

Title: **The Use of Connectors in the Texts of Non-Native Speakers of Czech**
Author: **Pavel Pečený**
Department: Institute of Czech Language and Theory of Communication
Faculty of Arts, Charles University
Supervisor: doc. RNDr. Vladimír Petkevič, CSc.

Abstract

Over the last few years two learner corpora of authentic texts of non-native speakers of Czech have originated (MERLIN Corpus and CzeSL Corpus), giving linguists an important source of data for researching Czech as a foreign language. Ergo, for the first time it is possible to carry out the language analysis of non-native speakers of Czech using tools of corpus linguistics to formulate evidence-based research findings.

The presented thesis uses that as well, focussing on the study and description of the use of connectors in the written text production of non-native speakers of Czech, being primarily based on evidence from the learner MERLIN Corpus, which as opposed to other corpora is characterized by linking written text production reliably to the proficiency levels of the *Common European Framework of Reference for Languages (CEFR)*, including the levels of A2–B2. At the same time, it also contains a text error annotation, thus enabling to ascertain what effect the language proficiency has on the extent of the repertoire and frequency of using connectors, as well as to determine how the proficiency is reflected in using them correctly.

The theoretical part of the thesis summarizes the approaches to the classification of connectors in Czech, pointing out some difficulties in their classification. Furthermore, it deals with the description of language proficiency within the CEFR with respect to using connectors. Last but not least, it presents the basic principles of an error analysis and the interlanguage theory. The second part of the dissertation characterises the used corpora, explaining the procedures for retrieving and processing data employed for the subsequent analysis, which is included in the third part of the thesis. Firstly, the repertoire of connectors that the corpus contains is described there. Subsequently, the tendencies in their use are researched. Using a correspondence analysis, the relationship of connectors to individual proficiency levels of the CEFR is shown. The description is in the end completed with a basic quantitative-qualitative error analysis.

Key words: connectors, Czech as a foreign language, interlanguage, error analysis, language proficiency level, CEFR, MERLIN corpus