A Contextual Analysis of Concrete and Abstract Words

Daniela Naumann, Diego Frassinelli, Sabine Schulte im Walde
Universität Stuttgart

{daniela.naumann,frassinelli,schulte}@ims.uni-stuttgart.de

Tag
Datum
Zeit
Raum

In the ongoing debate about meaning representation in the human mind, the precise description and comparison of concrete and abstract words plays a key role. Barsalou & Wiemer-Hastings (2005) provided experimental evidence that concrete words are grounded in the sensory-motor system, while the meaning of abstract words is derived from the activation of concrete words related to them. According to the Distributional Hypothesis (Firth, 1957), similar linguistic contexts tend to imply similar meanings of words. Thus, we suggest to use distributional semantics to perform a detailed analysis of the contextual cues of concrete and abstract words.

Based on previous psycholinguistic evidence, we hypothesise that: 1) the contexts of both concrete and abstract words are mainly composed of concrete words; 2) concrete words occur in a limited set of distinct contexts while abstract words appear in a broader range of different contexts; 3) concrete words are easier to predict than abstract words, due to the greater contextual variability of abstract words. We evaluate these assumptions within three context studies based on the ENCOW14A corpus (Schäfer & Bildhauer, 2012) and concreteness ratings from Brysbaert et al. (2014).

Our studies show consistent differences in the contexts of concrete and abstract words. Concrete words primarily co-occur with other concrete words, and their contexts can be predicted with greater certainty than the contexts of abstract words. However, abstract words mainly co-occur with other abstract words, which challenges the grounding theory of cognition and requires further investigation.

References: ● Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In: Pecher, D. and Zwaan, R. (eds.), *Grounding cognition: The role of perception and action in memory, language, and thought.* ● Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3). ● Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. In: Palmer, F.R. (ed.), *Selected papers of J. R. Firth 1952–59.* ● Schäfer, R., & Bildhauer, F. (2012). Building large corpora from the web using a new efficient tool chain. In: *Proceedings of LREC.*