Asia Signopedia: An open information system of Asian sign languages

Kai-Hung Ken Wong\textsuperscript{a}, Gladys Tang\textsuperscript{a} and Ronald Chung\textsuperscript{b},\textsuperscript{*}
\textsuperscript{a}Centre for Sign Linguistics and Deaf Studies, The Chinese University of Hong Kong, Hong Kong
\textsuperscript{b}Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong

Abstract. To promote deaf awareness and natural sign language in Asia, we created an open platform named “Asia Signopedia”. The web page allowed both deaf and hearing people to input and access entries of different Asian sign languages and their dialects in either video or text mode. This paper describes how the data structure and user interface of the web page were designed. The distributive authoring scheme of the web page allowed the database to be input and corrected by those who used it.

Keywords: Sign language, database, Wikipedia, multi-modality

1. Introduction

Research on sign language as a linguistic system emerged in 1960s. Over time, more and more sign languages are being uncovered and the need to archive sign languages for different research or applied purposes becomes imminent. To date, the sign languages under study are mainly from the western countries and research into the Asian sign language varieties is very much lagging behind. This is due to a lack of research expertise in sign linguistics as well as a generally low level of deaf awareness in the societies. Yet, there is an increasing number of isolated reports on Hong Kong Sign Language, Taiwan Sign Language, Chinese Sign Language, Korean Sign Language, Japanese Sign Language, Thai Sign Language, Ho Chi Ming Sign Language, and Cambodian Sign Language. On the Asia Pacific rim, we found reports on Australian Sign Language (Auslan) and New Zealand Sign Language. Such developments have led to a call for using computer software technology to archive sign language data for both research and applied purposes like sign language teaching and interpretation. There are three main concerns about sign documentation: (1) how to document the information of individual signs in particular their phonological structure and categorical status, (2) how to capture the streams of signing collected from monologues or dialogues and transcribe them in a format that makes information transparent to the reader, and (3) how to compare the signs from different regions in Asia or different parts of the world in order to capture their relationship from the perspective of sign language typology. There have been reports suggesting that American Sign Language and French Sign Language, Hong Kong Sign Language and Shanghai Sign Language, British Sign Language and Australian Sign Language share a great deal in common linguistically [11,24]. In this paper, we report on a recent unprecedented effort to develop a web interface ‘Asia Signopedia’, an open information platform that encouraged communities of sign language users in the Asian region to document the sign language varieties themselves, in the hope of promoting use of natural sign language in various social domains such as deaf communication, deaf education, sign language interpretation, etc. The paper is organized as follows: in Section 2, we will briefly summarize the current state of sign language developments in Asia as well as the technology that has been adopted so far. Section 3 outlines some of the previous works relevant for the development of our project. Sections 4 and 5 outline the data structure design and the user interface design of the database respectively. Section 6 offers some descriptions as to how the web site had been developed in terms of visitor count. Section 7 highlights some feedback from users as to the usage experience.

\textsuperscript{*}Corresponding author. E-mail: rchung@cuhk.edu.hk.
related to the web site. We summarize our study and highlight possible future developments in Section 8.

2. Sign language documentation: Technology and motivation

Commercial software has been used to document sign languages. FileMaker Pro is used to document British Sign Language for dictionary production [21] as well as an earlier version of a database on Auslan [8]. 4th Dimension has been used for a sign phonology database called SignPhon [2]. An earlier attempt to compile a software program named ‘SignStream’ specifically to document streams of signing was developed by Neidle [14]; and recently most researchers adopt ELAN (i.e. EUDICO Linguistic Annotator) developed by Max Planck Institute of Psycholinguistics in Holland for a similar purpose. SignStream is a MacOS application that allows users to view, annotate, analyze and search through video and/audio data. This software however imposes restrictions on text space in the transcription process, thus making the documenting of linguistic information largely impossible. ELAN is more versatile than SignStream in terms of sign language annotation and has been used in documenting and analyzing a good many sign languages. Some researchers also use ELAN in sign language teaching [13] and sign language acquisition research [10,19]. Recently, iLex has been promoted to supplement ELAN in sign language corpus building. Although both systems share a lot in common, iLex offers better support to tier dependencies especially when a number of sign languages are brought together for comparative purpose [1].

