
An Alternative Method for Building A Database for American Sign Language

Noriko Tomuro, Karen Alkoby, Andre Berthiaume, Pattaraporn Chomwong, Mary Jo Davidson, Jacob Furst, Brian Konie, Glenn Lancaster, Steven Lytinen, John McDonald, Lopa Roychoudhuri, Jorge Toro, Rosalee Wolfe

School of Computer Science, Telecommunications and Information Systems
DePaul University
Chicago, IL
asl@cs.depaul.edu

Abstract

Research shows that English literacy of deaf people in the United States is much lower than that of hearing people. Most deaf adults read English at the third or fourth grade level. That is because English and American Sign Language (ASL), the native language used by the Deaf community in the US, have very different syntactic structures. Because of the lack of reading fluency, English-based assistive technology such as closed captioning on television offers some help in making the hearing world accessible to deaf people, but can only serve as a partial aid. In order to fully bridge the gulf between the deaf and hearing worlds, alternative assistive technologies based on sign language are needed.

1. Introduction

American Sign Language (ASL) is a natural language used by members of the North American Deaf community and is the third or fourth most widely used language in the United States [Ster96]. While ASL shares some vocabulary with English, research in linguistics shows that it has a concise and elegant syntax which differs radically from English grammar [Bake80][Vall93]. This makes the acquisition of English rather challenging for ASL signers. Research shows that most native ASL signers read English at the third or fourth grade level [Holt94]. Therefore, English-based assistive technologies such as closed captioning on television do not provide a sufficient solution [Wolf99]. At present deaf people rely on sign language interpreters for access to spoken English. However, they can not depend on interpreters every day in life mainly due to the high accommodation cost and the difficulty in finding and scheduling qualified interpreters.

A better solution would be to use a personal digital English-to-ASL translator. An English-to-ASL translator would convert written or spoken English into a three-dimensional graphic animation depicting ASL. It would be an economical as well as flexible solution. Since personal computers have become lighter and are available for less than \$1000, deaf people can bring a laptop whenever and wherever they go. Such a translator would also convert full English

sentences to ASL. Thus, it could convey complex ideas expressed in English to ASL and is much more useful and flexible than sign dictionaries or phrase books.

Not only would this ASL synthesis technology assist deaf people, it can also be used as a valuable tool for educators and researchers. By using three-dimensional graphics, signs can be viewed from any position, from the signer's standpoint, or from the observer's stand point. It is even possible to view a signer from the side to see how far forward a sign extends. Also, animation can show the timing of signs visually. Thus, this synthesis technology offers the potential for a rich and flexible environment for ASL education and research.

However, developing this kind of English-to-ASL translation system imposes serious technical challenges. Because of the unique modality of ASL as a sign language -- visual/gestural rather than aural/oral -- the translation system must store the linguistic and geometric aspects of ASL signs and generate graphic animations on screen in real time. Also, a sequence of signs as an ASL sentence must look smooth and natural. Therefore, building such an English-to-ASL translator requires expertise in a wide range of areas, including linguistics, machine translation, computer graphics, mathematics and kinesiology.

We are currently in the process of developing a database of ASL signs that will be used as the lexical database in our translator. The database scheme draws on the experiences of Dutch [Cras98], German [Pril89], and Japanese [Lu97] researchers who are working on similar projects for other sign languages. It includes such items as position, orientation and shape of the hands as well as motions.

The largest task in creating a database of this type is data entry, which involves transcribing ASL. To transcribe sign language, several researchers used motion capture, which recorded the motion and position of hands through gloves with sensors [Eren96][Fels98]. However, despite a considerable amount of financial investment required for equipment, animations produced from the recorded data are often inaccurate [Zord99]. Motion capture also has a critical disadvantage in that recorded numerical data is hard to modify and abstract to symbolic level into linguistic features of ASL signs, such as fingers being "hooked" or "bent".

Our approach is to use an animation software package and customize it for fast sign transcription. Customizing a general animation package has a great advantage in speeding transcription time. Normally, learning a general package requires a significant time investment. Our students reported that working through the tutorials of a commonly used animation package took between 40 and 100 hours. Few volunteers from the deaf community are willing to invest such a large amount of time in training before beginning the transcription process. By eliminating features in the package that are irrelevant to ASL and facilitating an interface that is intuitive and ASL-specific, minimal learning is required for transcription.

