Therefore, automated translation must be done if information accessibility is to be possible without requiring the Deaf community to become fluent in the larger hearing community's spoken language.

Currently, there does not appear to be any commercially available machine translation (MT) for written text to ASL. Dr. Huenerfauth, currently at CUNY, appears to be the closest to developing such an application. While working on multiple development paths, the application is still in a preliminary state: corpora are in development as well as algorithms for displaying prosody and phrasing. For testing, they are using Sign Smith Studio to produce their avatar (Huenerfauth, 2004, 2006, 2008; Huenerfauth, et al. 2008).

Huenerfauth's team uses motion capture, going beyond the hands to include facial expression, in the development of a "designed corpus" of formal, fluent ASL. Presently, they are in the evaluation stage of development and do not have a complete corpus. However, the data collected from their current ASL performances have been used in their evaluation of prosodic elements. Unsurprisingly, Huenerfauth confirms the need to include prosody within a virtual avatar's ASL performance as it significantly improves understandability of the ASL performance (Huenerfauth, et al. 2011).

While his work appears to be the most promising solution, it is far from complete. Given the current state of MT development, it should be readily apparent that serving pre-authored video alongside or in place of text is necessary to improve access for the low literate within the Deaf community.

4. Considerations for support

Content development

While pre-lingual deafness is a low-occurrence event, it has a high rate of co-occurrence with some other disability. Unfortunately, the research is vague concerning which cognitive disabilities and the specific likelihood of co-occurrence, reporting a range between three and sixty percent of co-occurrence (Spencer & Marschark, 2010). It would be prudent to err towards the upper end of the above range. Therefore, if the content allows, a reduced vocabulary combined with short paragraphs employing a simplified sentence structure may be appropriate. Low-literacy readers often avoid reading long, complex prose or may read word-by-word (Chadwick-Dias, A., McNulty, M., & Tullis, T., 2003; Summers, K., & Summers, M., 2005). However, a tersely written text with a limited vocabulary does not increase literacy and fails to address situations where reducing the vocabulary is neither possible nor desirable. For those instances, expanding the text to include term definitions, examples, and synonyms may increase comprehension and may be a viable solution for those whose low-literacy is due to the content being provided in the end user's second language. Also, the heavy use of graphics in addition to the above recommendations may increase comprehension (Watanabe, *et al*., 2010). However, this contradicts Mayer's recommendation of eliminating redundancy to avoid channel overload (Mayer, Heiser & Lonn, 2001).

Presentation and Markup

As English is not the native language of the Deaf in North America, the Watanabe, *et al*. recommendation to include informational graphics may be valid, however, a mechanism to expand and contract the delivered content is also likely to be valuable if provided. For text, possible solutions include annotation or expandable/collapsible text divisions such as that provided by Spry and jQuery but with markup to allow assistive technology users to skip the collapsed text. This recommendation is predicated upon Mayer's research showing that channel overload, which occurs when information is presented through competing streams, reduces comprehension (Mayer, et al, 2001). In this situation, one must remember that text and sign are both verbal streams regardless of the manual code used in the signing stream (Pinker, 1994). Therefore, it would be unwise to provide a signing-avatar video along side of either the English text or the fingerspelled text. There should be a means to hide one while using the other.

However, this does not address the language barrier. The fact of the matter remains: for the American signing community, written English is a second language at best. If the deaf reader is already starting with delayed language development, a simplified second language may not increase comprehension and should not be the only alternative. The Rehabilitation Act may not go far enough by not mandating sign language in addition to English-based captioning. However, this is a contentious issue as it takes the nucleus of American Deaf culture, ASL, and places it at the level of an accommodation; this is not done for any other language (Clark).

