

**Dominik Schlechtweg**

### **Exploitation of Co-reference in Distributional Semantics**

The aim of distributional semantics is to model the similarity of the meaning—the semantics—of words via the words they occur with—their distribution. Thereby, it relies on the distributional hypothesis implying that similar words have similar contexts. Deducing meaning from the distribution of words is interesting in that it can be done automatically on large amounts of raw text freely available, e.g., on the internet. It is because of this convenience that most current state-of-the-art-models of distributional semantics operate on raw text, although there have been successful attempts to integrate other kinds of—e.g., syntactic—information to improve distributional semantic models. In contrast there has been less attention to semantic information in the research community. One reason for this is that the extraction of semantic information from raw text is a complex, elaborate matter and in great parts not yet satisfyingly solved. Nevertheless, recently there have been successful attempts to integrate a certain kind of semantic information, i.e., co-reference—the knowledge about which words refer to the same entities, into distributional models of the semantics of words. We will examine the general potential of co-reference to improve distributional semantic models. The contribution of co-reference is split into two basically different types of co-referential information which are integrated stepwise into current models of distributional semantics and evaluated on a wide range of existing and newly defined tasks on attributional similarity. The above-described way of deducing parts of the meaning of words has significant implications for the formulation of the distributional hypothesis, the methodological cornerstone of distributional semantics. Also, there is disagreement on the question whether or to which degree the distribution of words is involved in the cognitive process of learning the meaning of words reflected by the distinction between a weak and a strong version of the distributional hypothesis. We will also take a look at the possible cognitive implications of the ideas presented above.