2 Transcription System

Our transcription system has a bi-level structure. The lower level, the hand transcriber, is used to build the handshape data. The upper level, the sign transcriber, relies on the handshape database and allows users to specify the location and motion of the two (left and right) hands.

Both transcribers utilize familiar controls such as checkboxes, selection lists and slider bars. Most labels are symbolic, linguistic features of ASL signs, and the underlying mathematical information is completely hidden from users. The interface is what-you-see-is-what-you get (WYSIWYG), which allows users to see the signs graphically as they enter data. To verify that a sign appears correct from all angles, users can "walk around" any handshape or sign they create.

2.1 The Hand Transcriber

The hand transcriber allows users to specify handshapes. Users can select one or more fingers at a time and move them to a desired position. Slider bars specify the configuration of the selected fingers, for instance somewhere between "together" and "spread", or between "flat" and "hooked". Figure 1 shows an example screen capture of the hand transcriber.

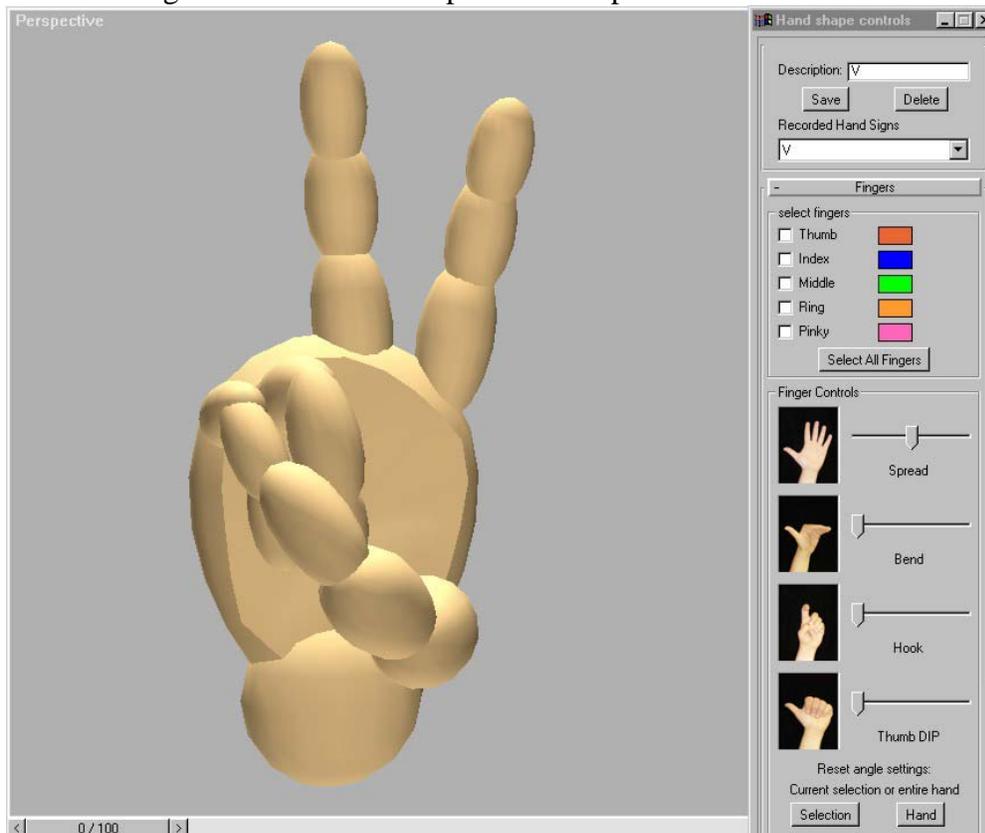


Figure 1: The Hand Transcriber

2.2 The Sign Transcriber

The sign transcriber is built on top of the handshape database. It allows users to specify ASL signs in terms of handshape, location and orientation for both hands. Motions are entered by specifying this information for various time steps. Then the sign transcriber generates a sign as an animation, with the configuration of step1 as the start key frame and the configuration of the last step as the end key frame.