From a developer's point of view, WAI and ADA compliancy begins with writing standards-compliant HTML. However, no matter how compliant the HTML is, it only defines document structure, leaving presentation to CSS. Presently, CSS2.1 presentation offers minimal support for text-alternatives, primarily in the form of aural style sheets and Media Types. As of this writing, the aural media is deprecated, though still supported by those browsers that have offered support to aural media in the past, and the replacement, media=speech, has not been defined (w3.org, 4/12/2011). Similar but expanded presentation rules will be defined in CSS3. However, there does not appear to be any sort of CSS specification for text-to-ASL readers. However, this does not mean that CSS support for ASL could not be included in future drafts or recommendations. At this time, the reasons why they are not included are only speculation. Those reasons include:

1. No sign language browser currently exists.
2. Sign language display would be translation (Clark).
3. Currently, sign language sites for the deaf are not accessible to the blind. Many of those sites are not accessible unless you speak ASL (Clark).
4. The Americans with Disabilities Act (ADA) does not mandate this level of accommodation, as a browser is a tool and not considered a channel of

communication despite the precedent set by the Television Decoder Circuitry Act of 1990 (United States Access Board).

5. Translation would be better handled through the use of a plugin or other adaptive technology.

The first item, the creation of an ASL browser, is not on the horizon. The cost would be prohibitive to build for such a small user population though it may be the optimal choice.

To expand upon items two and three, multiple language support is not seen as an area to be addressed under section 508 of the ADA. Language is acquisition is a seen as a matter of education and culture. For example, Spanish, spoken at home by 12.8 percent of the population of the United States, is the second largest language group after English yet there are no requirements to offer Spanish language support within the current ADA despite the fact that 44.7 percent of that population, 5.7 percent of the total United States population, does not speak English very well (U.S. Census Bureau). Requiring ASL support through the ADA may open the door for a similar requirement for Spanish. This would also require ASL-oriented sites, which often rely heavily on video, to closed-caption those videos to provide text marked up for Braille as well as an audio channel so that the blind and deaf-blind could access the content, provided the site would normally be required to adhere to ADA requirements. The latter scenario, support for blind users, is mandated by the ADA and requires text if it is to support deaf-blind users as they commonly use braille haptics. In any case, there is no federally mandated official language in the United States.

Four, the ADA's lack of ruling on whether or not a tool must be accessible is self-explanatory: if not mandated, there is no need to accommodate it. Five, the belief that using a plugin or third-party software to handle NLP, offers multiple technical issues despite making the most sense for the end user. Third-party plugins are banned from iOS devices (iPod, iPad and iPhone), poorly supported on early Android mobile browsers, and no longer available for current Android releases as Adobe has stopped development of the various mobile browser plugins to focus on Adobe Air, their current software development kit. This effectively removes Flash from all mobile devices. Java is still available for Android devices but the inconsistent operating system rollouts from service providers make development of robust apps difficult. While such voice recognition apps exist on both Android and Apple devices, these apps require access to a server to handle the actual processing, as NLP is resource-intensive.

Technology Support

Modern browsers, including all current native browsers for mobile devices, support JavaScript, including jQuery, and both the HTML5 Canvas and HTML5 Video elements to varying degrees. The latter allows for hardware decoding of modern video formats, specifically H.264/MPEG-4 and Theora OGV, which run natively inside the browser (Jobs, 2010). However, during my functional test of the above-mentioned technologies, this did not hold true for iOS 4 on the iPhone 3G despite claiming full support, though the supported file type will play in *Quicktime/iTunes*, and video-playback was spotty on Android devices with full support available on the current Samsung Galaxy tablet though there was no support beyond English text display on the Kindle Fire. Both iOS and Android mobile devices supported hosted web-fonts with the expected exception of the Kindle Fire. The Blackberry was able to display the signed movie. This is discussed in the following section.

5. Application and recommendations

Initially, there were no plans to build the wireframe or movies. However, during the writing of this paper, the current releases of the major browsers began to phase-in support for HTML5 and some of this new standard's associated technologies. Therefore, the development of a platform-agnostic website that would not require plugin use was in order. The only goal was to see if a working prototype could be built and to test its compliance with mobile device browsers. That there were any signers available to test the wireframe was purely serendipitous.

The tested wireframe is perpetually hosted at http://www.dougmcnamara.com/textnotanswer/ and will remain available, unaltered. Screenshots are to be found in the appendix. The first three are of the original site. The following two are of the simplified version. The movie, a Signed English equivalent of the text seen in the first tab, is a scripted animation built in *Sign Smith Studio* using a pre-developed avatar. While the prosody is incorrectly coded due to the author's unfamiliarity with ASL, Signed English often drops prosody to better preserve the source language's structure (Heunerfauth).

