Abstract

The bachelor thesis deals with the variability of the non-verbal component of verbal phrasal idioms in journalistic texts. It is based on data from the Czech National Corpus, specifically from the SYNv9 corpus.

Verbal phrasal idioms are among the most frequent in Czech, a large part of them shows high variability of the verbal component, and a smaller part of the other components. The aim of this paper is to focus on the variability of the non-verbal component and try to describe its types. Using automatic phrasal idiom annotation, phrasal idioms variants are selected, then cases with a variant non-verbal component are manually selected, using context to distinguish when it is a deliberate update by the author and when it is a variant. Then we classify and describe the variants found. The results of this work can be used to extend variant verb phrasal idioms in the LEMUR database of multi-word units.

Keywords