# Experiments in Sign Language Machine Translation Using Examples

Author: Sara Morrissey
Supervisor: Dr. Andy Way
IBM Mentor: Alexander Trousov

National Centre for Language Technology, Dublin City University, Dublin 9, Ireland

## Introduction

Sign langauges (SLs) are the first langauge of the Deaf communities worldwide and, just like other minority languages are poorly resourced and in many cases lack political and social recognition. As a result of this, users of minority languages are often required to have multi-lingual competencies in non-L1 languages. In the case of SLs, this causes considerable hindrance to Deaf people as the average literacy competencies of a Deaf adult are equated with those of a 10-year old. To alleviate this, we propose the development of an automatic machine translation system to translate from spoken language text to SLs through the medium of a signing mannequin.

## Sign Languages

- 1st languages of Deaf
- Articulators
  - Hands – 30 % information of utterance
  - Non-Manual features (NMFs) such as squinting of eyes, furrowing of brows, movement of the body – 70 % information of utterance
- Sign in SL similar to *morpheme* in spoken language
- Articulations of hands similar to *phonemes* in speech

- Spoken language is phonetically linear: *not =*
- SL phonemes occur simultaneously: *not =*
  - Handshape:       'n'
  - Orientation:      palm down
  - Location:         RHS neutral signing space
  - Movement:        RH to right
  - NMFs:            furrowed brows, shaking head

- Lack of formally adopted or recognised writing system
- Available systems (Stokoe Notation, HamNoSys, SignWriting) are phonologically based but don't accommodate simultaneous phoneme production or need secondary representation
- Manual annotation fills this gap allowing for descriptive detail of SL video data, caters for the poly-phonemic structure and allows for time-aligned divisions of the annotations, as may be seen in the interface below.
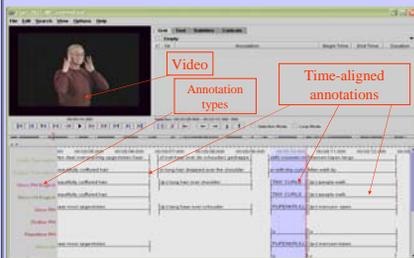


Annotaion user interface using ELAN software

## Example-Based Machine Translation (EBMT)

We want to derive translations for input form a bilingual dataset of source and target sentences. We do this in 3 steps:
1. Matching: close matches to input strings identified in source side of database
2. Alignment: corresponding target language chunks identified
3. Recombination: target language chunks combined to give target language strings

### Similarity Metric: Marker Hypothesis (Green, 1979)
- Segments spoken language sentences according to a set of closed class words (e.g. determiners, prepositions, pronouns)
- Chunks start with a closed class word and usually encapsulate a concept forming concept chunks, e.g. the concept of darkness in *<PRON> it was almost dark*

- Marker Hypothesis not suited for SL side of corpus due to natural lack of closed class lexical items in SLs
- Alternative method was developed to form chunks that are successful in forming potentially alignable chunks
- Time spans of *gloss* tier of annotation used and all other annotations occurring within same time-frame grouped together to form an SL chunk

**English chunk:**            **SL chunk:**
<DET> the hare takes off       <CHUNK>
                              (Gloss)      running hare
                              (Mouth)      closed-ao
                              (Cheeks)     puffing

## Experiments

- Difficult to find suitable SL corpora
- 561 Dutch Sign Language/Nederalndse Gebarenataal (NGT) sentences from annotated data provided by ECHO project to begin with

### Spoken Language → Sign Language

- 90:10 training-testing splits, total 55 test sentences
- Best match found on sentential, sub-sentential (chunk) or word level
- Manual analysis shows central concepts get translated
  **Input:**     They visit real life
  **Output:**    (Gloss) LIFE        (Gloss) VISIT
                (Mouth) 'leven'      (Mouth) 'bezoeken'
                (Eye Gaze) rh        (Dir&Loc) rh

- Output in annotation format cannot be evaluated using traditional, precision-based evaluation metrics such as:
  - BLEU score
  - Word error rate (WER)
  - Position-independent WER (PER)
- Necessary to reverse translation direction for evaluation purposes

### Sign Language → Spoken Language

- NGT → English/Dutch
- Automatic evaluation metrics may be used
- Sample output (a) compared with reference sentence (b):
  a. mouse promised help
  b. "You see," said the mouse, "I promised to help you".

- Central concepts still translated
- Evaluation results:

| BLEU: | 0 |
|---|---|
| WER: | 119% |
| PER: | 78 % |

- Low scores despite correct concept translation
- Precision based metrics don't cater for lack of closed class items in SLs or NMF detail.
- Better to adopt more syntactic- or semantic-based evaluation

### SLMT Data Suitability

- **NGT poetry and fables as data not suited to MT:**
  - Open domain
  - Very descriptive
  - Unrepeated content
  - Colloquial terms
  - Limited practical use

- *We felt we had exhausted the NGT data as much as possible yet had retrieved unsatisfactory results, namely due to the size and domain of the dataset*

- **More suitable dataset:**
  - Closed domain
  - Small vocabulary
  - Repetition of words, phrases and sentences structures
  - Possible practical use for the Deaf

- *For these reason we chose to develop our own corpus, tailored to the requirements of an MT system*



## Irish Sign Language (ISL) Dataset

- Flight Information Corpora = 1429 sentences from dialogue systems
- Almost three times that of the NGT data
  - 577 Air Travel Information System (ATIS) corpus
  - 852 SunDial corpus
- Closed domain, suitable topic for MT, repetition, potentially useful for Deaf

- Data signed by *native ISL signers* for video

- ELAN toolkit used for the hand-annotation of the data at *right and left hand gloss level* and *English translation*

### Experiments to Date

- 400 annotated sentences
- 44 test sentences
- Same MT system as NGT data
- Translating from ISL → English

| BLEU: | 0.06 |
|---|---|
| WER: | 89% |
| PER: | 55 % |

**c. 30 % improvement in WER and PER**

**Higher scores** across all metrics despite smaller training and testing set!

Encouraging results for further work with remainder of ISL data being incorporated.

## Conclusions and Further Work

- Annotation works best for trancribing the simultaneous phonetic structure of SLs

- NGT data shows promising results but falls bacaus of the unsuitability of data

- ISL corpus results show that already the change in domain topic has improved scores dramtically

- NMF information and phonetic detail will be incorporated into annotations so that a fully functioning English to ISL MT system can be developed where the sentences will be signed in ISL by a mannequin such as the one shown below from Poser Software.