Body at Work: Using Corpora in Sign Language Machine Translation

Sara Morrissey
CNGL, Dublin City University, Ireland
20 min presentation 50% own, 50% Qs

- **CHALLENGE 1: SYMBOLIC TRANSLATION**
  - In this area we want to discuss possible approaches for translating a spoken language sentence (in its written form) to a signed utterance representation, e.g. a sequence of glosses.
  - How much linguistics and lexicography is required?
  - How much world knowledge is required?
  - Do you use a corpus to inform the process, and how does the corpus have to look like?
  - What triggers the use of sign-specific means of expression, such as productive signs and classifier constructions, or constructed action?
  - What are good evaluation measures?
  - How much detail must the target signed utterance representation contain?
Corpus-based approach: statistical/example-based

Parallel Corpus
Aligned sentences
Aligned phrases
Aligned words

Decoder

Source text
Target text

- English → Sign Language of the Netherlands (Nederlandse Gebarentaal – NGT)

- ECHO project data (561 sentences)

- Glossing

- Hand-crafted basic example-based machine translation (MT) system

- Basic manual evaluations of gloss output
Preliminary experiments

The hare takes off

(Gloss RH English) (p-) running hare :
(Mouth) closed-ao :
(Mouth SE) /AIRSTREAM/ :
(Cheeks) p :
(Gloss LH English) (p-) running hare :
(Gloss RH) (p-) rennen haas :
(Gloss LH) (p-) rennen haas
Early work: phase 2 (2006-2008)

- English, German, Irish Sign Language (SL), German SL (Deutsche Gebärdensprache - DGS),
- Air Traffic Information System corpus (595 sentences)
- Manually glossed using ELAN (lacked NMF detail)
- Bidirectional MT using MaTrEx
- Avatar database of pre-created videos using Poser software – also addressed lack of NMF detail in glosses
- Manual and automatic evaluations
- Automatic scores: broadly comparable with contemporaries: 35-50% correct
- Manual evaluations: generally well received, but subjective, only 4 people
Example

English input: I’d like a flight

ISL Gloss output: LIKE FLIGHT
Sentence 4

Intelligibility: How would you rate this video sentence in terms of understandable and correct ISL?

- Understood and correct
- Understood but somewhat incorrect
- Difficult to understand but I grasp the gist of it
- Incorrect or too confusing to grasp the meaning

Fidelity: How would you rate the video sentence as a translation of the English?

Please click on the box below to view the original English sentence. English sentence...
Early Attempts: what we learned

- Detailed, multi-level transcriptions
  - cause difficulty with alignments in training and for the evaluation process
  - Are labour intensive to create
  - Few standards, subjective
- Using glossing is ‘cheating’
  - Using one language to represent another is not adequate (Pizutto & Pietrandrea, 2001)
  - Pidgin form of spoken language
  - Becomes computer-assisted translation rather than MT (ambiguity resolved, lexical gaps filled subjectively…)
- Individual pre-made videos don’t make for natural sign transition
- EBMT showed no significant improvement
Current Work

- Patients with Limited English (2008 – present)
- English → Irish Sign Language
- Multimodal patient—medical-receptionist appointment booking corpus (396 sentences)
- Manual transcription using HamNoSys
- Animated avatar using software of University of East Anglia
- Corpus collection complete, exploratory MT beginning
- The results are amazing! The problem is solved and we can all go home now!
Question: how do we maintain expressibility and content in the SL but keep text to minimum for MT?

English: hello, can i make an appointment.

ISL: sid5 sid14 sid117 sid27 sid118 sid7
Transcription of utterance in SiGML

<utterance spoken="1.26 Hello, can I make an appointment.">
  <sign gloss="HELLO_AS_2" signid="5"><mouth>hiUUUO</mouth><src editable="false"/><gol editable="false"/>
  <loc editable="false"/><hand/><limbs/><facialexpression/>
  <hamnosys>hamflathand,hamextfingerul,hampalmr,hamforehead,hamlirside,hamseqbegin,hamtouch,hamindexfinger,hamseqend,hamreplace,hamfinger2345,hamfingernbendmod,hambetween,hamflathand,hamextfingerol,hambetween,hamextfingeru,hampalmr,hamshoulders,hambetween,hamshoulders,hamlrat,hamextendend</hamnosys>
</sign>

<sign gloss="GST-GET_ATTENTION" signid="43"><mouth><src editable="false"/><gol editable="false"/><loc editable="false"/><hand/><limbs/><facialexpression/>
  <hamnosys>hamflathand,hamthumboutmod,hambetween,hamfinger2345,hamextfingeruo,hamextfingerol,hambetween,hamextfinger,hampalmr,hamshouldertop,hamarmextendend,hamparbegin,hamnodding,hamfingerplay,hamparend,hamrepeattfromstart</hamnosys>
</sign>

