

Improving Accessibility to Web Documents for the Aurally Challenged with Sign Language Animation

Jin-Woo Chung
Computer Science Department
KAIST
Daejeon, South Korea
jwchung@nlp.kaist.ac.kr

Ho-Joon Lee
Computer Science Department
KAIST
Daejeon, South Korea
hojoon@nlp.kaist.ac.kr

Jong C. Park
Computer Science Department
KAIST
Daejeon, South Korea
park@nlp.kaist.ac.kr

ABSTRACT

In this paper, we describe how to improve accessibility for the aurally challenged in a web environment, focusing on utilizing a signing avatar for web pages. Many systems were previously proposed to make a web environment more accessible for the deaf people by providing signed expressions, i.e. translating written text into sign language animations and presenting them in a proper way, based on the observation that deaf users normally have much difficulty understanding text-based information as well as audio contents. We analyze the strengths and weaknesses of these systems with respect to discussed design criteria, and propose a system that presents a signing avatar for web page documents via a mobile device, which is expected to overcome the shortcomings of the previous systems and to improve the accessibility of deaf users to textual contents in a web environment. The proposed system has three main parts based on a client-server architecture: 1) a client that executes a web browser and transmits selected text to the server, 2) a server that takes text as input and translates it into signed expressions through a sign language generation module, and 3) a mobile device that displays signing animation transmitted from the server by streaming. We also present some linguistic issues raised by the difference between Korean and Korean Sign Language. To the best of our knowledge, this is the first approach to the use of a mobile device for web document access by the aurally challenged people. We discuss implications of our study and future directions.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User interfaces – *interaction styles, natural language*

General Terms

Design, Human Factors

Keywords

Web accessibility, Deaf signers, Sign language

1. INTRODUCTION

Many contents in a web environment are usually presented in a written form, i.e. through text, which makes it extremely difficult for the deaf people to access and understand information, contrary to the common (and mistaken) belief that deafness and difficulty of hearing are just physical symptoms and that the aurally challenged people are thus able to read and understand textual contents without any problem. According to many researches on the reading ability of hearing-impaired users [1-3], they often show much poor reading performance compared to hearing users, particularly if they are prelingually deaf. For deaf signers, written text is nothing more than that written in a second (and foreign) language, making its understanding quite difficult in a way similar to that hearing users find it hard to understand expressions in a foreign spoken language. Deaf signers also tend to rely on a relatively smaller amount of working memory for verbal information than hearing users, which leads to much trouble performing the task of information retrieval and comprehension in a mostly text-based web environment [4-6]. Moreover, most studies for web accessibility are primarily focused on providing text-based equivalents for audio contents, e.g., closed caption subtitles, to compensate for their inability to hear, rather than providing signed information directly.

To overcome these difficulties, there have been a number of studies on generating signed expressions, including synthesizing natural signing animation [7-9], suggesting a proper description for representing sign language expressions [10-12], and converting written text into equivalent sign language notations or videos [9, 10, 13]. In a similar line, the present paper describes a way of providing signed information in conjunction with the contents of web pages, reflecting the deaf users' preference and cognitive characteristics and the result of previous studies. First, we analyze the strengths and weaknesses of the previously proposed systems from various aspects with respect to four design criteria. Then we present a prototype of a new system that incorporates their good features but that overcomes their shortcomings. The use of a mobile device for web document access by the aurally challenged people as proposed in this paper is the first such approach, to the best of our knowledge.

The rest of the paper is organized as follows: Section 2 describes the design goal to improve web accessibility of deaf users. Section 3 discusses the related work in terms of the design goal presented in Section 2. Section 4 proposes a new system that incorporates the discussion in Sections 2 and 3. Section 5

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WTMS'11, May 25-27, 2011, Sogndal, Norway.

Copyright © 2011 ACM 978-1-4503-0148-0/11/05...\$10.00.

discusses implications of our study. Section 6 concludes the paper and presents future research directions.

2. DESIGN GOAL

We present four design criteria to provide deaf users with more improved usability and accessibility to the web, based on the results of experiments conducted in previous studies [4-6].

Portability. Signed expressions on a web page should be delivered on a screen without having to modify the structure and layout of the original web page, so that deaf users can easily invoke signed expressions in any existing text-based websites by just installing simple software, e.g., through a plug-in. It would reduce the cost of a web designer to rebuild a whole new website for deaf users. It would also provide an equal opportunity for deaf users as well as non-challenged people to access information in currently available web services.

Location adjustability. Signed expressions that are presented should be physically (and conceptually) close to the corresponding text as much as possible. This is based on the understanding that deaf signers show better performance in searching and navigating the web when signed expressions are presented closer to the corresponding text [4]. It is also reported that the spatial contiguity between sign language videos and the corresponding text reduces the burden of deaf users on the use of their verbal working memory, which is normally lower than that of hearing users. Therefore, it helps much to make the distance as close as possible between signed and textual information to provide a better understanding.