Figure 2 shows a screen capture of the sign transcriber. To enter the configuration for a particular time step, users first specify either left or right hand and select a handshape from the database. Figure 3 shows the pop-up window of handshapes. Then users specify palm orientation for the hand by selecting from a list box ("up", "down" etc.). Next comes the specification for the hand location. A location is defined by two parameters: vertical height and horizontal space. Height is specified by raising a bar in the side view of a human body ("business woman"), and the horizontal space is specified by selecting a button at the position in the top view of a human body. After the user repeats this process for the other hand, specification for a time step is complete. At this point, hand information is displayed graphically as part of a full body ("bubbleman") in the Perspective window. The user continues to enter specification for more time steps. Then by selecting the Sign menu from the top menu-bar, an animation is displayed in the Perspective window.

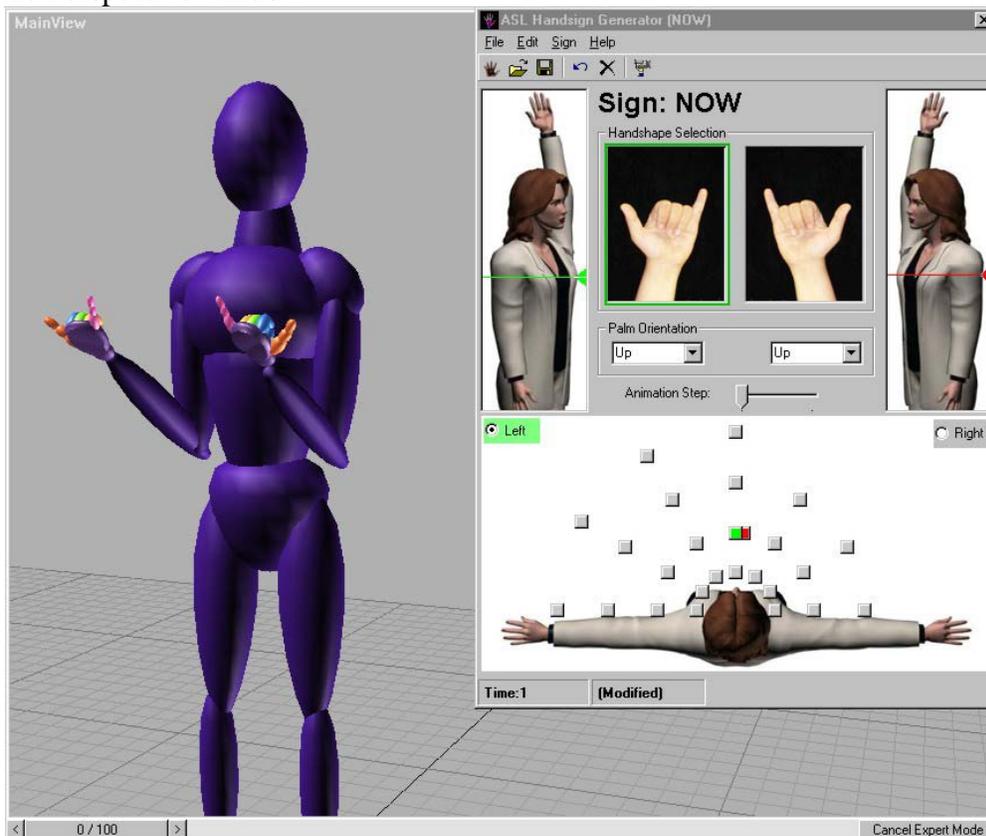


Figure 2: The Sign Transcriber

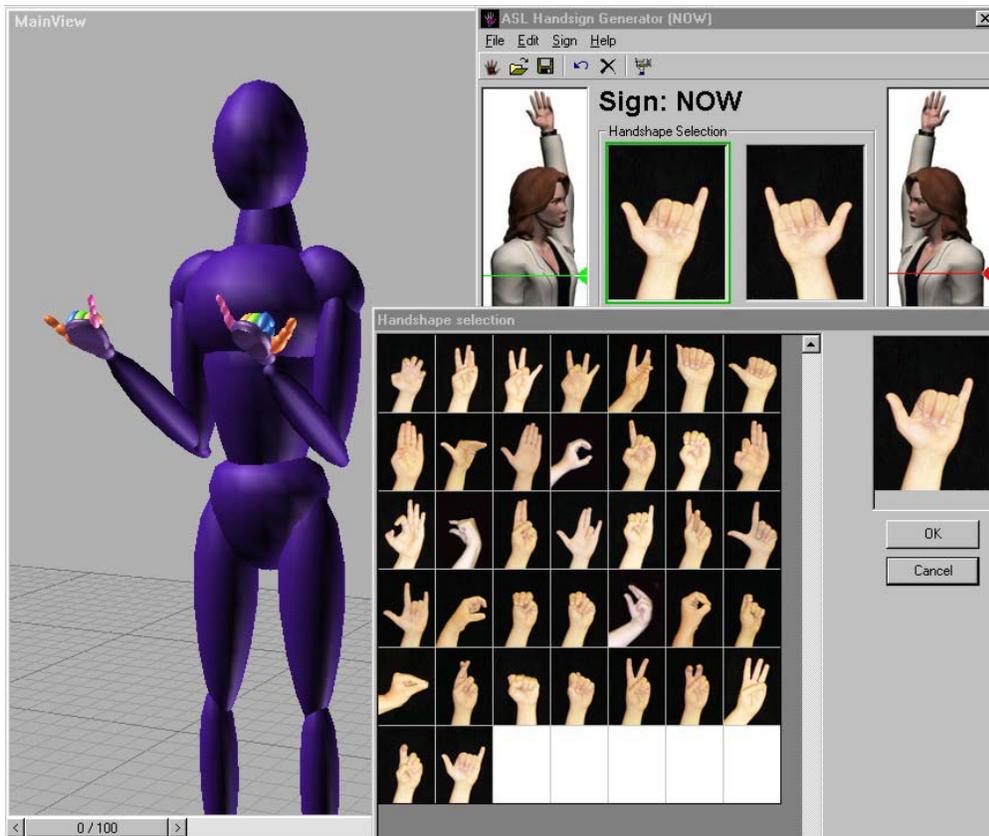


Figure 3: Hand Selection in the Sign Transcriber

3. Future Work

Our initial usability studies have been quite promising. On average, it took 10 minutes for native ASL signers to learn enough about the handshape transcriber to create handshapes. Transcribing a handshape took an average of 82 seconds. In order to make the transcription system truly useful, more in-depth usability testing by ASL signers of various computer literacy levels is needed.

Another important extension is to incorporate non-manual features of ASL, namely facial expression and body motion, in the database. In ASL, non-manual signals are just as critical as hands in defining signs. A transcription system which incorporates all these aspects of ASL could be used as a tool for teaching ASL, computer-aided tutoring in ASL and many other tasks.