Despite these advancements in the research community, not much has been done in Asia to document sign languages or to make them accessible to the Deaf or the laymen in the society, except for Taiwan Sign Language Online Dictionary [18]. In fact, only a handful of Asian sign language dictionaries have been compiled that are based on sign linguistics principles [7,19].

One obvious advantage of sign language documentation is the raising of deaf awareness not only among the Deaf but also people in the majority hearing societies. One aspect of such awareness is the importance of distinguishing natural sign language from those artificial signing systems, which will consequentially promote the use of natural sign language in deaf education and to dispel the misconception that exposure to natural sign language impedes spoken language development of deaf children [17]. The emergence of sign bilingualism has been gaining grounds in recent years, which advocates that Deaf people have the rights to develop both sign and spoken language [5,6,13]. This is evident by a recent proposal in the UN Convention on the Rights of Persons with Disabilities that sign language should be used in various social domains including deaf education. In addition, the 21st Organizing Committee of the 2010 International Congress on the Education of the Deaf made a public apology to the Congress attendants and Deaf Communities worldwide for having deprived Deaf people of access to sign language and Deaf teachers in deaf education for 130 years. This apology came also with a landmark decision to urge the restoration of sign language and Deaf teachers in educating deaf children in whichever educational setting where this need is called for. These high-profile demands for use of sign language to support deaf developments worldwide require enormous resources delivered in sign language. In deaf education for example, as many as 95% of the deaf children are born of hearing parents and receive no sign language input after birth to support the acquisition of sign language alongside spoken language. Most only receive sign language input when the parents are willing to hire adult deaf signers to teach the language at home, or when the deaf children are enrolled in a school system involving deaf teachers who know sign language. In higher education or various social domains that require sign interpretation, the lack of sign language resources as reference materials is more of a norm than an exception.

On the technical front, Deaf peoples in Asia are usually distributed over isolated clusters and seldom do they interact with the hearing people in the larger community or even with each other. Although the advancement of computer technology has greatly enhanced deaf communication and there have been several websites for Deaf forums in the world, they primarily support textual input and display. It is only in recent years have we seen the emergence of Deaf vlogs such as DeafRead [3]. Where Deaf people from different regions have a chance to meet, one interesting observation is that they always compare signs among themselves, probably to check similarities and differences of their signs in order to facilitate the exchange of Deaf information across the region if not the globe. Given these needs from the various sectors, sophisticated database systems are called for that support cross-linguistic and cross-cultural communication among the Deaf communities as well as between the Deaf and hearing communities. The question is whether such systems are manned as a closed information system or
shared and co-built by members of the Deaf and hearing communities interested in sign language and deaf developments.

According to the Declaration of Principles in World Summit on the Information Society (WSIS) held in Geneva on 12 December 2003, joint efforts are needed to “build a people-centered, inclusive and development-oriented Information Society, where everyone can create, access, utilize and share information and knowledge” and “the special needs of persons with disabilities and disadvantaged and vulnerable groups should be taken into account” [4]. In view of this declaration and in response to the urgent need of promoting sign linguistics research and deaf developments in Asia, we developed a project to create a database on Asian sign language varieties, which we hoped would reflect the philosophy underlying the declaration above and the recent trends in promoting natural sign language as language of the Deaf. At the time of development, we found that such a database required great effort to compose due to the immense, undocumented varieties of Asian sign languages; we also faced an even more demanding task of how to ensure the correctness of sign entries. Borrowing the concept of Wikipedia (The Wikimedia Foundation), we experimented on passing the composition effort to the database users themselves through the use of an open web page. This article presents the design of the web page, which we refer to as “Asia Signopedia”, on both its data structure and user interface. The web page had the address http://www.asiasignopedia.org/signopedia/ until March of 2010. One objective of the Asia Signopedia database was to set up a cross-modality information system where Deaf and hearing people can share and exchange information about the sign language varieties in the Asia-Pacific region. On this, the effort was the first of its kind in the Asia.

