SLOVAK NATIONAL CORPUS

Ľ. ŠTÚR INSTITUTE OF LINGUISTICS

The Fifth International Seminar

SLOVKO 2009
NLP, Corpus Linguistics, Corpus Based Grammar Research

25 – 27 November 2009, Bratislava

Smolenice/Bratislava, Slovakia

PROGRAMME

25 November 2009, WEDNESDAY

circa 8:00   Transport from Bratislava to Smolenice

9:30 – 10:00 REGISTRATION WITH COFFEE

10:00 – 10:30 *Opening*

10:30 – 10:50 Cristina Borrell and Antoni Oliver, Universitat Oberta de Catalunya, Spain
IT: Moving Towards Real Multilingualism

10:50 – 11:10 Adam Przepiórkowski, Rafał L. Górski, Marek Łaziński, and Piotr Pęzik, Institute of Computer Science, Polish Academy of Sciences, Poland
Recent Developments in the National Corpus of Polish

11:10 – 11:30 Emmerich Kelih, Institut für Slavistik, Universität Graz, Austria
Preliminary Analysis of a Slavic Parallel Corpus

11:30 – 11:50 Magda Ševčíková, Institute of Formal and Applied Linguistics, Charles University, Czech Republic
The Meaning of the Conditional Mood within the Tectogrammatical Annotation of Prague Dependency Treebank 2.0

11:50 – 12:30 ACCOMMODATION
12:30 – 13:30 LUNCH

13:30 – 13:50 Dana Hlaváčková, Klára Osolsobě, Karel Pala, and Pavel Šmerk), Faculty of Informatics, Masaryk University, Brno, Czech Republic
'Deriv' a Tool for Testing Automatic Derivation for Czech (and Other Languages)

13:50 – 14:10 Tatiana Bobkova, Kyiv National Linguistic University, Ukraine
Corpus of Computational Linguistics Texts

14:10 – 14:30 Doug Davidson and Adriana Hanulikova, MPI for Human Cognitive and Brain Sciences, Germany; Max Planck Institute for Psycholinguistics, Netherlands
Inflectional Entropy in Slovak

14:30 – 14:50 Vaclav Brezina, Department of Applied Language Studies and Linguistics, University of Auckland
We Only Say We Are Certain When We Are Not: A Corpus-informed Study of Epistemic Stance

14:50 – 15:20 BREAK

15:20 – 15:40 Elena Grishina, Institute of Russian Language RAS, Russia
Multimodal Russian Corpus (MURCO): General Structure and User Interface **(RUS)**

15:40 – 16:00 Svetlana Savchuk, Institute of Russian Language RAS, Russia
Spoken Texts Representation in the Russian National Corpus: Spoken and Accentologic Sub-corpora **(RUS)**

16:00 – 16:20 Olga Mitrofanova and Viktor Zakharov, Faculty of Philology, St. Petersburg State University, Russia
Automatic Analysis of Terminology in the Russian Corpus on Corpus Linguistics

18:00 – 19:30 DINNER

26 November 2009, THURSDAY

9:00 – 9:20   Annelen Brunner and Kathrin Steyer, Institute for the German Language, Mannheim, Germany
A Model for Corpus-driven Exploration and Presentation of Multi_word Expressions

9:20 – 9:40   Marcin Zabawa, Institute of English, University of Silesia, Katowice, Poland
A Corpus of Spoken Language and Its Usefulness in the Research on Language Contact

9:40 – 10:00 Marek Grác, Faculty of Informatics, Masaryk University, Brno, Czech Republic
Shallow Ontology Based On VerbaLex

10:00 – 10:20 Marek Nagy, Faculty of Mathematics, Physics and Informatics, Comenius University, Slovakia
Using Speech and Handwriting Recognition Enhanced Electronic worksheets to Learn Reading and Writing Skills

10:20 – 10:50 BREAK

10:50 – 11:10 Ružena Kozmová, FF UCM Trnava, Slovakia
Methodological Foundations for Contrastive Model of Verb Valence **(SLK)**

11:10 – 11:30 Natalia Darchuk and Viktor Sorokin, Ukraine
Text-oriented Thesaurus Retrieval System for Linguists **(UKR)**

11:30 – 11:50 Lukasz Jedrzejowski, Freie Universität Berlin, Germany
On epistemicity, Person and Speaker Deixis in Polish (Based on the Polish National Corpus) **(POL)**

11:50 – 12:10 Jana Kesselová, Katedra knižničných a slovakistických štúdií Fakulta humanitných a prírodných vied, Prešovská univerzita, Slovakia
Operators for Extending and Developing an Utterance

