

Title: Valency of verbs in the Prague Dependency Treebank

Author: PhDr. Zdeňka Urešová

Department: Institute of Formal and Applied Linguistics MFF UK

Supervisor: Prof. PhDr. Eva Hajičová, DrSc.

Abstract: This dissertation describes PDT-Vallex, a valency lexicon of Czech verbs, and its relation to the annotation of the Prague Dependency Treebank (PDT). The PDT-Vallex lexicon was created during the annotation of the PDT and it is a valuable source of verbal valency information available both for linguistic research and for computerized natural language processing. In this thesis, we describe not only the structure and design of the lexicon (which is closely related to the notion of valency as developed in the Functional Generative Description of language) but also the relation between the PDT-Vallex and the PDT. The explicit and full-coverage linking of the lexicon to the treebank prompted us to pay special attention to diatheses; we propose formal transformation rules for diatheses to handle their surface realization even when the canonical forms of verb arguments as captured in the lexicon do not correspond to the forms of these arguments actually appearing in the corpus.