There are numerous benefits for such a research effort. Sign language represents a unique manifestation of Deaf culture and a heritage rich in native linguistic resources. Thus a database that is built upon the different varieties of sign languages as well as their cross-references is invaluable not only in the promotion of deaf developments in Asia, but also in the understanding and preservation of local Deaf cultures. By cross-referencing in sign language documentation, we mean, as far as possible, how a data entry on say “apple” in sign language A can be linked to the data entry of the same meaning in sign language B. An objective of the design of Asia Signopedia was that cross-referencing could be built automatically by the distributive users without themselves making conscious efforts of it.

Establishing such a platform was in line with the recent development that an increasing number of Deaf people access the internet and use multimedia technology to capture communication in sign language in recent years. Therefore, visually signed video clips are more Deaf-friendly than textual communication. To place this observation in the context of our design, two points can be noted. Firstly, an online database would only become more useful to more Deaf people over time. Secondly, the sign language database needs to be multimodal in the mode of interaction, like allowing both video and textual access, so that both the technical needs of Deaf and hearing people can be satisfied. In our system, we also deployed the concept of Web 2.0, in which the database was a web interface that encouraged any reader to become an author himself. The database contents could thus be read and composed by users all over the world who could access the internet. This concept also encouraged self-correction through the collaboration of the internet users.

In addition, we expected that the sign entries in the system not to be composed by users who were novice to sign language but by those who had been using the language in daily communication, in other words, users of a sign language community in which regional and territorial sign languages were often the preferred communication mode. A critical issue of the design was thus how to allow people of various territorial sign languages to access the database or to input data entries of a particular sign language variety using text and/or video clips in the interface. To achieve the above, our system designed a special database’s data structure and web page users’ interface. The principle was that the system should be multi-textual and multi-lingual in both sign input and explanation, to optimize communication among signers, Deaf and hearing, as well as providing a platform for novice users to learn about the sign language varieties and their cultures.

3. Previous works

3.1. Wiki websites

Information and communication technologies have been gradually transforming our perception about how knowledge may be transferred, maintained or utilized for different purposes. On this there is the problem of how people with conflicting perspectives and values can all come to agreement on the information. A number of online communities are trying to resolve such
a problem by allowing people to meet and voice their views. However, given the endless arguments often found in traditional forums, inviting a large group of virtual users to reach a consensus online may seem impossible [22]. Therefore, the concept of computer supported cooperative work (CSCW) and Web 2.0 have been established with this specific goal in mind. One example of this is a special kind of web site known as “wiki”. Invented in 1995 by Ward Cunningham, wiki has the defining feature that any reader of the site may also be an author. Each page has an “edit this page” link allowing users to change the content of the page. This interface supports a higher level of consensus building because a user who disagrees with a statement can suggest deleting it very easily. In this sense, the ultimate text on wiki pages is content that has survived the critical eye of the community. In sum, Wikipedia is the largest public wiki in the world. It currently contains articles of more than 100,000 subjects. Wikipedia is multilingual even on similar subjects. However, there has not been a wiki website that is cross-modal between a sign and a spoken language. Deaf people can hardly enjoy the benefits of such technology with the concept of Web 2.0.

