

Abstract

The case in Czech is the basic morphological means by which nouns express their function in a sentence. The objective of this thesis is to describe, from a frequency point of view, the relation between form and function of nouns, or, more precisely, how frequently cases (both simple and prepositional) are used to realise syntactic functions in sentences. The thesis is based on one of the largest corpora of written synchronic Czech: 100-million-token corpus SYN2005.

In order to obtain data on frequencies of syntactic functions of nouns in relation to their cases, we annotated the corpus SYN2005 with a dependency syntactic annotation. For this annotation, we adopted the format of the analytical layer of the Prague Dependency Treebank. The syntactic annotation has been performed by a stochastic parser: the MST parser. Since the reliability of this annotation was not high enough, we have built an automatic correction module, which identifies errors of syntactic annotation in the output of the stochastic parser and corrects these errors by means of linguistic rules. We have implemented 26 different rules, but annotation errors have been reduced by merely 6–8%. However, this correction module can be further developed. It can be used to correct the output of any dependency parser trained on the data from the Prague Dependency Treebank.

We used the syntactically annotated corpus SYN2005 as the basis for our research of frequencies of syntactic functions of nouns in relation to their cases. Using manual probes into the corpus, we have verified and corrected the statistical data obtained. We analysed and described in detail all the relevant combinations of syntactic functions and simple or prepositional cases. The frequencies of the syntactic functions of nouns and their realisation have been presented in synoptic tables, presenting data from the corpus SYN2005 and its three subcorpora of fiction, scientific/technical literature and journalism. We compared these three subcorpora according to the frequencies of syntactic functions of nouns, as well as according to the details of usage of several combinations of syntactic function and case.