Abstract

The aim of the thesis is to propose a tagging system for a learner corpus of spoken English which would, apart from tagging errors, focus also on the features specific for spoken language. Theoretical part, therefore, introduces basic concepts including learner language, the development of learner corpora in the last 20 years and both classical and computer-aided error analysis. Features typical of spoken language are described in the theoretical part as well since these are the focus of the research part of the thesis. The Louvain tagging system used for error-tagging of a leaner corpus of written language is used as the basis for the tagging system proposed in this thesis. Based on the analysis of 20 transcriptions taken from the Czech part of spoken learner corpus LINDSEI, modifications of the categories taken from the Louvain error-tagging system are proposed and new categories necessary for a better description of spoken language are introduced. The tagging system proposed in this thesis should make further analysis of the tagged corpus easier.