An Assessment of Appropriate Sign Language Representation for Machine Translation in the Healthcare Domain

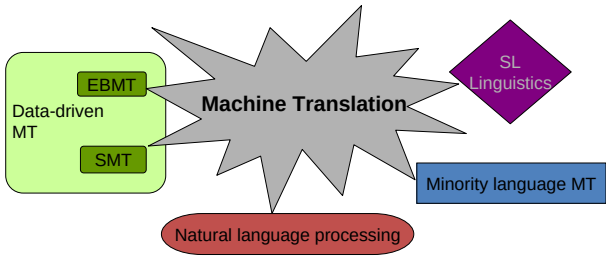
Sara Morrissey

Centre for Next Generation Localisation (CNGL), School of Computing, Dublin City University, Ireland

Project Background

Prescription:

- Assistive cross-communication technology
- Native language use
- Cost-effective
- Focus on high translation quality
- Medical secretary interactions



Prerequisite: corpus

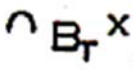
Transcription method

N.B. - Sensitive context
- Domain specific data
- Maximal representation of SLs
- Computability of representation

L–I–N–E–A–R

HamNoSys

Stokoe Notation



- ☑ detailed phonetic description
- ☒ possible unnecessary use of symbols for MT

MULTI–LEVEL

Linguistic Annotation

- ☑ comprehensive, flexible
- ☒ subjective, uses one language to describe another

↑↳☺∩↑↳

SignWriting

- ☑ concise
- ☒ not widely used/not currently machine readable

Discussion

A full representation is of prime importance for both SLs and MT. The extent of the language must be honoured, but also full phonetic detail is required for complete MT systems in order to facilitate on-the-fly animation. Missing detail means reduced accuracy, a quality that is imperative in an MT system focusing on the healthcare domain.

All the above representation approaches can provide phonetic detail necessary, but it is the coded version of this detail rather than the symbols themselves that is of use to the MT system.

HamNoSys, being the most developed of the linear systems, machine readable, and having individual phonemes described that could be used in the animation process, it appears the most appropriate for current MT.

We envision that if SignWriting was widely accepted, used and machine readable, it could positively facilitate the translation process by containing compound information in one symbol, and removing the need for the complex animation process.

References

Huenerfauth, M. 2005. American Sign Language Generation:Multimodal NLG with Multiple Linguistic Channels. In *Proceedings of Information Technology and Telecommunications Conference 2006*, Ann Arbor, MI., pp.37-42.

Leeson, L., J. Saeed, A. Macdu, D. Byrne-Dunne and C. Leonard. 2006. Moving Heads and Moving Hands: Developing a Digital Corpus of Irish Sign Language. In *Proceedings of the ACL Student Research Workshop (ACL 2005)*, Carlow, Ireland.

Liddell, S. and R. E. Johnson. 1989. American Sign Language: The Phonological Base. In *Sign Language Studies*, 64 pp.195-277.

Prillwitz, S., R. Leven, H. Zienert, T. Hanke and J. Henning. 1989. *HamNoSys. Version 2.0. Hamburg Notation System for Sign Languages: An Introductory Guide*, Hamburg Signum, Hamburg, Germany.

Morrissey, S. and A. Way. 2006. Lost in Translation: the Problems of Using Mainstream MT Evaluation Metrics for Sign Language Translation. In *Proceedings of the 5th SALTMIL Workshop on Minority Languages at LREC'06*. Genoa, Italy, pp.91{98.

Morrissey, S. 2008. *Data-Driven Machine Translation for Sign Languages*. PhD Thesis. Dublin City University.

Somers, H. 2007. Theoretical and Methodological Issues Regarding the Use of Language Technologies for Patients with Limited English Proficiency. In *Proceedings of the Eleventh Conference on Theoretical and Methodological Issues in Machine Translation (TMI-08)*, Skövde, Sweden, pp.206-213.

Stein, D., J. Bungeroth and H. Ney. 2006. Morpho-Syntax Based Statistical Methods for Sign Language Translation. In *Proceedings of the 11th Annual conference of the European Association for Machine Translation (EAMT'06)*, Oslo, Norway, pp.169-177.

Stokoe, W. C. 1960. An Outline of the Visual Communication Systems of the American Deaf. In *Studies in Linguistics: Occasional papers*, No. 8,Department of Anthropology and Linguistics, University of Buffalo, Buffalo, New York, pp.78.

Sutton, V. 1995. Lessons in Sign Writing, Textbook and Workbook (Second Edition). *The Center for Sutton Movement Writing*, Inc., La Jolla, CA.

This research is funded by Science Foundation Ireland (SFI) as part of the National Development Plan (NDP) 2007-2013