Synchronizability. Signed expressions should be provided without blocking the users' view of the original textual contents; that is, deaf users should be able to access not only the signed expressions but also the corresponding original text. The experiment conducted in [1, 4] indicates that deaf users actually want the original textual information together with the signed expression and that they do not fully rely on the signed expression alone in most cases. In fact, not all deaf or hard-of-hearing users have difficulty reading and understanding textual information, especially if they are not prelingually deaf. They might thus want to see signed expressions selectively, i.e. some word or phrase but not a full sentence, depending on the degree of the challenges that they face. In addition, providing textual and signed information simultaneously would make it effective to improve the reading comprehension level of the deaf users since they can easily obtain feedback by comparing two different kinds of expressions at the same time.

Scalability. It is better to generate signed expressions by directly translating written text into signing animation, than to play pre-recorded videos. In particular, if we are to allow deaf users to see sign language expressions for all textual contents in any existing web pages, it would be infeasible to pre-record and keep track of these exponentially growing web contents. Also, while it is certainly obvious that sign language expressions delivered by a human signer would be much more expressive and natural than computationally animated ones because sign language expressions normally contain various subtle movements and facial expressions, it would be quite difficult to meet this level of performance by computer animation, much as synthesized voice tend to sound unnatural. However, providing pre-recorded videos has many

problems of its own from a practical perspective as well, including those of cost, consistency, and management, which do not occur on this scale with signing animation [1]. Therefore, we believe that it is reasonable and practical to use a signing avatar for an extensive and scalable service.

We presented four design criteria to improve the web accessibility and usability for deaf users. However, they are certainly not exhaustive. In particular, there may be others that assess the situation more properly. For example, if we must offer web services only for the prelingually deaf users, it might be better to develop a fully optimized new website that provides mostly video-based contents even if it is fairly costly. However, our goal is to help deaf users with a diverse range of disabilities to exploit information and services already available on the web as much as possible. In this regard, we expect that our design criteria would be appropriate for accomplishing the goal. In Section 3, we discuss the strengths and weaknesses of previous systems with respect to these criteria.

3. RELATED WORK

There have been many studies that dealt with sign language from a computational perspective of providing deaf users with a better access to textual information. Several research projects have focused on improving the quality of a signing avatar as synthesized by computer animation [7-9]. Others have proposed automatic translation of written text into sign language notations or signing animation [10-12]. The major part of the work in the translation task is to reflect properly on the linguistic difference between written and sign languages and to deliver manual and non-manual expressions as simultaneously as possible. In this paper, we focus on how to present and interact with a signing avatar on web pages based on the design criteria as described in Section 2. There are a number of studies on this branch of work [1, 4-6, 14-18]. They can be categorized according to how and where signed expressions are presented: on a separate layer, within the body of a page, or within a completely optimized page.

Some researches such as [6, 16, 17] propose to present signed expressions on a separate layer as in a pop-up window. This approach is intuitively attractive in that it is readily applicable to most existing websites, since a web designer does not need to alter the layout and structure of web pages to find extra space for the signed expressions. Moreover, it is possible to allow the user or web designer to adjust the size of a window for a better view and to change its position to avoid blocking the content of a page. However, it also has some disadvantages: First, although a user can move around the window as desired, the pop-up window may still obstruct the view of original text when it appears first. In addition, since the position of contents on a page could vary depending on its structure, the user may also have to move the window repeatedly when migrating to another page, which inconveniences the web navigation further. Second, it might be mentally distracting and disturbing to manually click on a button to open or close the pop-up window, especially when it should be frequently performed. The experimental result of [16] indicates that deaf users much preferred the option to automatically close the window. We can also witness this consideration from Google Translate, one of the free online translation services, which allows a user to check the original sentence by just putting a cursor on

the corresponding translated sentence, not by clicking on any button or text to open or close the window (see Figure 1).

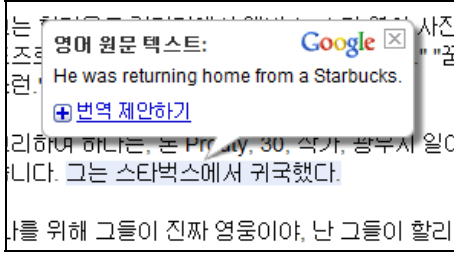


Figure 1. Google Translate

Other studies propose a design where a signing avatar appears within a subframe as specified by the web designer [18-20]. This approach can overcome the shortcoming of a pop-up window in that signed expressions never block any other parts of a page. Users also do not need to click on a button all the time to open or close the window. One problem with this design, however, is that the method is not universally applicable because it requires the structure of a page to be fixed, which does not meet our first design criterion. Moreover, in some cases, if the frame in which signing animation is displayed is quite distant from its corresponding text, it would be difficult to read text and signing at the same time, which is against our second criterion. It would thus be important in this design to make the distance between signed and textual information as close as possible. Nonetheless, if the purpose of the website is to provide specialized contents solely for deaf users, this approach would probably be most acceptable.

