

Evaluating the Alignment of Utterances in the Swedish Sign Language Corpus

Carl Börstell
University of Bergen
carl.borstell@uib.no

Introduction

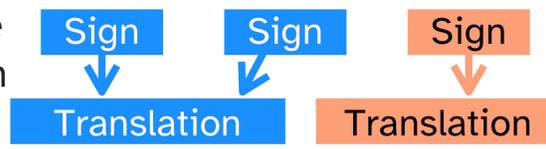
- There is **no segmentation** of the STS Corpus¹ beyond the individual sign glosses
- Some work^{2,3} used **translation-tier** segmentations, or **pauses** as segment-breaks, to **infer utterance units**
- The **syntactic UD annotation** of the STS Corpus in principle = syntactic units

How do these all align?

Inferring units

Translation units do not align exactly with sign units, assignment is approximate (see figure)

Prosodic units = pause between signs > median between-sign duration



Syntactic units = segmentation in the STS UD data⁴

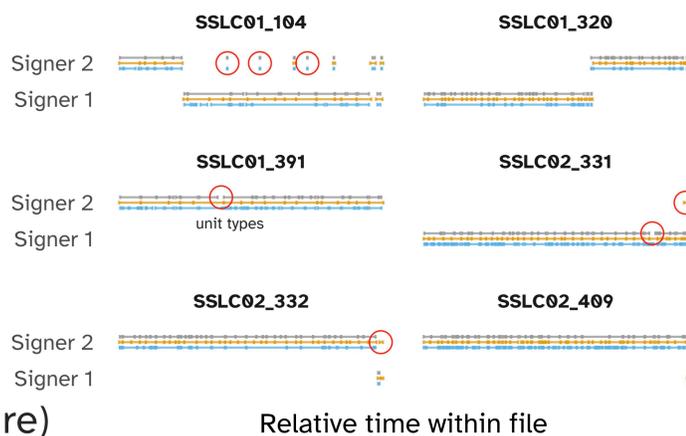
NB: Based only on a subset of the STS corpus!

Correspondence

Utterance units do not always align. Some gaps between unit types (see figure) – for instance, not all manual back-channels have translations, but some non-manual back-channels have translations

Alignment of syntactic, translation and prosodic utterance units

Major gaps and discrepancies marked with red circles



Endpoints of utterance units are not always aligned (see figure)

MediaPipe metric

The subset video files were analyzed with MediaPipe⁷ computer vision software: the **movement of the head and hands** was compared to locations with **endpoints** across utterance unit types (see figure)

No clear alignment between computer vision-based activity and endpoints across utterance units

Relative movement of hands and head

Landmark coordinates for wrists (solid) and nose (dotted) of signers 1 & 2: vertical bars show segmentation points aligned across utterance units

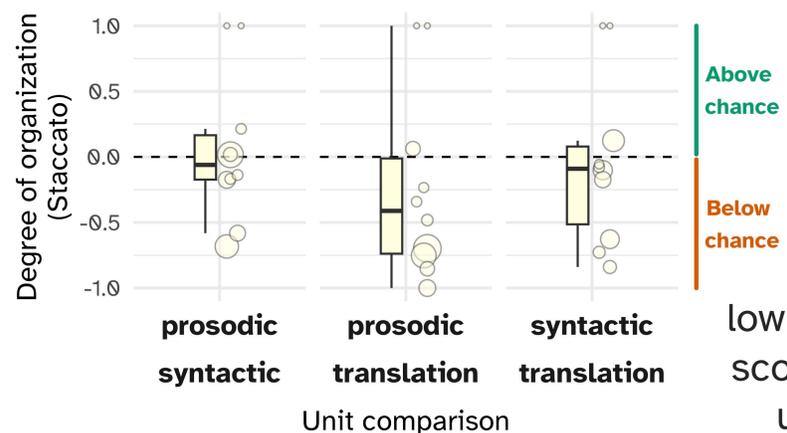


Alignment score

Alignment across unit types was calculated as *degree of organization* with the Staccato method⁵ in ELAN⁶

Degree of organization between utterance units

Distribution of scores for each unit comparison: each circle represents one tier comparison



≈1600 signs
≈200 units

Overall, low alignment scores across unit types...

Conclusions

- ◆ **Low degree of alignment** across inferred utterance units in a subset of the STS Corpus
- ◆ **Little overlap** between endpoints of inferred utterance units' & CV-extracted "prosody"
- ◆ **Preliminary study** based on a **subset** of the STS corpus: more data is crucial for validity!

References

1. Mesch, J., Wallin, L., Nilsson, A.-L. & Bergman, B. 2012. Dataset. Swedish Sign Language Corpus project 2009–2011 (v.1)
2. Sjons, J. 2013. Automatic induction of word classes in Swedish Sign Language. Master's thesis, Stockholm University
3. Östling, R., Börstell, C. & Wallin, L. 2015. Enriching the Swedish Sign Language Corpus with Part of Speech Tags Using Joint Bayesian Word Alignment and Annotation Transfer. In Proceedings of the 20th Nordic Conference of Computational Linguistics (NoDaLiDa 2015), 263–268, Vilnius: Linköping University Electronic Press
4. Östling, R., Börstell, C., Gärdenfors, M. & Wirén, M. 2017. Universal Dependencies for Swedish Sign Language. In Proceedings of the 21st Nordic Conference on Computational Linguistics, 303–308, Gothenburg: Association for Computational Linguistics
5. Lücking, A., Ptock, S. & Bergmann, K. 2011. Staccato: Segmentation Agreement Calculator according to Thomann. In Proceedings of the 9th International Gesture Workshop: Gestures in Embodied Communication and Human-Computer Interaction, 50–53, Athens.
6. Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. 2006. ELAN: A professional framework for multimodality research. In Proceedings of the 5th International Conference on Language Resources & Evaluation (LREC 2006), 1556–1559
7. Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M.G., Lee, J., Chang, W.-T., Hua, W., Georg, M. & Grundmann, M. 2019. MediaPipe: A Framework for Building Perception Pipelines. Publisher: arXiv Version Number: 1.



UNIVERSITY OF BERGEN