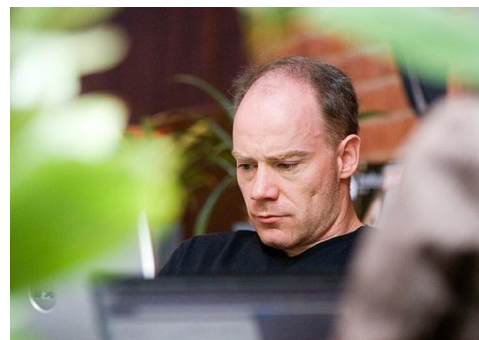# Stefan Th. Gries

We are pleased to announce that Professor **Stefan Th. Gries** has accepted the invitation of the Institute of the Czech National Corpus and is coming to Prague to give **two lectures at the Faculty of Arts**. Both lectures are open to the public.

Corpus data and aspects of the mental lexicon from a cognitive-linguistic perspective: frequency, contingency, recency, and context

Wednesday, April 26th 2017, 6 pm
Faculty of Arts, Charles University (room 104), nám. Jana Palacha 2

Over the last decades, linguistic research has become more diverse both theoretically and methodologically. With regard to the former, after a long period in which "theoretical linguistics" was synonymous with "generative linguistics", now a wider variety of approaches have emerged; for this talk of interest are cognitive/usage-based approaches, which assume a less-than-modular linguistic system that is 'governed' to a large extent by domain-general mechanisms such as frequency, contingency, recency, context etc. With regard to the latter, linguists are now routinely using a wider range of data and it is corpus data that have seen a particular increase. Against this background, I will discuss in this talk ways in which corpus-based work can help explore the lexicon/construction in ways that properly operationalize the above domain-general determinants of processing and learning: frequency, contingency, recency, context. I will discuss two brief case studies – one on phonological similarity within lexical units (involving frequency), one on multi-word identification (adding contingency) – before I turn to a broader discussion of how to involve recency and context properly to our corpus-linguistic toolkit.

What statistical methods have to offer to linguistics: three (differently complex) case studies of spelling, morphological change, and foreign language learning

Thursday, April 27th 2017, 6 pm
Faculty of Arts, Charles University (room 104), nám. Jana Palacha 2

This talk is basically a demonstration of how quantitative methods of different degrees of sophistication can inform linguistic research on various levels of linguistic analysis. I will report on three case studies. First, I will show how very simple statistics can be used to explore aspects of Spanish Internet Orthography, specifically how standard spellings are changed in online forums and comments and how even speakers' typing is influenced by semantic and articulatory characteristics of what they are typing. Second, I will address a frequent question in historical data, namely how to study morphological change given the inherent noisiness and multidimensional nature of the data using exploratory as well as hypothesis-testing statistics. Finally, I will discuss a fairly new method designed to facilitate the exploration of how speakers of a certain kind (e.g., non-native speakers or indigenized variety speakers) differ from a 'standard/reference' group of speakers even when human annotators of, say, learner data are not available.

ČESKÝ NÁRODNÍ **KORPUS**