There are also some studies on improving the web accessibility of deaf users in a very different way [6, 15]. They propose video-based hyperlinking as a new navigation mechanism that is completely optimized for deaf users, especially for the prelingually deaf. This design is innovative in that it provides a much improved access for the people who understand sign language but cannot read text at all. However, the applicability and portability of the system are fairly limited and the web designer should build an entirely new website to make this service available on the web, which is also against our design criteria.

4. PROPOSED SYSTEM

We propose a system that presents a signing avatar in a mobile device. This approach is designed to avoid the shortcomings of the previous systems and to meet the four criteria described in Section 2. Figure 2 illustrates the schematic flow of the system.

4.1 System Overview

The system has three main parts: a client, server, mobile device. The simple description of how each part interacts is as follows: First, the user executes a web browser in the client and invokes its plug-in to communicate with the server and mobile device. Once the plug-in is set to start, it keeps checking for the user's mouse behavior and captures the drag over the sentence. Then it transmits the selected sentence to the server, which in turn translates the sentence into corresponding signed expressions by its sign language generation module. A more detailed description of this module is presented in Section 4.2. After signing animation is generated in the server, it is played on the screen of the mobile device by streaming (see Figure 3).

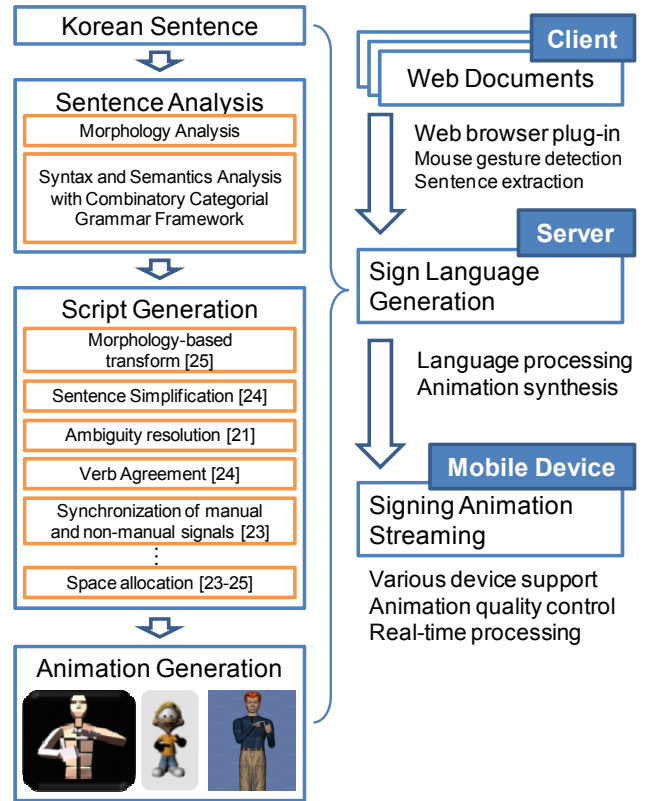


Figure 2. System architecture

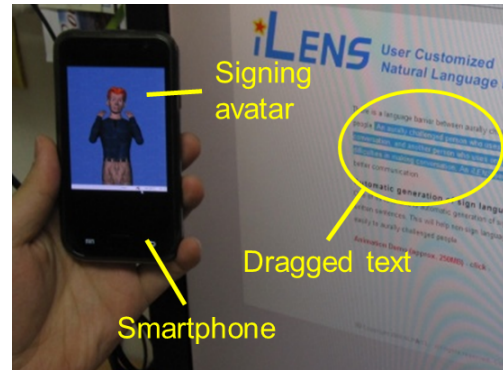


Figure 3. System demonstration

4.2 Sign Language Generation

As mentioned earlier, sign and spoken languages are significantly different in terms of their linguistic characteristics and phenomena. These differences are also found between Korean and Korean Sign Language (KSL).¹ Korean is an agglutinative

¹ There is another form of sign language in Korea, called 'Signed Korean'. It is generally known that its grammar follows that of Korean, but Signed Korean and KSL have been used without a clear distinction and their differences are not much revealed [21, 38]. We therefore assume in this paper that the term 'Korean Sign Language' represents both Signed Korean and KSL.

