An Annotation of Semantic Change based on Usage Relatedness
Dominik Schlechtweg*, Sabine Schulte im Walde*, Stefanie Eckmann†
*Institute for Natural Language Processing, University of Stuttgart
†Historical and Indo-European Linguistics, LMU Munich

Introduction
- aim:
  - build a computational model detecting semantic change in corpora
- benefits:
  - test theories empirically on a large scale
  - inspire research by detecting new changes
- problem:
  - models should be evaluated on a sufficient number of semantic developments
  - there is no reliable test set of semantic change for any language
- solution:
  - collect a set of words displaying semantic change in a specific corpus in an annotation study

Semantic Change
- semantic innovation creates polysemy (cf. Fritz, 2006, p. 57)
- following Blank (1997, p. 113) we distinguish two types of semantic change:
  i) innovative meaning change: the emergence of polysemy
    - case example 1: German packen, ‘to grab’, earlier: Ich wollte reden, er packte mich aber mit starker Faust und worf mich zur Thüre hinaus...
    - case example 2: Der alte Schäferle packte ihn am Arme und rief: ‘The old Schäferle grabbed his arm and shouted.’
  ii) reductive meaning change: loss of a fully-established meaning
    - case example 1: German schinden, ‘to torture; to skin’, earlier: Wer sich stundenlang schindet / who later publicly gave testimony against him and sufficiently proved / how he had pressured...‘...solt få minde dengre schindet skindet / at vi ikke kan se...’
    - case example 2: Ich wollte reden, er packte mich aber mit starker Faust und worf mich zur Thüre hinaus...

Related Work
- various polysemy annotation studies of use pairs on proximity scales

Annotation Scale

4: Identical
3: Closely Related
2: Distantly Related
1: Unrelated
0: Cannot decide

Table 1: Four-point Scale of Relatedness derived from Brown (2008, p. 250).

Semantic Proximity
- Blank (cf. 1997, p. 406–419) distinguishes four degrees of semantic proximity for pairs of word uses relying on prototype theory:
  1. identity: sem. identical, both uses have same meaning
  2. context variance: sem. very near, referents belong to same category
  3. polysemy: sem. distant, ref. from diff. cat. but have sem. relation (similarity, contiguity...)
  4. homonymy: sem. very distant, ref. from diff. cat. and no sem. rel.
- basic idea: we measure the mean semantic proximity of uses of a word over time
  - increase suggests innovative meaning change (polysemization)
  - decrease suggests reductive meaning change

References

Acknowledgements
The research was supported by DFG Collaborative Research Centre SFB 732.