References

[Bake80] Baker-Shenk, S. and Cockely, D., "*American Sign Language, A Teacher's Resource Text on Grammar and Culture*" (the Green Book), Gallaudet University Press, 1980.

[Eren96] Erenshteyn, R., Laskov, P., Foulds, R., Messing, L., and Stern, G., "*Recognition Approach to Gesture Language Understanding*". In Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria, 1996.

[Fels98] Fels, S. and Hinton, G., "*Glove-TalkII: A neural network interface which maps gestures to parallel formant speech synthesizer controls*", IEEE Transactions on Neural Networks, pp. 205-212, Vol 9, No. 1, 1998.

[Holt94] Holt, J., "*Demographic, Stanford Achievement Test – 8th Edition for Deaf and Hard of Hearing Students: Reading Comprehension Subgroup Results*".
<http://www.gallaudet.edu/~cadsweb/sat-read.html>

[Lu97] Lu, S., Igi, S., Matsuo, H. and Nagashima, Y., "*Towards a Dialogue System Based on Recognition and Synthesis of Japan Sign Language*", Gesture and Sign Language in Human-Computer Interaction, 1997.

[Pri189] Prillwitz, S. et al., "*HamNoSys Version 2.0; Hamburg Notation System for Sign Languages*", International Studies on Sign Language and Communication of the Deaf; Signum 1989.

[Ster96] Sternberg, M., "*The American Sign Language Dictionary*", Multicom, 1996. (CD ROM)

[Vall93] Valli, C. and Luca, C., "*Linguistics of American Sign Language*", Gallaudet University Press, 1993.

[Wolf99] Wolfe, R. et al., "*Interface for Transcribing American Sign Language*". In Proceedings of the 26th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), Los Angeles, 1999.

[Zord99] Zordan, V. and Hodgins, J., "*Tracking and Modifying Human Motion with Dynamic Simulation*". In Proceedings of the 26th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), Los Angeles, 1999.