language, whose words and grammatical relations are formed and represented by joining morphemes together. In contrast, KSL is similar to an isolating language such as Chinese and Vietnamese where ‘morphemes’ are considered syntactically full-fledged expressions and their variations do not occur frequently [30]. One of the most significant differences is that KSL does not have case markers and ending markers. Most of such grammatical relations are determined by the order of morphemes and words or by separate non-manual signals produced along with manual signals. Therefore, such differences should be carefully considered when translating expressions of spoken language into those of sign language. It might otherwise lead to unnatural expressions which may not be fully understandable to deaf users. To address this problem, we have already studied various methods of generating appropriate expressions in KSL from ones in Korean [21-27], and have their results incorporated into the proposed system as the sign language generation module. This module consists of three parts: sentence analysis, script generation, and animation generation. The first part analyzes a syntactic and semantic structure of a given sentence in Korean with a grammar formalism called Combinatory Categorical Grammar. The script generation part produces an animation script for sign language animation using syntactic and semantic information obtained from the sentence analysis part. In this step, some linguistic phenomena are addressed to bridge the gap between the two languages, which is a key contribution of this study. The animation generation part takes scripts and synthesizes animation in KSL.

4.2.1 Sentence Analysis

Combinatory Categorical Grammar. It is necessary to first analyze the syntactic and semantic structure of a given sentence in order to appropriately map two sentences in different languages. This task is performed within the framework of Combinatory Categorical Grammar (CCG), an efficiently parseable, yet linguistically expressive grammar formalism. It has a transparent interface between surface syntax and underlying semantic representation and thus can model difficult linguistic phenomena, such as coordination, long distance extraction and topicalization, without further stipulation [28, 29]. This grammar formalism has been widely used for computational modeling of natural language sentences in Korean [31-33] as well as in English [34, 35]. The process of deriving syntax and semantics from a written sentence using CCG is exemplified in Figure 4. Each word or morpheme is assigned one of the possible categories, and two adjacent words are combined into one by combination rules (Table 1) in every derivation step, eventually producing the only one representation corresponding to a whole sentence. For further detailed explanation and computational motivation for this theory of grammar, the reader is directed to [28, 29].

Anna	met	and	married	Manny
<i>np</i>	<i>(s\vp)/np</i>	<i>conj</i>	<i>(s\vp)/np</i>	<i>np</i>
:anna'	:meet'	:and'	:marry'	:manny'
<i>(s\vp)/np</i>				< ϕ^n >
:λx. λy.and'(meet'xy)(marry'xy)				>
<i>s\vp</i>				
:λy.and'(meet'manny'y)(marry'manny'y)				<
s:and'(meet'manny'anna')(marry'manny'anna')				

Figure 4. CCG derivation

Table 1. CCG combinatory rules

Combinatory Rule	Name	Combinator
$X/Y \ Y \rightarrow X$	Forward Application	$>$
$Y \ X/Y \rightarrow X$	Backward Application	$<$
$X/Y \ Y/Z \rightarrow X/Z$	Forward Composition	$>B$
$Y/Z \ X/Y \rightarrow X/Z$	Backward Composition	$<B$
$X \rightarrow T/(T/X)$	Forward Raising	$>T$
$X \rightarrow T/(T/X)$	Backward Raising	$<T$
$X \text{ conj } X \rightarrow X$	Coordination	$<\$>^n$

CCG lexicon. In this paper, we include the necessary information for generating sign language expressions into the CCG lexicon. The template for such information consisting of CCG categories is shown in Figure 5. *np*, *cl*, and *s* to the left of colons represent a noun phrase, clause, and sentence, respectively, which are basic syntactic units used in conventional CCG categories. *Case* is the case of a noun or pronoun indicated by case markers in Korean such as *nom* (nominative), *acc* (accusative), and *dat* (dative). *Tag* represents a semantic role of nouns such as an agent and theme, also indicated by case markers. *Name* is the name of words in sign language expressions. *ClType* indicates the type of clause. Its value can be *adnc* (adnominal clause), *qc* (quotative clause), *adv* (adverbial clause), *coorc* (coordinate clause), and so on. *NFeature* is a feature of a noun phrase in sign language such as *classifier* (the type of classifier – e.g., gender), *numhands* (the number of hands), and *nms* (non-manual signals). *VFeature* is a feature of a verb including *verbType* (the type of verb – simple, agreement, and spatial verbs [37]), *hasMob* (indication for mobility of a verb), and *hasPass* (indication for verb passiveness). *ClFeature* is a feature of a clause – tense and politeness. *senType* represents the type of sentence such as *dec* (declarative) and *imp* (imperative). For instance, the transitive verb ‘사랑하다’[*sa-lang-ha-ta*] (meaning ‘love’) has the category as shown in Figure 6. The variables ‘A’ and ‘B’ here are slots to be filled with two noun phrases (a subject and object, respectively) both of which are taken from the left side of ‘love’ during the derivation process. The derivation result for the sentence “I love you.” is shown in Figure 7.

```
np(Case):Name@[NFeature]
cl:ClType&[[Tag=[Name@[VFeature]], Tag=[Name@[Feature]],
... ]# [ClFeature]]
s:[ClType&[[Tag=[Name@[VFeature]], ... ]# [ClFeature]],
... ]$[SenType]
```