3.2. Online databases

Some online sign language dictionaries and database systems have been developed for not only Deaf and hearing individuals, but also for scholars and researchers to access information about sign languages. The Auslan Signbank (Johnston 2009 http://www.auslan.org.au) offers a platform to collect and disseminate information about Auslan. It consists of four modules (a) SignBank, (b) Medical SignBank, (c) Finger spelling, and (d) Number signs. Of relevance to our project are websites on SignBank and Medical SignBank where signs are grouped alphabetically and each comes with a video clip showing sign demonstration and definition. A special interactive feature is built into the system whereby viewers can offer feedback on the various aspects about the signs they view. This interactive feature is absent in other systems like The Taiwan Sign Language Online Dictionary (http://tsl.ccu.edu.tw/htmltext/browser.htm) developed by Tsay et al. [18] or Asia SignBank. As for the latter, rather than focusing on one specific sign language of a country, the Asian SignBank features a cross-linguistic database system first ever developed in Asia. It aims to document Asian sign languages to support research as well as teaching and learning at the tertiary level. The development of the Asia SignBank is supported by a team of sign linguists and professionally trained Deaf researchers through the Asia Pacific Sign Linguistics Research and Training Program http://www.cuhk.edu.hk/cslds/apsl.

However, the above database systems are ‘closed’ or ‘semi-closed’ because incoming information about sign language is usually screened by sign language researchers before it is incorporated into the web site. Nevertheless, such efforts point to the fact that compiling sign language dictionaries or knowledge systems about Asian sign languages are very much lagging behind and are largely works of sign linguists but not sign language users in the larger community. Given the fact that sign languages are not universal and they do have their own regional if not territorial variation, to capture them through sign linguistics research projects on a manual basis is going to be daunting. Seen in this light, community support may deem a plausible solution for speeding up the process. The wiki philosophy which encourages users of sign languages to inform the community about their signs or to reach a consensus on the signs they observe through discussions on an internet platform indeed serves the role of ‘information and opinion pool’ about Asian sign languages, from which sign language professionals and researchers can draw references for their own academic purposes. To support these developments in Asia, using the concept of Web 2.0 for the development of Asia Signopedia was of definite advantages: the database soon could reach a much wider audience; it invited participation from both hearing and deaf people; its data entries might grow at a faster rate; it was less labor intensive because it was the collective effort from the community of sign language users. It also served the role of informing the research community some general observations about the sign languages and deaf situations in Asia.

4. Data structure design

The Asia Signopedia was an open platform that accepted video input and presented both the textual and non-textual data in a readable and accessible format. This platform was built to complement the Asian Signbank described above and to pitch at a much wider community of users. There were two types of entries in the Signopedia. Users could upload a video for a particular sign as a sign entry. Each sign entry would have textual description as well as sample sentences in sign videos. Users could also compose articles about sign language and deaf cultures or any related topics. These articles
could be in textual format as well as in sign language videos so that deaf and hearing people could share their knowledge in a barrier-free environment.

4.1. Sign entries

Each uploaded video was treated as a single sign. A sign was associated with a region (sign language) and had one or more textual meanings. Those meanings could be coded in different written languages. An example is shown in Fig. 1.

In addition to textual meaning, a sign entry could have a brief explanation in both textual and signed video format. Explanations in multiple written and sign languages could be tagged to a single sign entry, as illustrated by the example in Fig. 2.

4.2. Sample sentences

Users could upload sample sentences for a sign entry in video format. Though not mandatory, users could provide textual glosses and translation for the sentences. For instance, a sample sentence for “father” was glossed as “I FATHER NOT-HAVE” and translated as “I do not have a father”. The textual contents for a sentence could be in multiple languages, provided that there was a sign meaning in the corresponding languages. A sample sentence is illustrated in Fig. 3.

4.3. Articles

Users could upload articles or documents about the sign languages and deaf cultures in the region. Each article would contain a title, keywords and an abstract so that it could be searched easily. A document storage space was provided for each user. They could upload their documents, either in text or signed video format, or both, into their dedicated document space. In this way, web users could also view the documents wherever there was a link with the article entry. Figure 4 shows how documents may be linked to each other in the user’s document space.

