Magdaléna Rysová
Discourse Connectives in Czech
(From Centre to Periphery)

Abstract

The thesis focuses on description and analysis of discourse connectives in Czech in broader sense, i.e. by which language means it is possible to express sense relation within a text. The thesis is not limited to any parts of speech (like conjunctions or structuring particles) but it tries to find and describe all language means in Czech with the ability to connect two pieces or units of a text into one coherent complex. The thesis investigates discourse connectives in Czech with respect to the so called **secondary connectives** (i.e. mainly multiword phrases like *to je důvod, proč – that is the reason why*; *kvůli těmto skutečnostem – due to these facts* etc., in opposition to primary connectives like *však – however*, *nebo – or*, *a – and*, *ale – but*, *proto – therefore* etc.).

Discourse connectives are (in general terms) understood as language expressions that signal semantico-pragmatic relations within a text. However, there are many theories that significantly differ in the concrete description of these expressions. Therefore, there is not a generally accepted and universal definition of discourse connectives and their description and characteristics is still a matter of linguistic discussion. The aim of this thesis is to contribute to the general delimitation and definition of discourse connectives (not only in Czech) on the basis of a detailed language research.

Firstly, we describe various language theories focusing on description of discourse (or text). We concentrate especially on the individual approaches to discourse connectives and on the general criteria according to which we may define these expressions in language.

The detailed analysis of discourse connectives (with the emphasis on secondary connectives) is based especially on the annotated data of the Prague Dependency Treebank (PDT) containing written newspaper texts, marginally on the data of the Czech National Corpus (concretely SYN2013PUB). In the thesis, we introduce the principles and results of the annotation of secondary connectives in PDT (carried out for purposes of this thesis). We focus mainly on the comparison of primary connectives and secondary connectives according to their frequency in authentic texts, according to sense relations they express or according to

whether they occur more often inter- or intra-sententially. We compare then the annotations of secondary connectives in Czech and the similar expressions in English (carried out on the data of the Pennsylvanian corpus Penn Discourse Treebank) and we shortly compare the occurrence and usage of secondary connectives in written and spoken Czech (in PDT and in the corpus DIALOG).

On the basis of the practical analysis of discourse connectives, we try to draw several theoretical conclusions about these expressions. We focus mainly on the definition of discourse connectives, on general principles describing how it is possible to detect these expressions in authentic texts, and on the delimitation of the boundaries between primary and secondary connectives (mainly with respect to their lexical, syntactic and semantic characteristics).

The detailed analysis of discourse connectives and the general conclusions presented in this thesis might improve the understanding of semantico-pragmatic relations in texts and in this sense also the understanding of a text structure as a whole.