Figure 5. The template for CCG categories

```
cl:[action=[love@[verbType=simple, handPos=no, handMob=no,
hasPass=no, handOrt=down, numHands=two]], theme=A,
agent=B]# [tense=present, polite=no]]np(nom):B \np(acc):A
```

Figure 6. CCG category for the verb ‘love’

```
s:[mc&cl:[[action=[love@[verbType=simple, handPos=no, handMob
=no, hasPass=no, handOrt=down, numHands=two]], theme=[she@
[classifier=female, hasCl=yes, numHands=one]], agent=[i@[classifier
=male, hasCl=no, numhands=one]]]#[tense= present, polite=no]]]
$[dec]]
```

Figure 7. Derivation result for “I love you.”

4.2.2 Script Generation

After the sentence is analyzed and converted into a semantic representation, it is passed to the script generation part in which the result go through some pre-processing tasks for mapping into sign language expressions reflecting its linguistic characteristics. They are very important steps to improve understanding and naturalness of the resulting expression.

Morphology-based transformation. As mentioned in the beginning of Section 4.2, Korean and KSL belong to different kinds of languages – agglutinative and isolating, respectively. Thus, some grammatical particles in Korean should be detached from words so that it is expressed as a separate unit of words [25]. For example, past tense in KSL should be explicitly expressed as a hand-motion meaning ‘end’ after the expression of the main verb while tense in Korean is usually indicated by attached endings of a verb. Certain postpositions in Korean also need to find their counterparts in sign language [25]. For proper nouns and unknown words, we deal with them by expressing their exact finger spelling.

Restructuring complex sentences. It is necessary to address the problem of the structural difference between Korean and KSL. A complex sentence in Korean has a structure where the matrix clause contains one or more non-matrix clauses with particular endings. In contrast, a sentence in KSL delivers the meaning with a combination of the related signs that do not carry the corresponding notions of case postpositions and endings in Korean. It means that the structure of sentences in sign language is simpler than that of sentences in written language. These syntactic differences are some of the important factors that prevent deaf people from completely understanding complex sentences in Korean [30]. In order to avoid confusion and to provide better understanding, it is important to split one complex sentence into a couple of simple sentences and to make closer the distance between a predicate and its subject to compensate for a short working verbal memory span [24]. For instance, the Korean sentence in the example (1) would be converted to the signed expression “I + He’ + come’ + happy’”. However, since the adjective ‘happy’ appears somewhat distant from its subject ‘I’ in the expression, it may lead to misunderstanding that “He and I came back, so someone is happy.” This happens because case markers and connective endings that indicate grammatical roles in Korean do not appear in the KSL expression; instead, that information is delivered by the order of words or by context. Therefore, it needs to be converted to “He’ + come back’ + I’ + happy’” by putting the predicate and its subject closer to each other. It has the same as taking the nested clause out of the sentence.

- (1) 나는 그가 돌아와서 기쁘다.
 I(nom)_i He(nom)_j came back_j happy_i
na-nun ku-ka tol-a-wa-se ki-ppu-ta.
 “I am happy because he came back.”

- I’ + He’ + come’ + happy’ (possibly misunderstood)
 → He’ + come’ + I’ + happy’ (clearly understood)

Verb agreement. Verbs in KSL are divided into three groups as in American Sign Language: plain, agreement, and spatial verbs [36, 37]. In the case of agreement verbs, the form and direction of hand movement change according to their arguments. The movement usually begins from the locus of a subject argument and ends at the locus of an object argument [37]. A pattern of movement is however slightly different depending on what kind of agreement verb is used. For example, the agreement verb ‘teach’ in KSL is performed by pointing out only its object argument. In order to reflect these kinds of difference, we use 95 agreement verbs and their categorization as previously studied in [37]. In a way similar to agreement verbs, a spatial verb allows the signer to specify where things are or how they are moved. It usually involves a change in the direction of movement depending on locative adverbs in KSL [24]. Consider the example (2).

- (2) 그는 학교에서 집으로 걸어갔다.
 He(nom) School(from) Home(to) Walked
ku-nun hak-kyo-ey-se cip-u-lo kel-e-kass-ta.
 “He walked home from school.”
 →He’ + School’ + (Loc1)’ + Home’ + (Loc2)’ +
 Walk_(Loc1→Loc2)’

The direction of hand movement in the verb ‘walk’ (Loc1→Loc2) is determined by the relative position of a source (Loc1) and destination (Loc2) indicated from locative adverbs. Verbs that involve mobility such as ‘move’, ‘go’, and ‘put’ belong to this group of verbs. This information should be also included in the CCG categories of this kind of verbs. Most locative adverbs and their semantic role such as ‘from’ or ‘to’ can be captured by postpositional participles including ‘-에서’ and ‘-로’ as shown in the example (2). More detailed discussion on dealing with spatial verbs of KSL is found in [24].

