Designing an English to American Sign Language Machine Translation System

Matt Huenerfauth
Computer and Information Science
University of Pennsylvania
matt@huenerfauth.com

Why build an English-to-ASL System?

Literacy and Deafness
- Only half of deaf high school graduates (age 18+) can read English at a fourth-grade (age 10) level, despite having fluency in ASL.
- Many deaf accessibility tools forget that English is a 2nd language for these students (and has a different structure than English).

Applications for a Machine Translation System
- TV captioning, teletype telephones.
- Computer user interfaces in ASL.

Signed English vs. American Sign Language
- ASL has word order and linguistic structure that is different than English.
- Can't just replace words for signs and get ASL (you'd get Signed English).

ASL is Difficult for Machine Translation (MT) Software
- No written form of ASL, so there are few parallel English-ASL language samples available. Can't use machine learning MT methods.
- Some ASL sentences contain Classifier Predicates, a complex spatial linguistic phenomenon not seen in written languages.
- During a classifier predicate, the signer's hands represent objects under discussion. The 3D layout of these objects is topologically mapped to the volume of space in front of the signer's torso.

What's a Classifier Predicate?

The car parked between the cat and the house.

Design Issues

Some classifier predicate hand motion paths are linguistically determined...
- For some classifier predicate expressions, the motion of the hand is not a direct visual representation of the movement of the object.
- This system stores classifier predicates as prototypical movement templates.
- Sometimes there is not one-to-one mapping from English sentences to classifier predicate...

Current ASL linguistic models are ill-suited to the representation of classifier predicates.
- Models of ASL phonology store too much information about the handspace and not enough information about orientation to be appropriate for classifier predicates.
- These models also make it difficult to specify the complex motion paths of CPs.
- This project has developed a new set of linguistic models that are specifically tailored to the representation of classifier predicates and are suitable for computational processing in a machine translation system.

The 3D processing approach is "over-kill" for some English sentences.
- Only those English sentences that produce ASL classifier predicates require the 3D processing approach outlined above. Other English sentences could be processed by the traditional MT technology used in previous ASL MT systems.
- This system uses a multi-path machine translation architecture to combine resource-light and processing-heavy approaches in a single system.

Previous Work Ignores CPs
- Traditional MT technologies (grammar rules, dictionaries) can't successfully arrange objects in a 3D scene or select 3D motion paths for the signer.
- Previous ASL MT projects ignored CPs in order to use traditional methods.
- But classifier predicates are important:
  - There's no way to convey some concepts in ASL without CPs.
  - Signs use CPs frequently (one to seventeen times per minute).
  - English sentences that translate into CPs can be difficult to read.
  - The original sentences look different structurally than the ASL version.
  - Classifier predicates are needed for several important applications.

How it Works

Converting English Text into Placeholders.
Arranging placeholders for GUI elements.

Discussion

Impact of this Design
- This is the first MT approach proposed for producing ASL Classifier Predicates.
  - It is unique in its use of a multi-path architecture and 3D graphics software.
  - The design has attractive implications for use in accessibility applications.
  - The problem of ASL generation has many similarities to other multi-modal generation tasks for written/spoken languages. In particular, research on the generation of gesture animations for embodied conversational agents could take advantage of the software architecture of this system.

Future Extensions
- The 3D graphics software at the heart of this translation system makes it easy to embed in a computer user interface. The classifier predicates produced by this system rely on laying out a set of invisible placeholders in the space in front of the signer. The GUI coordinates can be used to arrange these placeholders (as in the diagram above).
- While sign languages used in other countries have different signs and linguistic structure than ASL, they all have a system of classifier predicate expression. This 3D translation approach could easily be adapted to these languages.
- There are other linguistic phenomena in ASL (aside from classifier predicates) that could benefit from the rich way in which this MT system manages the space around the signing character.

Current Status
- A detailed software design has been created, and the system is in the early stages of implementation.
- A user-based evaluation of a prototype Classifier Predicate animation generator will be conducted with native ASL signers.

Acknowledgments
I would like to thank my advisors Dr. Mitch Marcus and Dr. Martha Palmer for their guidance, recommendations, and support during this research. I would like to thank Dr. Norman Badler for his assistance in obtaining additional funding for this project.