Designing an English to American Sign Language Machine Translation System

Matt Huenerfauth

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

Abstract

Common misconceptions about the English literacy rates of deaf Americans, the linguistic structure of ASL, and the suitability of traditional machine translation technology to ASL have slowed the development of English-to-ASL MT systems for use in accessibility applications. This English-to-ASL MT project has made translating texts important for literacy and user-interface applications a priority. In particular, this system is focused on producing ASL phenomena ignored by previous MT researchers: "classifier predicates." This project has proposed a new multi-pathway MT system that makes use of 3D graphics technology to manage the space around an animated ASL signing character. The use of graphics software in this system has several attractive design implications.

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).

Animated ASL signing character made by VCom3D Corporation.

Applications for a Machine Translation System

- TV captioning, teletype telephones.
- Computer user-interfaces in ASL.
- Educational tools using ASL animation.
- Access to information/media.

How it Would Work

- Input: English text.
- Output: Animated character performing ASL.
- There are off-the-shelf 3D virtual human characters.
- We must convert English into 3D instructions for one of them.

Signed English vs. American Sign Language

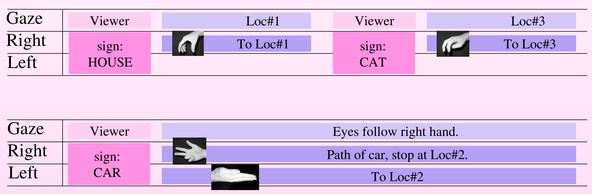
- ASL has word order and linguistic structure that is different than English.
- Can't just replace words for signs and get ASL (you'd get Signed English).

ASL is Difficult for Machine Translation (MT) Software

- There is 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.
(Loc#2) (Loc#3) (Loc#1)



Note: Facial expression, head tilt, and shoulder tilt are not included in this diagram.

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 hands.
- Previous ASL MT projects ignored CPs in order to use traditional methods.
- But classifier predicates are important!
  - o There's no way to convey some concepts in ASL without CPs.
  - o Signers use CPs frequently (one to seventeen times per minute).
  - o English sentences that translate into CPs can be difficult to read.
  - o The original sentences looks different structurally than the ASL version.
  - o Classifier predicates are needed for several important applications.

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.
- Some of the 3D path information to instantiate the template comes from the virtual reality scene. Other information is hard-coded inside the template.

Sometimes there is not one-to-one mapping from English sentences to ASL classifier predicate…

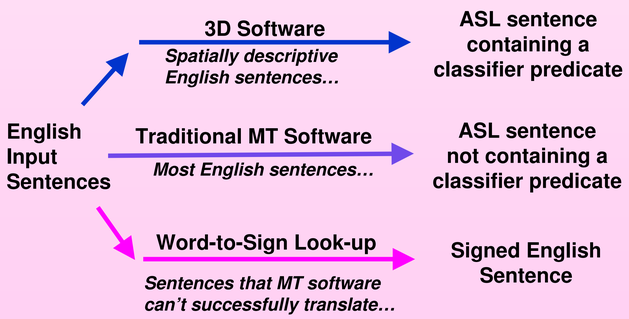
- Many-to-one, one-to-many, and many-to-many relationships are common.
- Sometimes multiple classifier predicates are necessary to convey meaning.
- To facilitate more flexible mappings from English sentences to classifier predicate templates, this system uses the same template formalism to record the structure *inside of* and *in between* classifier predicates.
- The templates are actually operators in a hierarchical planning system. This allows the effects of one classifier predicate to satisfy linguistic and spatial preconditions on later classifier predicates.

Current ASL linguistic models are ill-suited to the representation of classifier predicates.

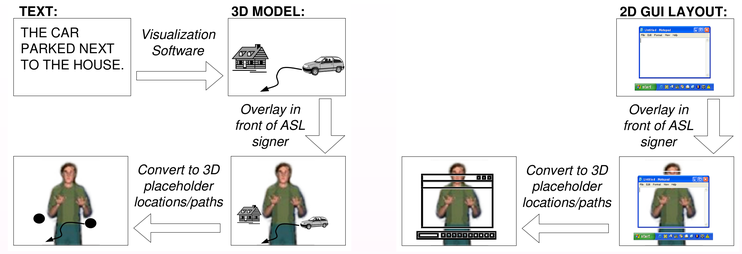
- Models of ASL phonology store too much information about the handshape and too little information about orientation to be appropriate for classifier predicates.
- These models also make it difficult to specify the complex motion paths of CPs.
- This project has developed a new set of linguistic models 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.



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.