Assistive Translation Technology for Deaf People: Translating into and Animating Irish Sign Language

Sara Morrissey

National Centre for Language Technology
School of Computing, Dublin City University
Dublin 9, Ireland

Abstract. Machine Translation (MT) for sign languages (SLs) can facilitate communication between Deaf and hearing people by translating information into the native and preferred language of the individuals. In this paper, we discuss automatic translation from English to Irish SL (ISL) in the domain of airport information. We describe our data collection processes and the architecture of the MaTrEx system used for our translation work. This is followed by an outline of the additional animation phase that transforms the translated output into animated ISL. Through a set of experiments, evaluated both automatically and manually, we show that MT has the potential to assist Deaf people by providing information in their first language.

1 Introduction

Communication is the essence of human interaction. However, this natural means of communication can prove a barrier in cases where languages differ. This is nowhere more evident than in the communication barrier between languages of different modalities, namely spoken and signed languages.

SLs are fully formed, natural and eloquent languages. This coupled with the right for all humans to communicate in a means that is most natural to them signifies that Deaf people should be able to communicate in their preferred language. But in Ireland, just over 1% of the Irish non-Deaf population know Irish Sign Language (ISL) (Leeson, 2003). Furthermore, most Irish Deaf people are not confident in their English language skills (Conroy, 2006). These can greatly hamper daily communication between the Deaf and hearing worlds. In addition, there is little by way of public services available through SLs and most of which is available is on a one-to-one basis (Ó'Baoill & Matthews, 2000).

SL interpreters (SLIs) play a valuable role in breaking down the barriers but they are often difficult to come by, particularly in Ireland where the ratio of interpreters to Deaf people is about 1:250 (Leeson, 2003). SLIs can be an expensive and impractical option if only required for a short period of time. There are also issues of confidentiality in legal or medical matters, for example.

Technological advancements such as teletype systems and subtitling can help overcome the auditory barrier, but they assume good literacy skills and speed of reading and understanding on the part of the users in their second language (Huenerfauth, 2006). In addition, often, for reasons of time and space they can be simplified resulting in a loss of information.

There is, however, one particular type of technology that can overcome the impractical and confidential issues regarding SLIs and allow Deaf people to access information and communicate through SLs, this is machine translation (MT). Over the last 60 years MT has been used to bridge communication between spoken languages, with data-driven approaches taking centre-stage in recent years. In this paper, we will discuss the development of our data-driven MT system for SLs, focussing on English–ISL translation for airport announcements. We include an overview of the state of the art and discuss the expansion of our work to include a computer-generated signing avatar.

2 State of the Art

Data-driven MT is the dominant methodology in the field of automatic translation, and allows for robust, extensible, broad-coverage translation. In general, linguistic information is eschewed in favour of probabilistic models collected from large parallel corpora. MT for SLs has been predominantly rooted in rule-based approaches, with significant work in this area only emerging within the last 20 years, including the Zardoz system (Veale et al., 1998) and the ViSiCAST project (Marshall & Sáfár, 2002).

There has been little work by way of data-driven MT, the dominant methodology for spoken language MT, for SLs. Three primary systems have been developed:

- (Bauer et al., 1999) employ a statistical approach translating German SL to German using gesture recognition technology. Their work falls somewhat outside the scope of the MT described in this paper, focussing on the opposite translation direction.
- (Stein et al., 2006) have attempted translation both to and from SLs using a phrase-based statistical MT (SMT) system. While they have incorporated some extra technological features such as gesture recognition and avatar production, their focus remains on MT and investigating SMT approaches.
- (Wu et al., 2007) combine rule-based and data-driven approaches in order to achieve translations for Chinese to Taiwanese SLs. Through using a rule-based approach their system lacks extensibility to new language pairs and directions. Also the system does not produce sign via avatar.

3 Methodology

Our methodology comprises three parts: data collection, data-driven MT and animation of output.

3.1 Data Collection

A pre-requisite for data-driven MT is a bilingual corpus of data. Suitable data are difficult to find and those available tend to focus on non-domain specific information, such as the children's stories of the European Cultural Heritage Online (ECHO) project corpus[1], that is not suitable for MT. Furthermore, data that contains a text-based representation of SLs are uncommon. Having performed previous experiments on the ECHO data with some success, we chose to create our own purpose-built corpus. We chose the Air Traffic Information System (ATIS) corpus (Hemphill et al., 1990) consisting of 595 sentences of English dialogue for its practical domain and suitability for MT. In order to create an authentic parallel ISL corpus we engaged the assistance of two native ISL signers to work in tandem translating, signing and monitoring the creation of our video corpus. The videos were then manually annotated with semantic glosses to create a parallel ISL 'text'. Due to the laborious annotation process, the videos were only annotated at the level of the hands meaning some important non-manual feature detail is lost. The EUDICO Linguistic Annotator (ELAN)[2] was used for annotation. An example of the annotated videos is shown in Figure 1. The resulting bitext is then used to feed the next phase: automated translation.

