Computational Approaches to Creative Language: Idiomaticity

Caroline Sporleder

Computational Linguistics
Universität des Saarlandes

Sommersemester 2010

25.05.2010

Types of Idiomaticity (Baldwin, 2010)

- lexical idiomaticity: at least one word not in lexicon ("ad hoc")
- syntactic idiomaticity: syntax of whole expression not directly derived from syntax of components ("by and large")
- semantic idiomaticity: meaning of whole expression not directly derived from meaning of components ("middle of the road")
- pragmatic idiomaticity: expression associated with fixed set of situations ("good morning")
- statistical idiomaticity: a particular combination of words occurs with markedly higher frequency than components or alternative phrasings ("black and white" vs. "white and black")

Terminology (Baldwin, 2010)

'Idiom'

- any kind of multi-word expression (MWE)
- or only those which are semantically idiomatic

'Collocation'

- a statistically idiomatic MWE
- narrower definition requires also syntactic non-idiomaticity and relative semantic transparency

'Phraseology'

- intersted in description and functional classification of MWEs