

Abstract

Czech dictionaries were created in the pre-corpus era; it is thus clear that some of their entries don't take into account semantic prosody – the fact that some lexemes occur in particular contexts (consequences are always far-reaching or catastrophic, while intention can be both evil and noble). The aim of this thesis is to compare selected parts of a dictionary (randomly selected probe into Czech adjectives) with corpus material, define the extent of missing information in the current state of Czech language description and explore, how many per cent of information about entries is missing in contemporary dictionaries. We based our research on *Slovník spisovného jazyka českého* (Dictionary of Written Czech) and representative corpuses of written language SYN2005, or else SYN2010. Context analysis was carried out by means of statistical methods and collocation rates. The difference between dictionary definitions and information inferred from the corpus research could become a further guideline for creating a new dictionary (besides adding a whole range of new entries that are still missing in Czech dictionaries).

Keywords

lexicon, corpus, semantic prosody