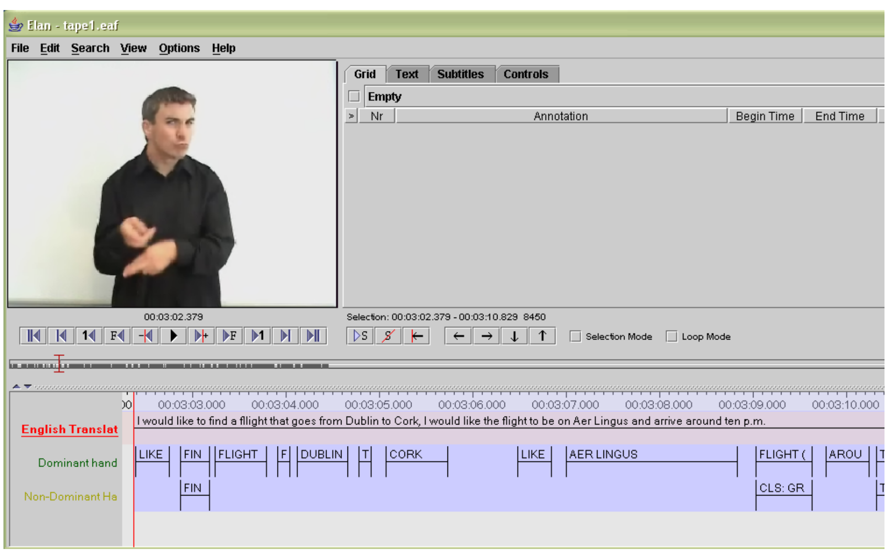

Fig. 1. MaTrEx Architecture

---
[1] http://www.let.kun.nl/sign-lang/echo/
[2] http://www.mpi.nl/tools/elan.html

3.2 The MaTrEx Data-driven MT System

In order to translate English text into animated ISL, we employ the MaTrEx data-driven MT system (Stroppa & Way, 2006). It is a hybrid system, developed at the National Centre for Language Technology, Dublin City University, that avails of both example-based MT (EBMT) and SMT approaches. A diagram of the system architecture is shown in Figure 2.

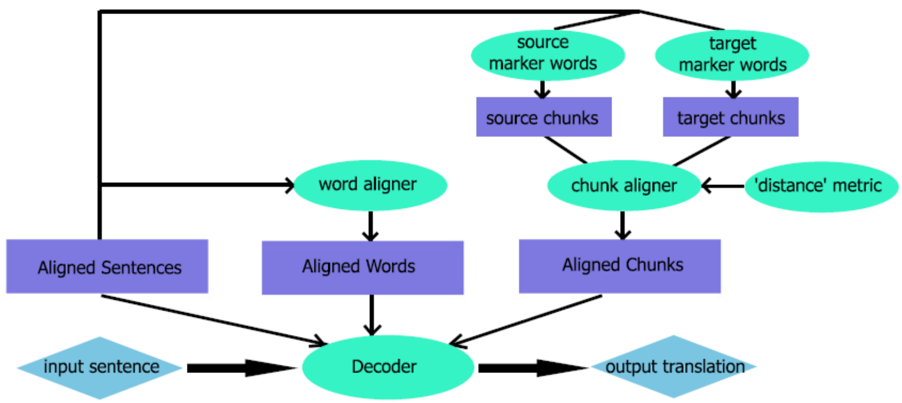

Fig. 2. MaTrEx Architecture

The decoder is the main engine that takes an input sentence and produces a candidate translation. It is fed by three bilingual data resources, namely aligned sentences, sub-sentential chunks and words. Each is derived from the bilingual corpus described in the previous section. The decoder employs these bilingually aligned resources as databases of translation examples against which it compares input data. Translation links are retrieved for the matches found and are recombined to produce the candidate target language translation. In the case of English to ISL translation, the output takes the form of glossed annotations. These are then fed into the final animation phase.

3.3 The Animation Process

For an SL MT system to be of practical use to the Deaf community it requires real SL to be produced rather than annotation. Real human signing, while preferred by most (Naqvi, 2007), is not a practical option for signed output. We chose to use computer-generated signing mannequins to animate the ISL. We used Poser Animation Software Version 6.0[3] to create our animated avatar. Individual videos were manually constructed for each sign using the positioning

[3] http://www.curiouslabs.de/poser6.html?&L=1

tools. It was not possible to produce fluid, seamless transitions between each signed video during sentence construction for this work. For this reason, there is a short pause between each sign where the mannequin returns to a neutral face and hand position. While we recognise this is somewhat unnatural, we hope to rectify this in the future. The animated signing avatar is shown in Figure 3.