5. User interface design

Asia Signopedia was designed and set up as a platform that was deaf-friendly and technologically user-friendly for data input purpose. It had an interface rich in graphics to attract people and clear in presentation, thus making the platform simple to use for both Deaf and hearing people. Figure 5 shows the front page of the Asia Signopedia web site.

5.1. Welcome page

The welcome page (Fig. 5) would be displayed in the browser when ‘Asia Signopedia’ was entered. It contained an introduction of the platform. This was important for web users who accessed to the web page the first time. There were other similar pages to explain the platform further. They were “About Us” and “About This Website”.

5.2. Search page

Users could search for a particular sign or article anywhere in the website because all pages contained a search box in its header. When searching to sign entries, users could limit the search by selecting a specific sign language. After clicking the search button, a list of sign entries would be shown on the search result page. Each entry was shown by its sign meaning, the sign language it belonged to, and the user who created the entry. The result items were themselves hyperlinks if viewers wished to watch the signs.

5.3. Sign page

A sign entry was usually documented with a signed video clip, meaning description(s) of the sign in textual and signed format, and sample sentences. There were three tabs on the sign entry page. The first tab showed the signed video clip and its meaning(s). If there were multiple meanings or meanings in multiple languages, a list of glosses would be shown. The second tab delivered description(s) about the sign in either or both text and signed video format. Users could select their own sign language to view the content if their respective signed description(s) were available. The third tab contained all the sample sentences together with their textual glosses and translations.
5.4. Uploading and editing signs

As mentioned before, each page in a wiki website had an “edit this page” link, allowing users to change the content of the page. In Asia Signopedia, sign pages had links and buttons to add new meanings, edit textual descriptions and add/remove signed description in video clips as well as sample sentences. To avoid in sign entries, users had to confirm the existence of the sign entries first before they uploaded a new sign into the system. To support this function, there was a “Create new entry for XXX” link on the search result page where XXX was the search string. This function checked automatically for duplications. Then, users might wish to upload either a completely ‘new’ sign or one that differed from the existing one(s) in terms of some phonological or semantic characteristics. To upload this new sign, there was a convenient hyperlink “Create your sign here” on the welcome page. It showed a dialog box for users to enter the sign meaning. Designed with such procedures, we observed the philosophy of sign language research that information about variation should be preserved in natural sign languages as much as possible. In time, the web could offer resourceful information regarding gender, social and regional differences among the varieties in Asia.

5.5. Listing articles

All pages contained a search box in its header. Users could select either sign entries or articles for the search.

There was a convenient button called “List Articles” when users switched the search box into search of articles. A click of this button would lead to a list of articles with titles posted on the platform by users.

5.6. Article page

If users knew the title of an article, they could type it in the search box and click the “Go to Article” button. A page of the abstract of the corresponding article would be displayed. If the article did not exist, users could compose a new article with the title they just typed in the search box. Every article page had the “Edit this article” link to allow the users to edit the article’s keywords, abstract and attached documents.
Fig. 4. The link between uploaded documents and articles (each in text or signed video format) in the document space is all up to users’ specifications.

Fig. 5. The front page of the Asia Signopedia web site.

5.7. Attaching documents in the article

Users had their own file storage space after they had registered to Asia Signopedia. They could upload document files, e.g. text documents, signed video clips and images, into their space. Users could click the “Attach Document” button on the article writing page to manage their storage space. They could then create hyperlinks for the documents as an article entry.

6. Database development

Since its inception in March 2008, a number of figures had been recorded everyday about the growth of
the database. They were (1) Visitor count and read access rate; (2) New sign entry count; (3) Textual meanings count; and (4) Cross references for signs that showed alternative ways of signing for the same meaning. Here we will report on (1) and leave the other domains for future exploration.

Since the launch of the Signopedia on March 31, 2008, the number of visitors grew steadily. Figure 6 shows that as at March 09, one year after the website had been launched, the number of users had soared to slightly below 3500, suggesting that this website had potentials for developing itself into an open platform to support the dissemination of sign language and deaf information.