In the test website, the non-English tabs were spelled out in fingerspelling using Sutton US. While not formally tested with anyone in the Deaf community, two non-native signers found them nearly impossible to understand due to unfamiliarity with Sutton SignWriting and the font. When asked, both said that no one would willingly read the fingerspelled text. This may confirm the animosity, seen on the AllDeaf.com forums, towards attempts at developing a written form of ASL as both this, an admittedly painful representation of English devised by hearing individuals, and Si5s, a writing system created by Robert Arnold, have received a considerable amount of criticism on the forum at AllDeaf.com. Arnold's offering appears based in a culturally Deaf understanding of English as an intrusive force. It is likely that the opposition to Si5s is as much of a reaction to Arnold's cultural stance as it is to the written language. They also commented on the signing avatar's performance, recommending that a fluent ASL signer should create the video as much of the Signed English was either fingerspelled or unneeded words like articles, which do not exist in ASL. The latter was as much the fault of the author as it was the fault of the limited dictionary. The English text is designed to be difficult with a Flesch Reading Ease score of fifty, the equivalent of twelfth grade. Both signers had no difficulty reading the text.

Based upon their recommendations, the second version, simplified.html, is far easier to read at a Flesch Reading Ease score of seventy-three. However, this score is too low as it places the reading level at seventh grade. Unfortunately, the text could not be reduced further without significant changes to the meaning of the text.

Recommendations

The current level of mobile device development along with the lack of support for non-native applications and plugins suggests that a browser plugin would be problematic at best. Until this situation improves, the following recommendations should be taken into consideration in situations where supporting a deaf audience is a concern.

Whenever possible, text should be reduced. Simplifying grammar, reducing the morpheme count per word and limiting the use of technical jargon will create a more accessible text. As the average pre-lingual Deaf reads English at a fourth-grade level, the Flesch Reading Ease score should be better than one hundred to include them. Furthermore, this suggests that the Signed English animation would still offer difficulty to those users. Therefore, the signing avatar should sign in the native sign language used, instead of a pidgin MCE. This follows Huenerfauth's research even though current ADA law does not require it. While this is contrary to Clark's response to supporting deaf internet users through sign language, Traxler's findings, along with those of Spencer and Marschark, suggests that this goes well beyond the "it's a matter of education" defense implied in Clark's response.

Moving the text display to different tabs using Spry, an implementation of AJAX, or using jQuery's Lightbox, a JavaScripted method of overlaying content – usually images or video – to separate redundant text may serve to eliminate the effect mentioned by Meyer, et al. Also, using HTML 5 as the display language allows the content to be loaded into mobile social media apps, such as Facebook, which often rely on HTML5 for presentation, served by some form of dynamic language like PHP, to handle text and graphic display. It also eliminates the need for a plugin provided multimedia elements are designed to run in the appropriate HTML 5 element.

Multimedia, like Adobe Flash, audio and video, should continue to offer captioning for deaf-blind users as braille haptics that require ASCII text input. However, to better support pre-lingual deaf users, a signing avatar should be provided in addition to the text. To avoid channel overload, selecting one option should deselect the other.

