

Mirjam Fried (Prague, Czech Republic)

## Grammatical analysis and corpus evidence

### PLENARVORTRAG/KEYNOTE SPEECH

#### Abstract

The goal of this talk is two-fold: (i) to explore the interdependence of qualitative and quantitative analysis in articulating empirically plausible and theoretically coherent generalizations and (ii) to draw attention to the central challenge in this enterprise, namely, capturing the inherently dynamic, variable nature of grammatical structure. In addressing these two general concerns, I will show that the use of large electronic corpora is indispensable to the grammarian's work primarily as a rich source of semantic and contextual information, which turns out to be crucial in categorizing and explaining grammatical forms. On the basis of corpus evidence we can arrive not only at sufficiently multi-faceted generalizations about a given form itself, but also capture subtle shifts in its distribution, depending on specific, well-defined criteria.

The theoretical issues will be illustrated on one particular syntactic form in Czech, concerning the use and classification of relative clauses (RC) with the non-declinable relativizer *co* + resumptive pronouns, against the background of the stylistically and semantically neutral RCs introduced by a fully declinable agreeing relative pronoun *který* 'which'. I will take a set of existing claims about the properties of these clauses and the 'rules' about their form, interpretation, and distribution in the Czech grammatical literature, and test their validity on corpus material. The relevant analytic categories, with implications for relativization strategies beyond the Czech facts, revolve around the referential type of the relativized noun, the interaction between relativization and deixis, and the semantic relationship between the proposition expressed by the RC and the head noun. The analysis, which takes into account also frequency-based patterns of usage, will demonstrate that some of the existing claims are either fully invalid, or too general to capture relevant semantic distinctions, while others are roughly on the right track but often not comprehensive or precise enough to capture the full richness of the facts.

On the basis of the corpus material, the presentation argues for a more dynamic approach to grammatical analysis, one in which grammatical generalizations can be organized in cognitively and communicatively coherent networks of overlapping grammatical patterns. The networks simultaneously provide a tool for (i) identifying points of fluctuations within the usage of a particular form and (ii) tracking incipient shifts between the form and/or function of a given form.