

Overview of Text-to-Gloss in Computational Sign Language Processing (SLP)

Achraf Othman

IEEE Senior Member

achraf.othman@ieee.org

Digital Accessibility to the content in web environments for people with hearing disabilities and with hearing impairment with a low level of literacy is becoming increasingly critical (Dena et al., 2020, Lahiri et al., 2020). Several applications have been developed to solve this challenge (Othman et al., 2019). Today, the existing solutions proposed by the literature to improve digital accessibility to web content and media for people with hearing disabilities remain limited and even not existing in Sign Language (SL) (Jemni et al., 2013). Some of these applications have been developed to set up conversational agents based on three-dimensional technologies called avatars for translating a written text to SL. All these tools do not take into consideration the intonation and rhythm of the uttered words (Kipp et al., 2011). This greatly reduces the quality of the translated web content for these people and can even make it incomprehensible only if we use video translating the content in SL by interpreters or Deaf users where the cost is very high. In any Sign Language, we're facing more challenges and issues based on the nature of the language itself. Shifting and interpreting from written text to SL requires many levels of linguistic processing toward reaching the cognitive meaning of the sentence.

Transcription is the operation that substitutes a grapheme or a group of graphemes of a writing system for every phoneme or for every sound. It thus depends on the target language, a unique phoneme that can correspond to various graphemes following the considered language. In short, it is the writing of words or pronounced sentences in a given system. The transcription also aims at being without loss, so that it should be ideally possible to reconstitute the original pronunciation from this one by knowing the rules of transcription.

Text-to-gloss—also known as sign language translation—is the task to translate between a spoken language text and sign language glosses. Zhao et al. (2000) used a Tree Adjoining Grammar (TAG) based system for translating between English sentences and American Sign Language glosses. They parse the English text and simultaneously assemble an American Sign Language gloss tree, using Synchronous TAGs (Shieber and Schabes 1990; Shieber 1994), by associating the ASL elementary trees with the English elementary trees and associating the nodes at which subsequent substitutions or adjunctions can take place. Synchronous TAGs have been used for machine translation between spoken languages (Abeillé, Schabes, and Joshi 1991), but this is the first application to a signed language.

For the automatic translation of gloss-to-text, Othman and Jemni (2012) identified the need for a large parallel sign language gloss and spoken language text corpus. They develop a part-of-speech-based grammar to transform English sentences taken from the Gutenberg Project ebooks collection (Lebert 2008) into American Sign Language gloss. Their final corpus contains over 100 million synthetic sentences and 800 million words and is the largest English-ASL gloss corpus that we know of. Unfortunately, it is hard to attest to the quality of the corpus, as they didn't evaluate their method on real English-ASL gloss pairs, and only a small sample of this corpus is available online.

References

- Abeillé, Anne, Yves Schabes, and Aravind K Joshi. 1991. "Using Lexicalized Tags for Machine Translation."
- Al Thani, D., Al Tamimi, A., Othman, A., Habib, A., Lahiri, A., & Ahmed, S. (2019, December). Mada Innovation Program: A Go-to-Market ecosystem for Arabic Accessibility Solutions. In 2019 7th International Conference on ICT & Accessibility (ICTA) (pp. 1-3). IEEE.
- Jemni, M., Semreen, S., Othman, A., Tmar, Z., & Aouiti, N. (2013, October). Toward the creation of an Arab Gloss for Arabic Sign Language annotation. In Fourth International Conference on Information and Communication Technology and Accessibility (ICTA) (pp. 1-5). IEEE.
- Kipp M., Heloir A., Nguyen Q., Sign Language Avatars: Animation and Comprehensibility, International Workshop on Intelligent Virtual Agents, pp 113-126, 2011.
- Lahiri, A., Othman, A., Al-Thani, D. A., & Al-Tamimi, A. (2020, September). Mada Accessibility and Assistive Technology Glossary: A Digital Resource of Specialized Terms. In ICCHP (p. 207).
- Lebert, Marie. 2008. "Project Gutenberg (1971-2008)." Project Gutenberg.
- Othman, Achraf, and Mohamed Jemni. 2012. "English-Asl Gloss Parallel Corpus 2012: Aslg-Pc12." In *5th Workshop on the Representation and Processing of Sign Languages: Interactions Between Corpus and Lexicon LREC*.
- Othmane A. And Jemni M., Designing High Accuracy Statistical Machine Translation for Sign Language Using Parallel Corpus: Case Study English and American Sign Language, Journal of Information Technology Research (JITR), pp 12–2, 2019.
- Shieber, Stuart M. 1994. "RESTRICTING the Weak-Generative Capacity of Synchronous Tree-Adjoining Grammars." *Computational Intelligence* 10 (4): 371–85.
- Shieber, Stuart, and Yves Schabes. 1990. "Synchronous Tree-Adjoining Grammars." In *Proceedings of the 13th International Conference on Computational Linguistics*. Association for Computational Linguistics.
- Zhao, Liwei, Karin Kipper, William Schuler, Christian Vogler, Norman Badler, and Martha Palmer. 2000. "A Machine Translation System from English to American Sign Language." In *Conference of the Association for Machine Translation in the Americas*, 54–67. Springer.