Appendix

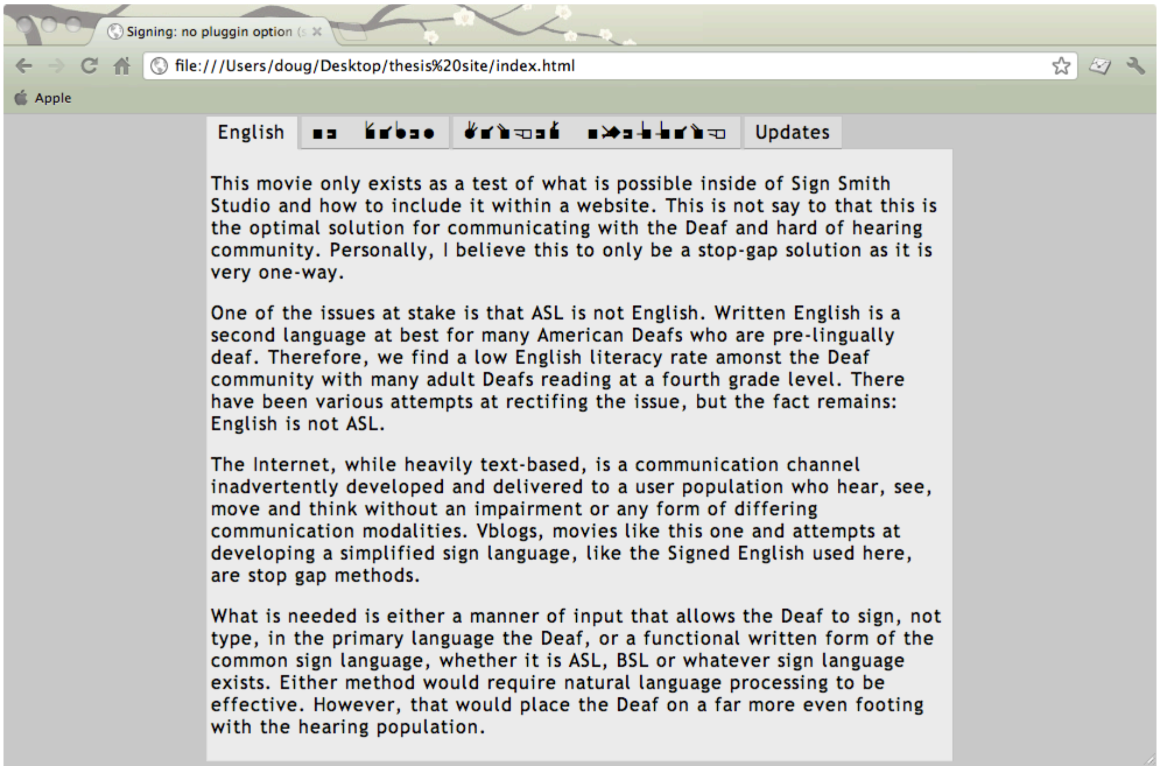

Figure 1: Full text. Flesch Reading Ease score: 50.