Space allocation. As mentioned earlier, deaf signers assign a certain part of signing space to entities when they are an argument of agreement or spatial verbs, but they differ in their use of space; while agreement verbs use space syntactically, spatial verbs use space topographically [36]. In KSL, deaf signers also utilize space when comparing two entities. The decision about where to place them in space depends on the geographical or conceptual relationship between them [39]. For example, ‘mountain’ is placed above ‘river’, and ‘good’ is placed to the left of ‘evil’. However, a systematic approach to assigning space to those two entities has not been sufficiently studied yet. We thus leave this problem for future work, and simply assign signing space only for agreement and spatial verbs. For agreements verbs, the left and right of space are assigned to subject and object arguments, respectively. For spatial verbs, the left and right side are assigned to the positions of source and destination, respectively.

After finishing all these pre-processing tasks, the resulting semantic representation is converted into animation scripts, as shown in Figure 8. The script has an XML-based format, which is an input form of the animation-generating tool we employ in this study. Figure 8 illustrates a script for signing animation corresponding to the word ‘father’ in KSL. Each single element (or word) in the script starts with the tag <SingleActionData> and ends with </SingleActionData>. A sentence or phrase consists of

a sequence of such single elements. The resulting script is then passed to the animation generation part.

```
<SingleActionData word="father.ges" desc="" path="" m="1" t=".25" h="0"
h0="0" filenames="R[[L;H3;22^H^L~L~L~U02;;l_BA_site;
neutral_site;;0,0,0;.051,143,0;-.727,-.155,-.669,2.973;;M3;H4;L+;;l_BA_
site;neutral_site;;0,0,0;.143,113,0;-.973,-.086,213,2.434;;]] duration=" .667"
basename="" regions="Universal">
<auxiliaryData>
<VoiceData label="" markers="" />
<Gaze headTrack="point" eyesTrack="point" trackPoint=".041,1.302,.296"
/>
<ExpressionData label="Right Smile" filename="smile_right" />
<SubActionData></SubActionData>
</auxiliaryData>
</SingleActionData>
```

Figure 8. Animation script

4.2.3 Animation Generation

In order to synthesize signing animation efficiently, we use SignSmith Studio², which is commercial software for generating sign language animations from scripts. The software synthesizes corresponding signing animation. It is then transmitted to the mobile device by streaming, and is played on its screen in real time, which means that the users do not need to install the synthesizing software in their mobile device separately. The system also allows a user to pause, stop, and replay the video by pressing buttons displayed on the screen. Figures 9 and 10 show the synthesized results for the sentences “I am happy because he came back.” and “He walked home from school.” As discussed in Section 4.2.2, the signed expressions in Figures 9 and 10 consist of sequences of signs, “he’ + come back’ + I’ + happy” and “He’ + School’ + (Loc1)’ + Home’ + (Loc2)’ + Walk_(Loc1→Loc2)”, respectively.



Figure 9. Synthesized animation 1



Figure 10. Synthesized animation 2

5. DISCUSSION AND FUTURE WORK

Delivering signing animation via a mobile device has several advantages, in particular satisfying the four criteria presented in Section 2: First, the content and structure of a web page do not need to be changed by a web designer because signing animation is displayed on the user’ separate device and the appearance of a web page is thus not affected at all. All the user needs to do is just

to open a web browser, to execute a plug-in, and to select the sentence he or she wants to translate into signed expressions. Second, the user can see textual information and signed expressions at a close distance by adjusting the position of a mobile device. For a clear view, the user may also enlarge the view of the screen much more easily; just bringing the mobile device close to the user may solve a potential problem, which is not possible in the previous systems. It is true that the size of a pop-up window can also be enlarged, but it may also inadvertently block some portion of contents on the screen in the process. Third, the mobile device does not interrupt the view of textual contents on a web page in any manner. Last, our system has a translation module that automatically converts written text into sign language animation by language processing and an animation-generating tool, obviating the need for recording videos and motion capture in advance. Table 2 briefly summarizes analysis results of the previous approaches to providing signed contents in a web page, with respect to the discussed criteria.

Table 2. Analysis result with the design criteria

Where to place Criteria	Separate layer	Body of a page	Optimized page	Mobile device
Portability	○	△	×	○
Location adjustability	○	×	△	○
Synchronizability	△	○	△	○
Scalability	×	×	×	○

However, the usefulness of each system we have discussed in this paper may be evaluated differently according to what design criteria are pursued. For instance, if we focus only on improving the quality of sign language expressions for deaf users who cannot accept textual information at all, a completely optimized web page and sign language videos pre-recorded by a native deaf signer would probably be most acceptable, even though they involve high costs and are not universally applicable. Nonetheless, we believe that the criteria we have presented will serve as

² <http://www.vcom3d.com>

important factors for evaluating accessibility in a web environment.

