American Sign Language Natural Language Generation and Machine Translation

Matt Huenerfauth
Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104 USA
matt@huenerfauth.com

MOTIVATIONS AND APPLICATIONS

Although Deaf students in the U.S. and Canada are taught written English, their inability to hear spoken English results in most Deaf U.S. high school graduates (18 year olds) reading at a fourth-grade (10 year old) level (Holt, 1991). Unfortunately, many Deaf accessibility aids, like television closed captioning or teletype telephones, assume the user has strong English literacy skills. Many Deaf people with English reading difficulty are fluent in American Sign Language (ASL); so, an English-to-ASL automated machine translation (MT) system could make information and services accessible when English text captioning is too complex or an interpreter is unavailable.

English-to-ASL MT software could translate English text into an animation of a 3D virtual reality character performing ASL. In this way, a variety of English text sources could be made more accessible to the Deaf, including: closed captioning, teletype telephones, computer interfaces, Internet information and services, etc. An MT system also has educational applications; software for Deaf students learning English or for hearing students learning ASL could use this technology to translate novel English input sentences into an ASL animation.

LINGUISTIC ISSUES

American Sign Language has a different word order and linguistic structure than English; so, English-to-ASL translation is as complex as translation between pairs of spoken languages. In fact, ASL's visual modality allows it to use linguistic phenomena not seen in any spoken language (Neidle et al., 2000; Liddell, 2003). The space around the signer can be used for many communicative purposes, including the production of constructions called "classifier predicates." These complex hand movements trace contours, identify locations, or draw paths through the space in front of the signer that topologically correspond to the shape, location, or movement of real world entities in a three-dimensional scene or event being discussed. When the ASL equivalent of an English sentence uses a classifier predicate, then the structure of two sentences is quite divergent. For example, the entire sentence "the car drove up the hill" may be expressed in ASL with a single hand movement tracing a hill-like 3D path through the air in front of the signer. Classifier predicates challenge traditional linguistic representations by their use of spatial metaphor and visualization (Liddell 2003) and thus cannot be generated by most MT software.

PREVIOUS WORK

Several previous MT projects have made initial efforts at translating English text into ASL animation; see survey of systems in (Huenerfauth, 2003). All of them use standard machine translation technologies, and so none of them model the 3D spatial arrangement of objects in a scene being discussed – a capability necessary for producing classifier predicates. Many English sentences lack an ASL translation without classifier predicates, and some ASL sign-frequency studies indicate that, depending on genre, signers produce a classifier predicate once per minute (Morford and MacFarlane, 2003). So, these previous systems can only produce ASL with limited fluency. However, it is the English/ASL sentences pairs these systems do not handle that would be particularly useful to translate for a Deaf user with limited English skills (since these are the pairs that are the most different structurally).

RESEARCH GOALS

The goal of this project is to design and build an MT system to generate animations of ASL classifier predicates from English sentences describing 3D scenes and events. There are two major facets to this work: (1) a new ASL representation that enables the generation of classifier predicates and (2) an overall software architecture design that will allow this classifier-predicate generation component to be part of a complete English-to-ASL translation system that handles a variety of sentence types.

RESEARCH ACCOMPLISHMENTS

After specifying the English-to-ASL problem space and surveying previous work (Huenerfauth 2003), recent ASL linguistic research was examined to understand how traditional MT techniques must be modified to handle the special properties of ASL. Several competing linguistic theories of classifier predicates were examined as starting points for potential classifier predicate generation designs (Huenerfauth 2004b). Finally, a design was created whose linguistic model accounts well for human-produced classifier predicate data and which minimizes the programming overhead required to implement it.

The design uses virtual reality software to calculate and store a model of the 3D coordinates of entities discussed in a spatially descriptive English text. The animated ASL signing character can use the 3D coordinate data to calculate the corresponding arm movements needed to produce a classifier predicate to describe each entity's 3D location or movement. The 3D model is analogous to the miniature invisible objects that human ASL signers imagine floating in space around them when using classifier predicates to describe a 3D scene. To decide the locations and movement paths of the objects in the model based on the English input text, this design uses existing "natural language control" virtual reality software (Bindiganavale et al., 2000). This software accepts English text containing instructions for a set of 3D virtual reality objects to follow, and it moves the objects in the 3D model to obey the English input sentences.

English input sentences not producing classifier predicates would not need to be processed by the virtual reality software; in fact, most of these sentences could be handled using the more traditional machine translation technologies in previous systems. For this reason, the classifier predicate generator is part of a translation software design that contains multiple processing pathways (Huenerfauth, 2004a). The pathway for handling English input sentences that produce classifier predicates includes the virtual reality software, while input sentences whose translation requires less sophisticated processing can use a traditional MT approach (that is easier to implement). In this way, the classifier predicate generation design could actually be built on top of some existing English-to-ASL MT system to give it the ability to produce classifier predicates.

OPEN ISSUES AND FUTURE WORK

Implementation of initial portions of this design is planned for the summer and fall of 2004, and it is anticipated that these efforts will motivate refinements to the specification. To motivate the development work, an initial application area for this technology will be selected in which classifier predicates frequently arise: domains containing a lot of spatial description, computer interfaces in which elements of the screen need to be spatially described, or educational software to help Deaf users understand spatial English text.

REFERENCES

1. Bindiganavale, R., Schuler, W., Allbeck, J., Badler, N., Joshi, A., and Palmer, M. 2000. Dynamically Altering Agent Behaviors Using Natural Language Instructions. 4th International Conference on Autonomous Agents.
2. Holt, J. 1991. Demographic, Stanford Achievement Test - 8th Edition for Deaf and Hard of Hearing Students: Reading Comprehension Subgroup Results.
3. Huenerfauth, M. 2003. Survey and Critique of ASL Natural Language Generation and Machine Translation Systems. Technical Report MS-CIS-03-32, Computer and Information Science, University of Pennsylvania.
4. Huenerfauth, M. 2004a. A Multi-Path Architecture for Machine Translation of English Text into ASL Animation. HLT-NAACL Student Workshop.
5. Huenerfauth, M. 2004b. Spatial Representation of Classifier Predicates for Machine Translation into American Sign Language. Workshop on Representation and Processing of Signed Languages, LREC 2004.
6. Liddell, S. 2003. Grammar, Gesture, and Meaning in American Sign Language. UK: Cambridge Univ. Press.
7. Morford, J., and MacFarlane, J. Frequency Characteristics of ASL. Sign Language Studies, 3:2.
8. Neidle, C., Kegl, D., MacLaughlin, D., Bahan, B., and Lee, R.G. 2000. The Syntax of American Sign Language: Functional Categories and Hierarchical Structure. Cambridge, MA: The MIT Press.