Figure 2: Signed English Movie

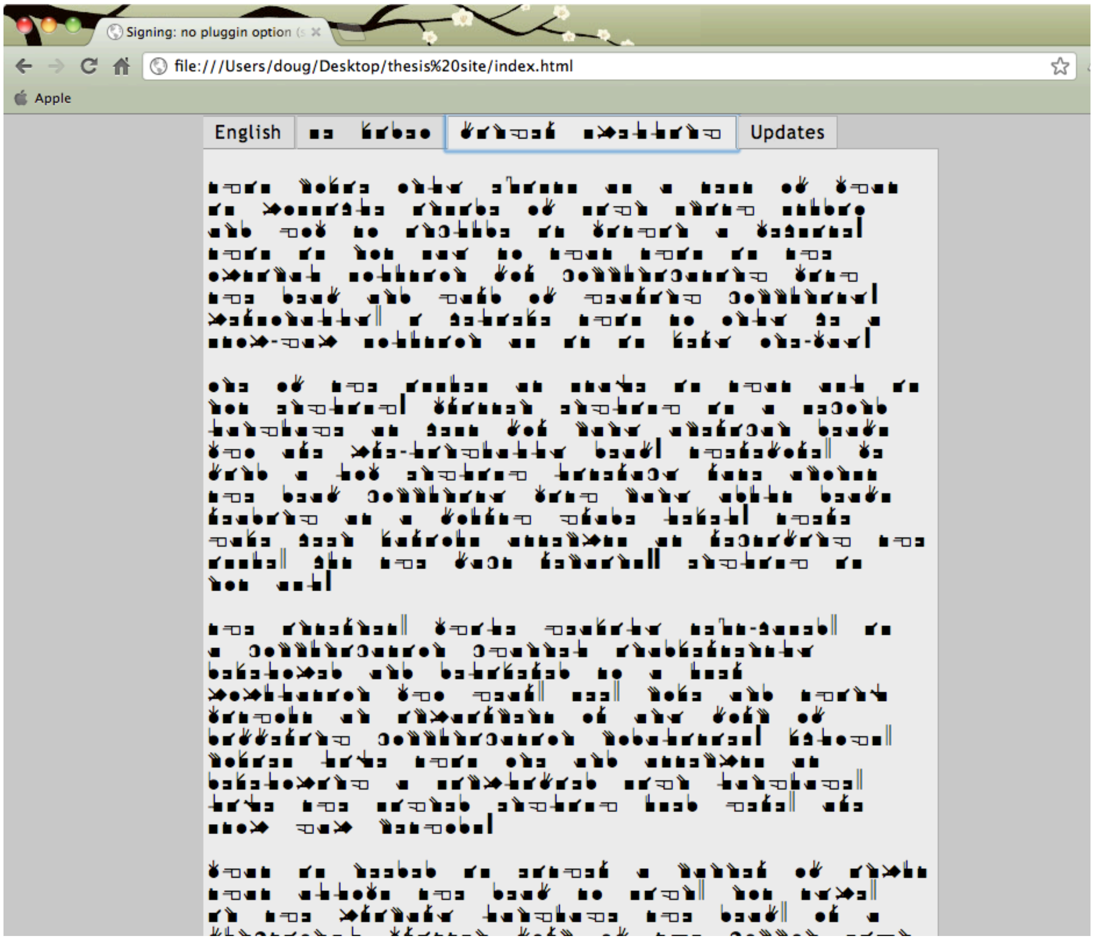

Figure 3: Fingerspelled English using Sutton US font.

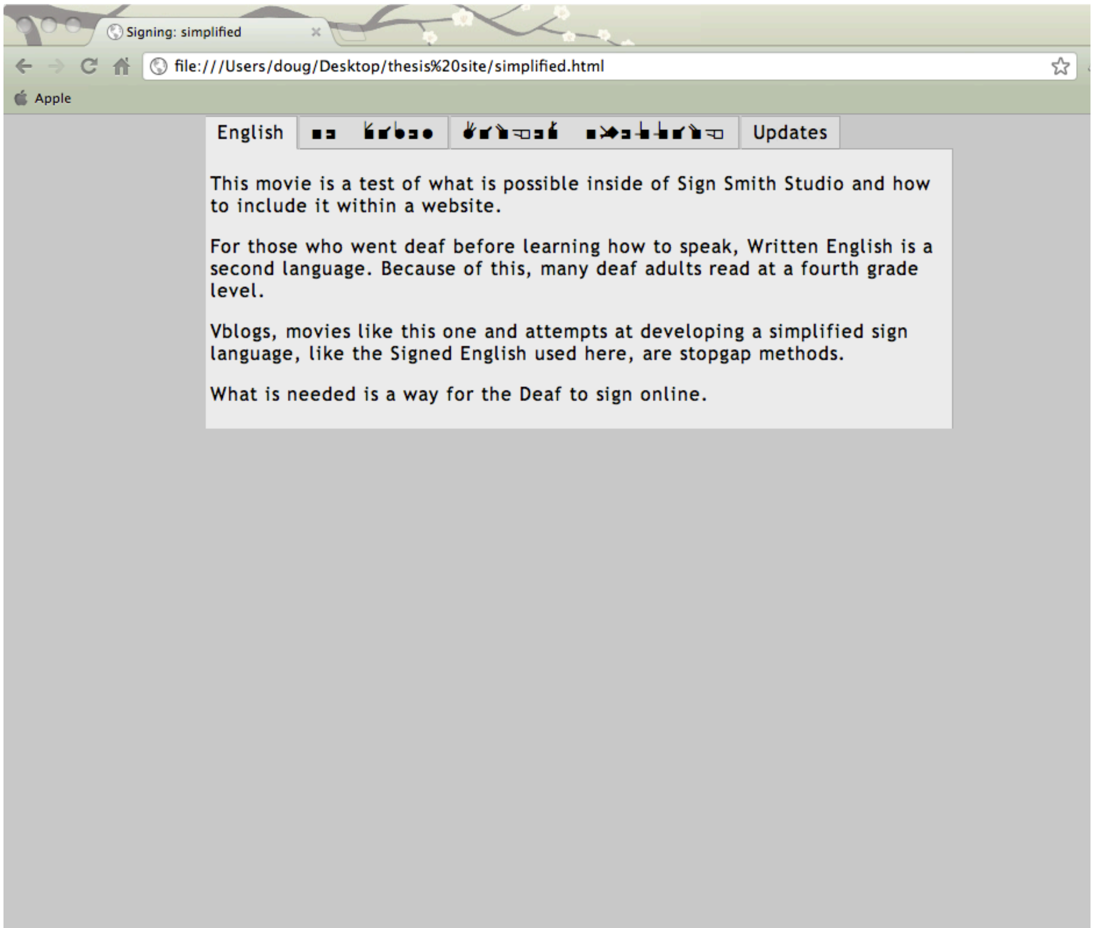

Figure 4: Reduced text. Flesch Reading Ease score: 73.