Fig. 3. The Animated Avatar

4 Experiments and Results

In the first phase, using the ATIS data described above, we feed the MaTrEx system. Three sub-experiments were run using different chunking methodologies in an attempt to improve translation. The baseline SMT system does not include any additional EBMT chunks. *Type 1* chunks are formed using the Marker Hypothesis and *Type 2* chunks are formed using the Marker Hypothesis for the English data and segmenting the ISL data according to the semantic sign glosses. In order to objectively assess the quality of the candidate translations produced independent of the avatar module, we employed automatic evaluation metrics, namely BLEU, word error rate (WER) and position-independent WER (PER). The results are shown in Table 1. The higher the BLEU score the better, the lower the error rate scores the better.

	System	**BLEU**	**WER**	**PER**
EN–ISL	MaTrEx Baseline	38.85	46.02	34.33
	Baseline + Type 1 Chunks	39.11	45.90	34.20
	Baseline + Type 2 Chunks	39.05	46.02	34.21

Table 1. Evaluation Scores for EN–ISL experiments

We can see here that Type 1 chunks perform the best for this language pairing by improving on the BLEU score and decreasing the error rates. A WER score of 45.90 for Type 1 chunks indicates that the system gets 54.10% of translated words correct and in the correct order.

In the second phase of our experiments, we fed the annotated output from the MT system into the avatar module to produce animated sign. The pre-made avatar videos for each sign in the translation are joined together to form a signed sentence. We engaged Deaf native ISL signers to manually evaluate the animated output for intelligibility and fidelity. 82% were considered intelligible by the monitors and 72% were considered good-to-excellent translations.

5 Impact and Contributions

This work addresses an MT methodology that has never been applied to SL MT to date, namely EBMT. Furthermore this approach describes a complete English to animated ISL translation system that forms part of a bi-directional multilingual MT system to assist in communication and comprehension for members of the Irish Deaf community.

6 Conclusions and Future Work

We have shown through experiments and comprehensive automatic and manual evaluation, that EBMT is a suitable and successful approach for translating into SLs in the practical domain of airport announcements. Also the addition of an avatar module enables the production of output in the preferred language of the Irish Deaf Community.

Future work includes researching improved chunking methodologies to improve translations, increased avatar development to enhance naturalness of signing and the extension of the MaTrEx system to a fully functioning bi-directional SL MT system, including provision for SL–SL translation.

References

B. Bauer, S. Nießen and H. Heinz. 1999. Towards an Automatic Sign Language Translation System. In *Proceedings of the International Workshop on Physicality and Tangibility in Interaction: Towards New Paradigms for Interaction Beyond the Desktop*, Siena, Italy.

P. Conroy. 2006. *Signing In & Signing Out: The Education and Employment Experiences of Deaf Adults in Ireland. Research Report.* Irish Deaf Society, Dublin.

C. Hemphill and J. Godfrey and G. Doddington. 1990. The ATIS Spoken Language Systems Pilot Corpus. In *Proceedings of the Workshop on Speech and Natural Language*, Hidden Valley, PA, pp. 96–101.

M. Huenerfauth. 2006. *Generating American Sign Language Classifier Predicates for English-to-ASL Machine Translation.* PhD Thesis, University of Pennsylvania, Philadelphia, PA.

L. Leeson. 2003. *M. Cronin and C. Ó Cuilleanáin(Eds.), Languages of Ireland.* Four Courts Press: Dublin, Ireland. Chapter 8, pp. 148–164.

S. Naqvi. 2007. End-User Involvement in Assistive Technology Design for the Deaf Are Artificial Forms of Sign Language Meeting the Needs of the Target Audience? In *Proceedings of the Conference and Workshop on Assistive Technologies for People with Vision & Hearing Impairments (CVHI)*, Granada, Spain, [no page numbers].

D. Ó'Baoill and P. A. Matthews. 2000. *The Irish Deaf Community (Volume 2): The Structure of Irish Sign Language*. The Linguistics Institute of Ireland, Dublin, Ireland.

I. Marshall and É. Sáfár. 2002. Sign Language Generation using HPSG. In *Proceedings of the 9th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-02)*, Keihanna, Japan, pp.105–114.

D. Stein, J. Bungeroth and H. Ney. 2006. The Architecture of an English Text-to-Sign Languages. In *Proceedings of the 11th Annual Conference of the European Association for Machine Translation (EAMT, '06)*, Oslo, Norway, pp. 169–177.

N. Stroppa and A. Way. 2006. MaTrEx: DCU Machine Translation System for IWSLT 2006. In *Proceedings of the International Workshop on Spoken Language Translation (IWSLT)*, Kyoto, Japan, pp. 31–36.

T. Veale, A. Conway, and B. Collins. 2000. The Challenges of Cross-Modal Translation: English to Sign Language Translation in the Zardoz System. *Machine Translation* **13**(1):81–106.

C-H. Wu and H-Y. Su and Y-H. Chiu and C-H. Lin. 2007. Transfer-Based Statistical Translation of Taiwanese Sign Language Using PCFG. *ACM Transactions on Asian Language Information Processing (TALIP)*, **6**(1):[no page numbers].