12:30 – 13:30 LUNCH

13:50 – 14:10 Mária Šimková, Slovak National Corpus, Ľ. Štúr Institute of Linguistics, Slovak Academy of Sciences, Bratislava, Slovakia
Frequency of Words and Forms in Contemporary Slovak (Based on the Slovak National Corpus **(SLK)**

14:10 – 14:30 František Štícha, Ústav pro jazyk český, AV ČR, v. v. i., Czech Republic
Frequency, Function and Grammaticality: The Case of the Anaphoric takový **(CZE)**

14:30 – 14:50 Ludmila Dimitrova, Mathematical Linguistics Department, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Bulgary
From Electronic Corpora to Online Dictionaries (on the example of Bulgarian Language Resources) **(RUS)**

15:30 – 18:00 ČERVENÝ KAMEŇ (trip to a castle)

19:00   Social event (conference venue)

27 November 2009, FRIDAY

9:00 – 9:20 Maria Khokhlova and Adam Rambousek, Faculty of Philology, St. Petersburg State University, Russia; Masaryk University, Brno, Czech Republic
Russian Lexicographic Framework – First Stage

9:20 – 9:40 Petr Pořízka and Markus Schäfer, Dept. of Czech Studies, Palacky University in Olomouc, Czech Republic; Institute of Computer Science. University of Bonn, Germany
Morph-Con – A Software for Conversion of Czech Morphological Tagsets **(CZE)**

9:40 – 10:00 Ján Genči, Technical University of Košice, Slovakia
Synset Building Based on On-line Resources **(SLK)**

10:00 – 10:20 Marie Mikulová and Jan Štěpánek, Institute of Formal and Applied Linguistics, Charles University, Prague, Czech Republic
Annotation Procedure in Building the Prague Czech-English Dependency Treebank **(CZE)**

10:20 – 10:50 BREAK

10:50 – 11:10 Patrice Pognan, LALIC-CERTAL (INALCO), France
Database Creation Based on a Dictionary Using a Corpus Method **(CZE)**

11:10 – 11:30 Maria Khokhlova and Viktor Zakharov, Faculty of Philology, St. Petersburg State University, Russia;
Corpus-based Analysis of Lexico-grammatical Patterns

11:30 – 11:50 Radovan Garabík, Jan Jona Javoršek, and Tomaž Erjavec, Ľ. Štúr Institute of Linguistics, Slovak Academy of Sciences, Bratislava, Slovakia & Jožef Stefan Institute, Ljubljana, Slovenia;
Evaluating Grid Infrastrucure for Natural Language Processing

11:50 – 12:10 Renata Šimůnková, Technical University of Liberec, Czech Republic;
Analysis of the Means Expressing 'Necessity not to' in English and Czech Based on General and Parallel Corpora

12:10 – 12:30 Zdenka Urešová, Charles University, Czech Republic
Diatheses in the Czech Valency Lexicon PDT-Vallex

12:30 – 13:30 LUNCH

13:30 – 13:50 Lyubov Struhanets, Department of the Methodology of the Ukrainian Language Teaching and Speech Culture, Ukraine
Representation of Changes of World Stock of the Ukrainian Literary Language of the 20th – Beginning of the 21st Centuries in the Information System "Lexika" **(UKR)**

13:50 – 14:10 Olena Siruk, Viktor Sorokin and Pavlo Grytsenko, Laboratory for Computational Linguistics, Institute of Philology, Kyiv National Taras Shevchenko University, Ukraine
Electronic Lexical Card Index for the Ukrainian Dialects (ELCIUD) **(UKR)**

14:10 – 14:30 Markéta Lopatková and Václava Kettnerová, Institute of Formal and Applied Linguistics, Charles University, Prague, Czech Republic
Changes in Valency Structures of Verbs: Grammar vs. Lexicon **(CZE)**

14:30 – 14:50 Olga Shypnivska and Sergiy Starykov, Taras Shevchenko University of Kyiv, Ukraine; Ukrainian Lingua-Informational Fund of the NAS of the Ukraine
The Creation of the Depositary of Morphological Ambiguity in Ukrainian Language

14:50 – 15:10 Irina Nekipelova, Russia
Composite Lexical Units as an Element of Lexicographical Historical Computer System **(RUS)**

15:10 – 15:20 CONCLUSION

16:00 Transport from Smolenice to Bratislava

Language of the presentation is indicated in ISO 639-3 code in brackets next to the presentation title. If omitted, the presentation will be in English.