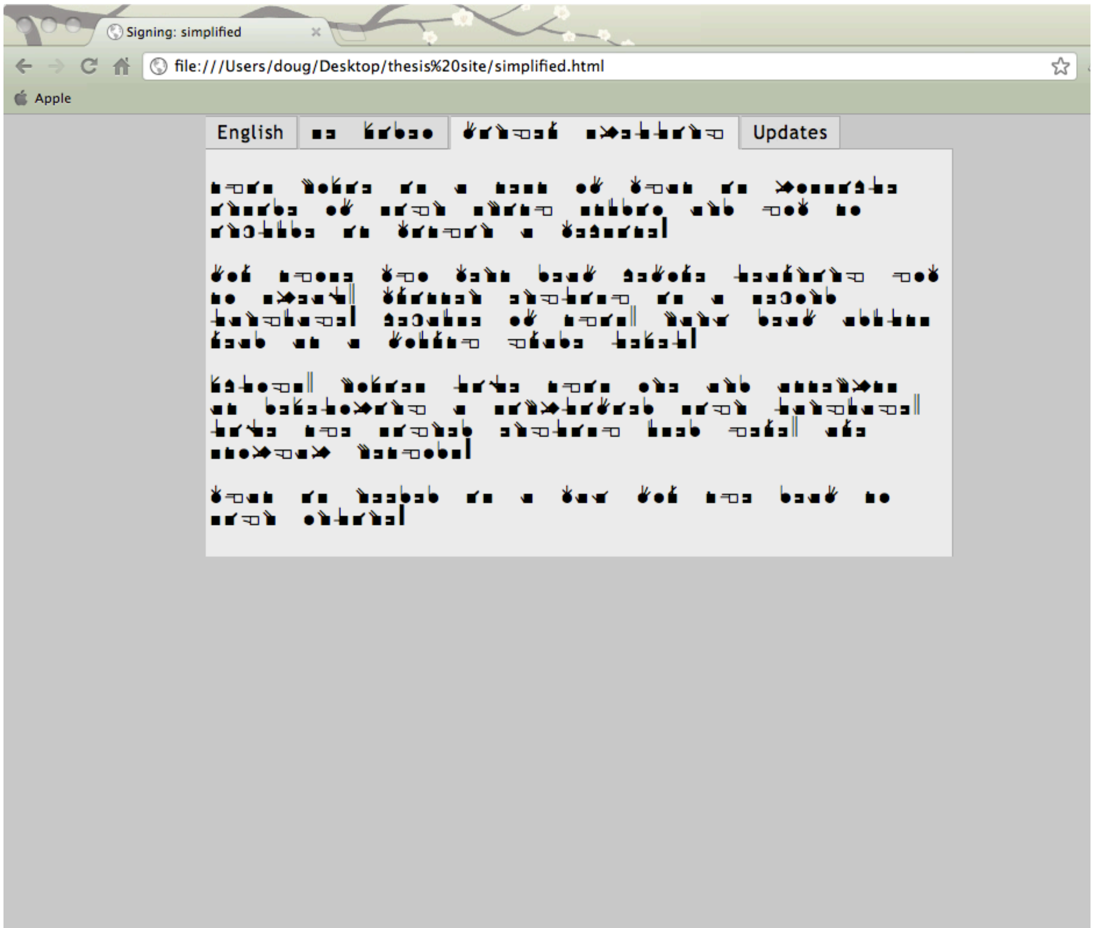

Figure 5: Reduced text in fingerspelled English using Sutton US font.

References

Allen, T. (1986). Patterns of academic achievement among hearing impaired students: 1974 and 1983. In A. Schildroth and M. Karchner (Ed.). *Deaf Children in America.* (pp. 161-206).

