

## **Corpora and cross-linguistic comparison: opportunities and challenges**

***Natalia Levshina, Max Planck Institute for Psycholinguistics, Nijmegen***

Over the last few years, the number of corpora that can be used for language comparison has dramatically increased. In my talk I will chart this new and changing territory, providing a few landmarks, warning signs and safe paths. Although no corpus at present can replace the traditional type of typological data based on language description in reference grammars, corpora can help with diverse tasks, being particularly well suited for investigating probabilistic and gradient properties of languages and for discovering and interpreting cross-linguistic generalizations based on processing and communicative mechanisms. At the same time, the use of corpora for typological purposes has not only advantages and opportunities, but also numerous challenges. I will also present several empirical case studies, including one addressing pertinent problems, namely, the role of text types in language comparison and the problem of using the word as a cross-linguistic comparative concept.