6.1. Patterns of cooperation and cross-referencing

The Asia Signopedia was a wiki website where every user was also the author of the content. Unlike Wikipedia which requires that submissions to the website have to observe factual accuracy, Asia Signopedia had a slightly different philosophy in rendering sign language information to the website. As sign languages vary not only from countries to countries, regions to regions, but also between individuals, Asia Signopedia allowed users to upload different ways of signing even if they were of the same meaning. Thus, cross referencing of sign entries was an important feature in the design as accumulating information would form the basis of future research and of cross-cultural communication among the Deaf communities. Therefore, each sign entry and article could only be edited by the composer. Only the textual glosses and translations for sample sentences could be edited by all users.

In order to maintain content stability, a team of sign linguists would be responsible for monitoring the sign entries. Screened entries would then be used for research and could be exported to other sign language databases.

7. Feedback from users

Before launching the website, we conducted a trial at the Centre for Sign Linguistics and Deaf Studies in which six users, half of them deaf and others hearing, completed a questionnaire after they had navigated the website, browsed through the data and uploaded some of their own productions onto it. Most users found that the procedures of uploading sign entries rather complicated. They asked for clearer instructions for installing the plug-in software if a user wished to create sign entries directly through a webcam. Some deaf people had difficulty in conducting a search. Nevertheless, given the facilities provided, most deaf people were willing to contribute their signs to the platform. In sum, they wanted to see a more friendly user interface with clearer explanations. Most users, either deaf or hearing, were not so sure if they would continue to post articles onto the platform.

During the one year period in which Asia Signopedia was launched, although we had recorded a high number of viewers, only 37 users had registered as regular users but no records of new sign entries were observed. There may be many reasons for this out of which we would like to highlight three: (1) The uploading procedures were too complicated, as conveyed by the user survey; (2) The processing of video
data and the associated uploading procedures posed difficulty when compared with textual input, suggesting that there might be a technological bottleneck in implementing multi-modal communications; (3) People were more interested in viewing the already documented sign entries than contributing their own as collecting signs could be quite time-consuming. Therefore, we had enhanced the ‘welcome’ page by adding a “Create your sign entry here” or “Create your article here” link as shown in Figure 5. We had also investigated the use of webcam to capture the user’s signing directly into the platform. A prototypical version of such feature had been demonstrated during the user survey; however, this procedure was not possible if a software program called the Java Media Framework had not been installed (http://java.sun.com/javase/technologies/desktop/media/jmf) which requires some technical configurations in the user’s computer in order to make it function.

8. Conclusion and future work

The Asia Signopedia was a sign language knowledge-base platform open to people or researchers interested in sign language and deaf culture in Asia or worldwide. It had the potentials for creating a virtual community with members who may eventually regard themselves as part of the developing team, gradually compiling facts and figures about the sign language varieties in the Asian region. Although the platform was suspended after a year of operation, we had at least acquired a set of technical know-how for developing a web site through a medium of information transfer that was more attuned to use by deaf users. Asia Signopedia confirmed that they particularly preferred visual rather than textual information and demanded a high degree of transparency in disseminating information about their language and culture and by doing so preserving it. Future researchers can certainly draw references from our experiences in developing Asia Signopedia in their attempt to build a wiki platform to document information about sign languages in future to support deaf developments, and ultimately to support research into sign linguistics, computer vision, automatic sign recognition, multimodal computing, virtual reality and computer supported cooperative works.

Acknowledgements

The work described in this paper was supported by a grant from Microsoft Research Asia (Project No. 6902262). The work was also technically sponsored by the Centre for Sign Linguistics and Deaf Studies, and the Department of Mechanical and Automation, of the Chinese University of Hong Kong. This work is affiliated with the Microsoft-CUHK Joint Laboratory for Human-centric Computing and Interface Technologies.

References