Auer, E., & Bernstein, L. (2008). Estimating when and how words are acquired: A natural experiment on the development of the mental lexicon. In *Journal of speech, language and hearing research, 51 No. 3* (pp. 750-8).

Banjar, S. (2008). Methods of translation presentation [Web]. Retrieved from http://www.slideshare.net/dr.shadiabanjar/methods-of-translation-presentation

Brentari, D. (2001). Borrowed elements in sign language: A window on word formation. In D. Brentari (Ed.), *Foreign vocabulary in sign languages: Cross-linguistic investigation of word formation (pp. ix-xx)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brentari, D., & Padden, C. A. (2001). Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins. In D. Brentari (Ed.), *Foreign vocabulary in sign languages: Cross-linguistic investigation of word formation (pp. ix-xx)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cavender, A. C., Otero, D.S., Bighamy, J. P., & Ladner, R. E. (2010). ASL-STEM forum: Enabling Sign Language to grow through online collaboration. *In CHI '10 Proceedings of the 28th international conference on Human factors in computing systems*.

Chadwick-Dias, A., McNulty, M., & Tullis, T. (2003). Web usability and age: How design changes can improve performance. Paper presented at the CUU, Vancouver, BC, Canada

Clark, J. (2003). Re: Accessible web design for the deaf: NOT Retrieved from http://lists.w3.org/Archives/Public/w3c-wai-ig/2003JulSep/0224.html

Crandall, K. (1978). Reading and writing skills and the deaf student. In *Volta Review, 80* (pp. 319-332).

Fisher, W., Petrie, H. L., Weber, G. (2005). Personalization, interaction, and navigation in rich multimedia documents for print-disabled users. *IBM systems journal, 44, No. 1* (629-35).

Ghadi, Alireza S. (2010). *All New Theories And Concepts About Translation In New Century* Retrieved from http://www.articlesbase.com/

Hanke, T. (2004), HamNoSys - representing sign language data in language resources and language processing contexts. In: Streiter, Oliver, Vettori, Chiara (eds): *LREC 2004, Workshop proceedings : Representation and processing of sign languages*. Paris : ELRA, 2004 (pp. 1-6).

Hanke, T., G. Langer and M. Metzger, (2001). Encoding non-manual aspects of sign language. In: T. Hanke(ed.), *Interface definitions*. ViSiCAST Deliverable D5-1 Retrieved from http://www.visicast.cmp.uea.ac.uk/Papers/ViSiCAST_D5-1v017rev2.pdf

Hickok, G., Bellugi, U., & Klima, E. S. (2001). Sign language in the brain. Scientific American, 284, No. 6, (pp.58-65).

Huenerfauth, M. (2004). "A multi-path architecture for machine translation of English text into American Sign

Language animation." In: *The proceedings of the Student Workshop at the Human Language Technology conference / North American chapter of the Association for Computational Linguistics annual meeting* (HLT-NAACL 2004), Boston, MA, USA.

Huenerfauth, M. (2005). Representing coordination and non-coordination in an American Sign Language animation. *Assets'05,*

Huenerfauth, M., Zhao, L., Gu. E., & Allbreck, J. (2008). Evaluation of American Sign Language generation by native ASL signers. *ACM Transactions on Access Computing, Vol. 1, No. 1, Article* 3.

Huenerfauth, M. (2008). Evaluation of a psycholinguistically motivated timing model for animations of American Sign Language. *Assets'08,*

Huenerfauth, M., Lu, P., & Rosenberg, A. (2011). Evaluating importance of facial expression in American Sign Language and Pidgin Signed English animations. *ASSETS '11,* October 24 - 26, 2011. Dundee, Scotland, UK.

iCommunicator. Accessed March 2012 http://www.icommunicator.com/

Jobs, S. (2010). Thoughts on Flash. Retrieved from http://www.apple.com/hotnews/thoughts-on-flash/

Leigh, G., Newall, J.P., & Newall, A.T. (2010). Newborn screening and earlier intervention with deaf children: Issues for the developing world. In M. Marschark & P. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education, vol 2* (pp. 345-359). New York: Oxford University Press.

Liddell, S. (2003/2004). *Grammar, gesture and meaning in American Sign Language*. New York: Cambridge University Press.

Lu, Pengfei, Huenerfauth, M. (2011). Collecting an American Sign Language corpus through the participation of native signers.

