

Institute of Formal and Applied Linguistics
cordially invites you to a talk by

Prof. Joakim Nivre

**“Lost in the Woods? Transition-Based
Dependency Parsing with Non-Projective Trees”**



The talk takes place
on Monday **October 10, 2011**, at **1:30 p.m.**
at the Faculty of Mathematics and Physics,
Malostranské nám. 25, 4th floor, room S1 (428).

Abstract: The first part of the talk introduces the transition-based approach to data-driven dependency parsing, where inference is performed as greedy best-first search over a non-deterministic transition system, while learning is reduced to the simple classification problem of mapping each parser state to the correct transition out of that state. The second part of the talk explores three different techniques for handling non-projective trees in transition-based dependency parsing: pseudo-projective parsing, augmented arc transitions, and online reordering. A comparative evaluation based on data from Czech, English and German shows that all three methods can accurately recover non-projective dependencies but with differences in precision and recall that can be related to language-specific properties.

Joakim Nivre is Professor of Computational Linguistics at Uppsala University. He holds a Ph.D. in General Linguistics from the University of Gothenburg and a Ph.D. in Computer Science from Växjö University. Joakim's research focuses on data-driven methods for natural language processing, in particular for syntactic and semantic analysis, and has so far resulted in over 150 scientific publications. He is one of the main developers of the transition-based approach to data-driven dependency parsing, described in his 2006 book *Inductive Dependency Parsing* and implemented in the MaltParser system. Systems developed using MaltParser were tied for first place in the shared tasks on multilingual dependency parsing at the Conference on Computational Natural Language Learning in both 2006 and 2007. Joakim's current research interests include the analysis of mildly non-projective dependency structures, the integration of morphological and syntactic processing for richly inflected languages, and the modeling of human sentence processing. He is the Secretary of the European Chapter of the Association for Computational Linguistics (EACL) and the President of ACL's Special Interest Group on Computational Natural Language Learning (SIGNLL).

<http://stp.ling.uu.se/~nivre>

Professor Dr. Frederick Jelinek (1932–2010), dr.h.c. Charles University in Prague, Julian Smith Professor at JHU, Baltimore, MD, USA, of Czech origin, was an outstanding researcher in Electrical Engineering and Computational Linguistics. His breakthrough ideas have led to a whole new research paradigm – application of stochastic methods – in the field of automatic speech recognition as well as in natural language processing in general. He held leading positions at Cornell University, IBM T. J. Watson Research Center and Johns Hopkins University, and was a guest professor of Charles University in Prague.

The program of the Fred Jelinek Seminar Series and video recordings of the talks are available at <http://ufal.mff.cuni.cz>