There are still a lot of issues to be resolved in improving naturalness (or grammaticality) of the resulting signed expressions. For example, non-manual signals play a very important role in communication between deaf signers, but in KSL they have not been much studied yet. Our proposed system considers only non-manual signals that are used when expressing emotion, e.g. 'happy'. In practice however, they are used for various purposes, and can be actually produced across multiple words. There is also a problem with the size of vocabulary. The vocabulary of sign language is much smaller than that of Korean. It is therefore necessary to develop a method of mapping Korean words to KSL words appropriately. Additional characteristics of KSL should be further explored and reflected in sign language generation. Since these issues need further discussion and examination, we leave them for future work.

6. CONCLUSION

In this paper, we investigated the method of improving the accessibility of deaf users to textual information in a web environment. Four design criteria were presented for enhancing convenience and applicability. We analyzed previous systems with respect to these criteria and proposed a new system that displays a signing avatar on the screen of a mobile device, which overcomes their weaknesses. The method of mapping Korean to KSL based on our previous studies was also presented. We showed that there are many linguistic issues to be addressed to produce more natural signed expressions.

In recent years, mobile devices such as smartphones and e-book readers have become widespread and lots of application software that makes use of their good functionalities has been developed. Using these functionalities of mobile devices, we can provide additional useful services. For instance, our system can be upgraded not only to generate signed expressions from text, but also to record and store them elsewhere to replay afterwards. This has the same effect as writing a memo by a hearing person. Moreover, if the OCR (optical character recognition) technique is employed to recognize the entire sentence or paragraph, deaf users may be able to obtain signed expressions directly not only from a web page on the screen but also from printed materials such as books just by taking their image and executing the OCR software. Through such efforts, we believe that our studies would overcome the information gap between hearing and deaf people and contribute to improving their quality of life one step further.

7. ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2010-0012527).

8. REFERENCES

- [1] Kennaway, J. R., Glauert, J. R. W., and Zwitserlood, I. 2007. Providing signed content on the Internet by synthesized animation. *ACM Trans. Comput.-Hum. Interact.* 14, 3, Article 15 (September 2007). DOI= <http://doi.acm.org/10.1145/1279700.1279705>
- [2] Huenerfauth, M. 2009. A linguistically motivated model for speed and pausing in animations of American Sign Language. *ACM Trans. Access. Comput.* 2, 2, Article 9 (June 2009). DOI= <http://doi.acm.org/10.1145/1530064.1530067>
- [3] Elliott, R., Glauert, J. R. W., Kennaway, J. R., and Marshall, I. 2000. The development of language processing support for the ViSiCAST project. The 4th International ACM Conference on Assistive technologies (Assets '00).
- [4] Fajardo, I., Parra, E., and Canas, J. J. 2010. Do sign language videos improve web navigation for deaf signer users? *Journal of Deaf Studies and Deaf Education*. 15, 242-262.
- [5] Fajardo, I., Abascal, J., and Canas, J. J. 2004. The role of working memory and long term memory in deaf users' hypertext navigation: review of guidelines for web accessibility. 8th ERCIM Workshop: User Interfaces For All (UI4All 2004).
- [6] Fels, D. I., Richards, J., Hardman, J., Soudian, S., and Silverman, C. 2004. American Sign Language of the web. ACM CHI Conference on Human Factors in Computing Systems (CHI 2004).
- [7] Elliot, R., Glauert, J., Kennaway, J., Marshall, I., and Safar, E. 2008. Linguistic modeling and language-processing technologies for avatar-based sign language presentation. *Univ. Acc. Inform. Soc.* 6, 4, 375-391.
- [8] Verlinden, M., Tijsseling, C., and Frowein, H. 2002. A signing avatar on the WWW. The 4th International Workshop on Gesture and Sign Language based Human-Computer Interaction (GW 2001).
- [9] Huenerfauth, M. 2006. *Generating American Sign Language classifier predicates for English-to-ASL machine translation*. PhD dissertation. University of Pennsylvania.
- [10] Marshall, I. and Safar, E. 2005. Grammar development for sign language avatar-based synthesis. The 3rd International Conference on Universal Access in Human-Computer Interaction (UAHCI 2005).
- [11] Jemni, M. and Elghoul, O. 2008. A system to make signs using collaborative approach. The 11th International Conference on Computers Helping People with Special Needs (ICCHP 2008).
- [12] Kennaway, R. 2004. Experience with and requirements for a gesture description language for synthetic animation. The 5th International Workshop on Gesture and Sign Language based Human-Computer Interaction (GW 2003).
- [13] Chiu, Y. -H., Wu, C. -H., Su, H. -Y., and Cheng, Chih. -Jen. 2007. Joint optimization of word alignment and epenthesis generation for Chinese to Taiwanese sign synthesis. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 1 (January 2007), 28-39. DOI= <http://dx.doi.org/10.1109/TPAMI.2007.15>
- [14] Verlinden, M., Tijsseling, C., and Frowein, H. 2002. A signing avatar on the WWW. The 4th International Workshop on Gesture and Sign Language based Human-Computer Interaction (GW 2001).
- [15] Kaibel, A., Grote, K., Knoerzer, K., Sieprath, H., and Kramer, F. 2006. Hypertext in sign language. The 9th ERCIM Workshop: User Interfaces for All (UI4All 2006).