Luft, P. (2011). Promoting resilience in deaf children. In D. Zand & K. Pierce (Eds.), *Resilience in Deaf Children: Adaptation through Emerging Adulthood*. Springer

Kennaway, J. R., & Glauert, J. R. W. (2007). Providing signed content on the internet by synthesized animation. *ACM Transactions on Computer-Human Interaction*, Vol. 14, No. 3.

Kingston, J. (1999). Extrapolating from spoken to signed prosody via laboratory phonology. *Language and Speech* 42.00238309 251-81. ProQuest Central.

Mayer, R. E., Heiser, J. & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding." *Journal of Educational Psychology* Vol. 93, No. 1 pp. 187-198.

Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies, 4(2),* (pp. 138-163). ProQuest

Nespor, M., & Sandler, W. (1999). Prosody in Israeli Sign Language. *Language and Speech* 42.00238309 143-146. ProQuest Central.

Ooyen, B, et al. (2003). Prosodic structure and syntactic Acquisition: The case of the head–direction parameter. *Developmental Science 6.2* 211-220. Academic Search Premier.

Pinker, S. (1994). *The language instinct: The new science of language and mind.* London: Penguin.

Ramscar, M. & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science 31*, (pp. 927-960).

Sinclair, J. 2005. Corpus and text - basic principles. In M. Wynne (Ed.). *Developing linguistic corpora: a guide to good practice*. Oxford: Oxbow Books: 1-16. Retrieved from http://ota.ahds.ac.uk/documents/creating/dlc/chapter1.htm

Singtel Inc. (March 20, 2012) http://www.signtelinc.com/

Spencer, P. & Marschark, M. (2010). *Evidence-based practice in educating deaf and hard-of-hearing students*. New York: Oxford University Press.

Stanovich, K. (1980). Toward an interactive compensatory model of individual differences in the development of reading fluency. In *Reading Research Quarterly 16* (pp. 32-71).

Summers, K, & Summers, M. (2005).“Reading and navigational strategies of Web users with lower literacy skills.” ASIS&T Proceedings, Oct 2005. Charlotte, NC.

Retrieved from http://iat.ubalt.edu/summers/papers/Summers_ASIST2005.pdf

Sutton, V & Slevinski, S. (2011) Encoding the graphemes of SignWriting. Internet Draft. Retrieved from http://tools.ietf.org/html/draft-slevinski-iswa-2010-00.html

Stewart, L. G. (1998). The essays of Larry G. Stewart. In T. Bertling (Ed.), *American Sign Language: Shattering the myth (pp. 12-47)*. Wilsonville, ON: Kodiak Media Group.

Tinwell, A., Grimshaw, M., Abdel Nabi, D., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. Computers in Human Behavior 27(2), 741-749.

Traxler, C. (2000). The Stanford Achievement Test, 9th Edition: National norming and performance standards for deaf and hard of hearing students. *Journal of Deaf Studies and Deaf Education*, 5, (pp. 337-348).

United States Access Board. (2/5/2012). Television Decoder Circuitry Act of 1990. Retrieved from http://www.access-board.gov/sec508/guide/1194.24-decoderact.htm

W3C. (2011). Cascading style sheets level 1 Revision 1 (CSS 2.1) specification. Bos, B., Çelik, T.,Hickson, I. & Lie H. W. (Eds.) Retrieved from: http://www.w3.org/TR/CSS21/

W3C. (4/14/2011). HTML to platform accessibility APIs implementation guide. Faulkner, S. & Shelly, C. (Eds.) Retrieved from: http://www.w3.org/TR/2011/WD-html-aapi-20110414/

Watanabe, W., Candido, A., Amâncio, M., de Oliveira, M. Pardo, T., Fortes, R. & Aluísio, S. (2010) Adapting web content for low-literacy readers by using lexical elaboration and named entities labeling. *W4A2010 – Technical*. Retrieved from ACM 978-1-4503-0045

Wilcox, S. (2004) Struggling for a voice: An interactionist view of language and literacy in deaf education. In B. J. Brueggemann (Ed.), *Literacy and deaf people: cultural and contextual perspectives (pp. 157-191)*. Washington, DC: Gallaudet University Press.