Determining whether a noun phrase is discourse-given (i.e. referring to something already mentioned, c.f. Prince, 1981, 1992) or discourse-new is an important step in information structure and discourse analysis. The goal of my work is to automatically accomplish this task. As coreference annotation subsumes discourse-givenness, I am using existing coreference-annotated corpora for training a classifier. The talk will comprise (i) a discussion of existing coreference corpora for English (incl. e.g. van Deemter & Kibble's (2000) comments), and (ii) a presentation of classification experiments and results (incl. Ritz, 2010). It will conclude with an outlook on the applicability of such classifiers in corpus linguistics and Natural Language Processing.