<sign gloss="I_VS_3" signid="117"><mouth>m</mouth><src editable="false"/><gol editable="false"/><loc editable="false"/><hand/><limbs/><facialexexpression/>
  <hamnosys>hamfinger2,hamthumbacrossmod,hambetween,hamfinger2,hamthumbacrossmod,hamfngstrectmod,hamindexfinger,hampalmr,hamshoulders,hamclose,hammovei,hamsmallmod</hamnosys>
</sign>

<sign gloss="BOOK_appointment" signid="7"><mouth>@pOm@</mouth><src editable="false"/><gol editable="false"/><loc editable="false"/><hand/><limbs/><facialexpression> FU </facialexpression>
  <hamnosys>hamsymmpar,hamflathand,hamparbegin,hamextfingeru,hamplus,hamextfingerol,hamparend,hampalmr,hamparbegin,hampalmr,hamclose,hamlips,hamhclose,hamplus,hamchest,hamparend,hamseqbegin,hamparbegin,hamseqbegin,hamreplace,hamextfingeru,hampalmr,hamchest,hamseqend,hamplus,hamnomotion,hamparend</hamnosys>
</sign>

<sign gloss="CAN" signid="52"><mouth>k}n</mouth><src editable="false"/><gol editable="false"/><loc editable="false"/><hand/><limbs/><facialexpression> RB WB </facialexpression>
  <hamnosys>hamfinger23,hamextfingeru,hampalmr,hamshoulders,hamclose,hamparbegin,hamreplace,hamextfingerol,hamparend</hamnosys>
</sign>
Encapsulating data

<sign gloss="HELLO_AS_2" signid="5">
  <mouth>hiUUUO</mouth><src editable="false"/><gol editable="false"/> <loc editable="false"/> <hand/>
  <limbs/><facialexpression/>
  <hamnosys>hamflathand,hamextfingerul,hamextent,between,hamtouched,head,hamlrbeside,hamseqbegin,hamtouch,hamshoulder,hamseqend,hamreplace,hamfinger2345,hamfingerbetween,hambetween,hampalml,hamshoulders,hambetween,hamshoulder,hamflathand,hamextfingerol,hambetween,hamextfingerul,hamextended</hamnosys>
</sign>
Current Work Considerations

- Do we maintain expressibility and content of the SL?
- Need to better investigate the MT process, the corpus and see how we can encode and transfer this information across
- Linguistic information will not come through via ID tags, may need to alter this to encode the information for EBMT
- HamNoSys offers a more comprehensive, faithful description of an SL than glossing
- Using the UEA avatar tool, MT can now be performed on-the-fly and without pre-loaded videos that can be unrealistic.
- Don’t yet know what we’re missing in terms of sign-specific linguistic information…let the fun begin!
Addressing the Challenges in the Context of Corpus-based Machine Translation

- The corpus, its form (notation) and what we can glean from it
- Output format and evaluation
A corpus is a highly valuable linguistic resource
- For linguistic analysis and collection, for seeding data-driven MT
- But the problem is getting enough AND in the same format
- It can take ages to create one
- Solution: central repository and standards?

Do we take advantage of glossing and how it makes MT easier?

Or do we consider it potential misrepresentation and choose an alternative format?
How much detail must the target signed utterance representation contain?
- Consider representation as transcribed, text-based annotation:
  - Less detail = easier translation
  - More detail = better animation
- Consider representation through an avatar.
  - Animated representation should be fully articulate, human-like, competently use non-manual features and the signing space accurately.
  - Realism of avatars in relation to the ‘uncanny valley’, how realistic can an avatar be before it gets disconcertingly real? Where is the line of acceptability?

Is perfect output necessary?
- Mainstream MT
- Gisting
- Not perfect, but maybe helpful
- Not trying to replace interpreters, and it shouldn’t!
Evaluation

What are good evaluation measures?

Automatic evaluation
- Not yet possible for avatars
- BLEU and error rates used in mainstream spoken language MT can be adopted for transcription-based output
- Internal progress of the MT can be compared

Human evaluation
- Artificial for transcriptions
- Imperative – evaluate actual signing
- Evaluates system as a whole
- Subjective evaluation
Conclusions

- We all need corpora – the more the better
  - Wouldn’t a central repository be great?
- Transcription should faithfully represent the SL…but:
  - The simpler it is, the less faithful it is?
  - The more complex it is, the more difficult it is for MT?
- Evaluation is imperative
  - Automatic and manual
  - Whole pipeline or individual components
  - Guidelines/rules set out
- Lots done, more to do!
Thank you/ Go raibh maith agaibh

Questions/Ceisteanna?

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (www.cngl.ie) at Dublin City University