

Computational Lexical Semantics

Vorbesprechung und Seminarüberblick

Timm Lichte & Younes Samih

HHU Düsseldorf

1. Sitzung, 12. April 2018



- 1 Worum geht's?
- 2 Seminarüberblick
- 3 Cruse (2001): The lexicon
- 4 Formalitäten

Aus der Kursbeschreibung:

Das Lexikon enthält diejenigen semantischen Informationen, die für die Bedeutungskomposition notwendig sind.

- 1 Welche Einheiten enthält das Lexikon?
⇒ Morpheme, **Worte** (Wortformen, Lexeme), Phrasen, ...
- 2 Was ist “die Bedeutung”?
⇒ Problem: Mehrdeutigkeit, Abstraktheit
- 3 Wie repräsentieren wir (lexikalische) Bedeutung?
⇒ Paraphrasen, logische Formeln, Merkmalsstrukturen, Typenhierarchien, Vektoren, ...

In diesem Seminar

eine **Auswahl** (2 SWS!) computerlinguistischer “Antworten”

Wissen:

- grundlegenden theoretischen Konzepte der lexikalischen Semantik
- entsprechende computerlinguistischen Ressourcen und Anwendungen

Fähigkeit:

- wissenschaftliche Texte lesen
 - über wissenschaftliche Texte referieren
 - über wissenschaftliche Texte diskutieren
 - über wissenschaftliche Texte schreiben
- ⇒ “Lektüreseminar”

X tun lernt man nur, indem man X tut.

- 1 Theorie: Lexikon & lexikalische Semantik
- 2 Lexikalische Ressourcen: WordNet, FrameNet, VerbNet
- 3 Semantisch annotierte Korpora: SemCor, PropBank, OntoNotes
- 4 Anwendungen:
 - Word Sense Disambiguation
 - Semantic Role Labeling
 - ...

Cruse, D. Alan. 2001. The Lexicon. In Mark Aronoff & Janie Rens-Miller (eds.), *The Handbook of Linguistics*, 238–264. Oxford, UK: Blackwell Publishers. doi:10.1002/9780470756409.ch10.

ein paar Grundbegriffe der Lexikographie ...

- lexicon = “basic units” = “words” (“no succinct definition”)
- words = word forms | lexemes | lexical units
- word forms = inflected forms
- lexeme = “a set of related meanings associated with a set of related word forms”
- lexical unit = lexeme/sound + meaning

distinguishing properties of “words”:

- graphic and phonetic clues
- “largest chunks of the sentence which cannot be interrupted by the insertion of new material”
- possible meanings: complete chains of “semantic dependencies”

Contextual/holistic approach

- “the meaning of a word is the position within a network of relationships” (is-a, is-not-a, has-a, is-part-of, lives-in-a)
- Firth: “Words shall be known by the company they keep.” [distributional vectors]

Componential/localist approach

- “a finite assemblage of elementary bits of meaning”: semantic components, semantic features, semantic markers

Conceptual approach

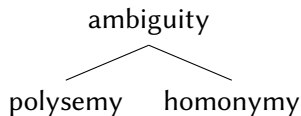
- [conceptual grounding of meaning]

Cruse (2001): How many meanings?

It is important to be able to decide whether two interpretations of a word in different contexts represent one semantic unit or two.

Ambiguity = separate **senses**

- identity test
- independent truth condition test
- zeugma test



Cruse (2001): Sense relations

To be interesting, a relation must recur with significant frequency throughout the vocabulary, and must be capable of supporting significant generalizations.

Relations of identity and inclusion:

- hyponymy (is-a relation)
- meronymy (part-whole relation)
- synonymy
 - absolute synonymy (substitutable salva grammaticale)
 - propositional synonymy (substitutable salva veritate)
 - near synonymy (“salient common core of meaning”)

Relations of opposition and exclusion:

- incompatibility and co-meronymy
- opposites (“inherently binary”): complementaries, antonyms, directional opposites, converses

paradigmatic (relations between words in isolation)

versus

syntagmatic (relations between words in a sentence)

- register clash
- semantic clash
- pleonastic
- [semantic roles of the verb]

Cruse (2001): Meaning extensions and change

- established readings and nonce readings
- literal and non-literal meanings
- metaphor
- metonymy
- spezialization and generalizations
- amelioration and pejoration

Cruse (2001): Larger groupings of words

- word fields
- word families
- domain specific vocabulary / layers of vocabulary
- mental lexicon
- vocabulary of a language

- **Leistungsnachweis:**
 - AP: Referat (unbenotet) + zwei Hausaufgaben (unbenotet) + Hausarbeit (benotet)
 - BN: Referat (unbenotet) + zwei Hausaufgaben (unbenotet)
- **Referate:**
 - maximal 2 Studierende pro Referat
 - vorgegebene Folienvorlagen (PowerPoint oder \LaTeX)
 - Folien auf Englisch (Vortrag auf Deutsch)
- **Hausaufgaben:**
 - zur Überbrückung der beiden Ferientage
 - kleine Programmierübung
- **Zielgruppe:** Fortgeschrittene Bachelorstudierende (Sprachwissenschaft, Computerlinguistik) mit Semantikvorkenntnissen
- **Website:** http://timm-lichte.de/?page_id=501

References