Carita Paradis, Lund University & Debela Tesfaye, Addis Ababa University

On linguistics categories as categories
The case of antonyms and synonyms

This corpus study addresses the question of the nature and the structure of antonymy and synonymy as categories in language. While quite a lot of empirical research using different observational techniques has been carried on antonymy (e.g. Roehm et al. 2007, Lobanova 2013, Paradis et al. 2009), not as much has been devoted to synonymy (e.g. Divjak 2010) and very little has been carried out on both of them using the same methodologies (Gries & Otani 2010). The goal of this study is to bring antonyms and synonyms together, using the same (semi-)automatic methods to identify their behavioral patterns in texts. We examine the conceptual closeness/distance of synonyms and antonyms through the lens of their DOMAIN instantiations. For instance, strong used in the context of WIND or TASTE (OF TEA) as compared to light and weak respectively, and light as compared to heavy when talking about RAIN or WEIGHT. In order to identify as many domains as possible for our synonyms and antonyms, we choose as our starting-point sets of both antonym and synonym pairs, and through their use we extract and cluster other words expressing properties of these various domains. Using an algorithm similar to the one proposed by Tesfaye & Zock (2012) and Zock & Tesfaye (2012), we mine the co-occurrence information of the pairs in different domains separately, measuring the strength of their relation in the different domains with the aim of (i) making principled comparisons between antonyms and synonyms from a DOMAIN perspective, (ii) enhancing the algorithm to mine co-occurrence information specific to given domains, and (iii) determining the structure of antonymy and synonymy as categories in language and cognition.

References