

Comparison of German Semantic Verb Classes



Sabine Schulte im Walde & Katrin Erk
 Computational Linguistics
 Saarland University, Saarbrücken
 {schulte,erk}@coli.uni-sb.de

Motivation

Example: *eilen, hasten* 'rush, hurry'
 → *Manner of Motion* : *Rush* (Schulte im Walde, 2003)
 → *Manner of Motion* : *Self Motion* (FrameNet, SALSA)

Computational Perspective:

- Acquisition of semantic verb classes
- Use of semantic verb classes for refinement and generalisation

Central Questions:

- Why are there so many verb classifications?
- Why and how do they differ?
- Is there any 'optimal' classification?

Criteria for Comparison ↔ Case Study on Manner of Motion

Background: Motivation & Goals

Structure: Organisation of classes
 Relations between classes

Decision Criteria: Verb sense distinction
 Grouping verbs into classes

Central Meaning Dimensions

| Type of mover | animate vs. inanimate | BB | GN | FN | SIW |
|------------------------|-------------------------|----|----|----|-----|
| Prominent criteria | group vs. individual | | GN | FN | |
| | source/path/goal | BB | GN | FN | SIW |
| | noise during motion | BB | GN | | |
| Movement in place | speed | BB | GN | | SIW |
| | vehicle | BB | GN | FN | SIW |
| | moving in place | BB | GN | FN | SIW |
| | body movement | BB | GN | FN | |
| Accompaniment | iterative movement | BB | GN | | SIW |
| | accompaniment/chase | BB | | FN | |
| Idiosyncratic criteria | cause motion | BB | GN | FN | SIW |
| | propel | | GN | | |
| | travel (long journey) | | | FN | |
| | movement by gravity | | | FN | |
| | erroneous movement | BB | | | |
| | preparation of movement | BB | | | |
| | reason for movement | BB | | | |
| non-movement | BB | | | | |

Examples of Verb Assignment

anschauen 'look at' - **BB**: idiosyncratic classification into *active motion* model *Aktivbewegung* in subclass *bemustern* 'judge' - **GN**: hyponym of *perception* verb *sehen* 'see' - **FN**: *perception active*.

fallen 'fall' - **BB**: *motion / erroneous motion* - **GN**: *motion with path* specified as *vertical* - **FN**: *directional motion*, which is motion determined by natural forces such as gravity.

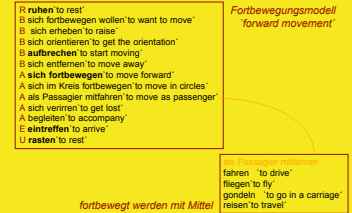
sitzen 'sit' - **BB**: *rest phase* in *motion* models - **GN**: *position* verb under *rest* - **FN**: *posture* describing stable body posture of agent - **SALSA**: *being situated*, the (geographic) object position - **SIW**: *position* verb *be in position*.

Findings and Conclusions

- Central vs. idiosyncratic features
 - Strengths and weaknesses in each classification
 - There is no 'true' semantic verb classification. One word may instantiate one or several dimensions of meaning.
- A combination of several resources is *advisable* and *promising*.
 → Application example: gold standard for cluster analysis:
ermorden, töten, verhaften, befragen, entlassen

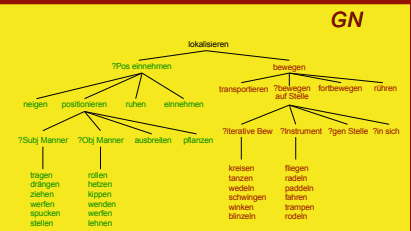
Process-based Classification

- **Goal:** thesaurus
- **Choice:** 8,000 non-prefixed verbs
- **Assignment to classes:** based on paraphrases
- **Organisation:** process models
- **Model:** start, transition, end; preconditions, results, consequences
- **Class relations:** causal, temporal, implicative, etc.
- **Motion domain:** agen/non-agent motion in place; forward motion; transport; external motion



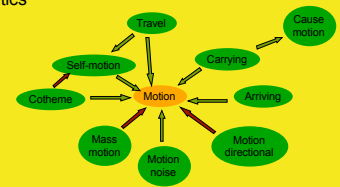
WordNet / GermaNet

- **Motivation:** lexical memory
- **Goal:** lexical semantic net
- **Organisation:** deep hierarchy
- **Class relations:** lexical and conceptual relations
- **Class formation:** synsets of synonymous verbs
- **Basis for synonymy:** substitution in prototypical context



FrameNet / SALSA

- **Background:** Fillmore's frame semantics
- **Goal:** lexicographic database
- **Organisation:** flat hierarchy
- **Frame:** situational knowledge
- **Frame inventory:** semantic roles, participants, situation properties
- **Frame relations:** inheritance, subframe, using
- **Verb meaning:** frame choice



Gold Standard for Automatic Induction

- **Goal:** gold standard
- **Choice:** 168 verbs → 43 classes
- **Class size:** 2-7 verbs
- **Ambiguity:** 8 verbs with 2 senses
- **High and low frequency verbs:** 8 ≤ freq ≤ 71,604
- **Class labels** on two conceptual levels
- **Relation to FrameNet:** conceptual scene description
- **Basis:** subjective conceptual knowledge, monolingual/bilingual dictionaries, corpora



References

Collin Baker, Charles Fillmore, and John Lowe. The Berkeley FrameNet Project. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics*, Montreal, Canada, 1996.

Thomas T. Ballmer and Waltraud Brennenstuhl. *Deutsche Verben*. Gunter Narr Verlag, Tübingen, 1986.

Katrin Erk, Andrea Kowalski, and Manfred Pinkal. A Corpus Resource for Lexical Semantics. In *Proceedings of the 5th International Workshop on Computational Semantics*, Tilburg, The Netherlands, 2003.

Claudia Kunze. Extension and Use of GermaNet, a Lexical-Semantic Database. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation*, pages 999-1002, Athens, Greece, 2000.

Sabine Schulte im Walde. *Experiments on the Automatic Induction of German Semantic Verb Classes*. PhD thesis, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 2003.