The DGS-Korpus approach to including frequent sign combinations in a corpus-based electronic sign language dictionary

Gabriele Langer, Anke Müller, Sabrina Wühl, Thomas Hanke
University of Hamburg, Institute of German Sign Language and Communication of the Deaf

Analysis of Frequent Sign Combinations

- Statistics (Mutual Information) helps to find candidates for frequent sign combinations.
- In the MI-query phonological and lexical variants are treated as instantiations of the same potential neighbour. Reason: fine-grained lemmatisation with regard to form and high lexical variation often result in small numbers of tokens per type.
- For lexicographic purposes the focus is not only on specific collocations but also on semantic preferences.
- Results provide a rough first impression but need to be evaluated and selected in the analysis stage.
- Candidates for frequent sign combinations help to identify different senses of a sign and support the process of Word Sense Discrimination (WSD).

Collocational Patterns

- Combinations of frequent neighbours.
- Meaning of combination can usually be inferred from individual meanings of both partners.
- Theory: One partner, the collocator, is usually less predictable and more restricted as to combinability; its defined and understood completely only through the base (comp. Svensén 2009, 161-162).
- Practice: In our dictionary we also include some selected frequent free combinations in the collocational patterns slot when they are chunks that are useful for clarifying the meaning of the respective sense or its typical usage to the learner.
- Collocational patterns are specific to a sense and therefore addressed as information to the learner.
- Collocational patterns slot when they are chunks that are useful for clarifying the meaning of the respective sense or its typical usage to the learner.

Phrases

- DGS-specific combinations of signs that as a whole have a separate meaning, usually not inferable from the meanings of the individual parts of the phrase or surpassing their combined meanings (principle of idiomaticity).
- One mouthed word or mouth gesture often stretches across both partners.
- The combined meaning of the signs is not transparent. Thus it is represented as a separate sense in the entry.

Loan Compounds and Compound-like Compositions

- Sign combinations that mirror the sequence and structure of German compounds; the German compound is usually simultaneously mouthed.
- Individual signs of the combination strongly correspond to individual parts of the mouthing but several different variants may be used for each element.
- Loan compounds are dominated by the German compound visible in the mouthing.
- A selection of compounds and compound-like compositions found in the corpus are documented as run-ons in the entry.

Within the query results three different types of sign combinations can be found:
- phrase-like structures
- collocational patterns
- loan compounds and compound-like compositions