- [16] Debevc, M., Kosec, P., Rotovnik, M., and Holzinger, A. 2009. Accessible multimodal web pages with sign language translations for deaf and hard of hearing users. The 20th International Workshop on Database and Expert Systems Application (DEXA 2009).
- [17] Yao, D., Qiu, Y., and Huang, H. 2009. Web-based Chinese Sign Language broadcasting system. The 6th International Cross-Disciplinary Conference on Web Accessibility (W4A 2009).
- [18] Othman, A., Ghoul, O. E., and Jemni, M. 2010. SportSign: a service to make sports news accessible to deaf persons in sign languages. The 12th International Conference on Computers Helping People with Special Needs (ICCHP 2010).
- [19] Straetz, K., Kaibel, A., Raithel, V., Specht, M., Grote, K., and Kramer, F. 2004. An e-Learning environment for deaf adults. The 8th ERCIM Workshop: User Interfaces for All (UI4All 2004).
- [20] Hilzensauer, M. 2006. *Information Technology for Deaf People*. Springer, Berlin/Heidelberg, 183-206.
- [21] Chung, J. -W. and Park, J. C. 2010. Detecting and resolving syntactic ambiguity for automatic Korean-Korean Sign Language translation. The 22nd Annual Conference on Human and Cognitive Language Technology (HCLT 2010).
- [22] Jung, S. 2010. *Automatic Sign Language Generation Reflecting the Relationship between Entities*. Master thesis. KAIST.
- [23] Jung, S., Chang, E., and Park, J. C. 2009. Synchronization of manual and non-manual signals in automatic generation of sign language expressions. The 21th Annual Conference on Human and Cognitive Language Technology (HCLT 2009).
- [24] Kim, S. 2008. *Computational Processing of Verb Agreement for Automatic Generation of Sign Language Animation*. Master thesis. KAIST.
- [25] Choi, J., Lee, H. -J., and Park, J. C. 2005. From text to sign language: exploiting the spatial and motioning dimension. The 19th Pacific Asia Conference on Language, Information and Computation (PACLIC 19).
- [26] Choi, J. 2004. *Automatic Translation of Korean into Korean Sign Language with Combinatory Categorical Grammar*. Master thesis. KAIST.
- [27] Lee, H., Kim, S., and Park, J. C. 2008. Syntactic construction of coordination in sign language generation. The 18th International Congress of Linguists (CIL 18).
- [28] Steedman, M. 1996. *Surface Structure and Interpretation*. The MIT Press.
- [29] Steedman, M. 2000. *The Syntactic Process*. The MIT Press.
- [30] Um M. S. 1996. *Analysis of Syntactic Characteristics of Korean Sign Language*. Master thesis. Daegu University.
- [31] Park, J. C. and Cho, H. 2000. Informed parsing for coordination with Combinatory Categorical Grammar. The 18th International Conference on Computational Linguistics (COLING 2000).
- [32] Lee, J. -B. 2001. *Computational Processing of Floating Quantifiers in Korean with Combinatory Categorical Grammar*. Master thesis. KAIST.
- [33] Cha, J., Lee, G. and Lee, J. 2002. Korean Combinatory Categorical Grammar and statistical parsing. *Computers and the Humanities*. 36, 431-453.
- [34] Bos, J., Clark, S., Steedman, M., Curran, J. R., and Hockenmaier, J. 2004. Wide-coverage semantic representations from a CCG parser. The 20th International Conference on Computational Linguistics (COLING 2004).
- [35] Clark, S., Steedman, M., and Curran, J. R. 2004. Object-extraction and question-parsing using CCG. Conference on Empirical Methods on Natural Language Processing (EMNLP 2004).
- [36] Padden, Carol. 1988. *Interaction of Morphology and Syntax in American Sign Language*. New York; Garland Publishing, University of California, San Diego.
- [37] Hong, S. -E. 2008. *Ein empirische Untersuchung zu Kongruenzverben in der Koreanischen Gebardensprache*. PhD dissertation. University of Hamburg.
- [38] Jhang, S. -E. 2009. Notes on Korean Sign Language. *The Handbook of East Asian Psycholinguistics*. 3, 361-376, Cambridge University Press.
- [39] Sutton-spence, R. and Woll, B. 1998. *The Linguistics of British Sign Language: An introduction*. NY: Cambridge University.