

24–26 September 2018, Charles University, Prague

MONDAY 24 SEP 2018

PANEL ON CORPUS-BASED LANGUAGE VARIATION

9:00 – 18:00 Registration

Chair: Václav Cvrček

9:30 – 10:00 Václav Cvrček, Zuzana Komrsková, David Lukeš, Petra Poukarová, Anna Řehořková and Adrian Jan Zasina: Multi-dimensional analysis of Czech

10:00 – 10:30 Roland Meyer: Register in Russian: Multi-dimensional analysis vs. functional stylistics

10:30 – 11:00 Randi Reppen: Using Multi-Dimensional Analysis to analyze university student register variation

11:00 – 11:30 Václav Cvrček, Zuzana Komrsková, David Lukeš, Petra Poukarová, Anna Řehořková and Adrian Jan Zasina: Register versus idiolect variation

11:30 – 12:00 Coffee break

Chair: Václav Cvrček

12:00 – 13:00 **Plenary:** Douglas Biber: Exploring cross-linguistic universals of register variation through multi-dimensional analysis



EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education



MAIN CONFERENCE

MONDAY 24 SEP 2018

13:00 – 14:20 Lunch

14:20 – 14:30 Conference opening

Chair: Michal Křen (room no. 1)

14:30 – 15:30 **Plenary:** Ruprecht von Waldenfels: Variation on many levels: why and how comparing corpora and (Slavic) languages makes sense

15:30 – 16:00 Coffee break

	Natural language processing 1 (room no. 1)	Contrastive research 1 (room no. 2)	Discourse (room no. 3)
	Chair: Adam Przepiórkowski	Chair: Neil Bermel	Chair: Laura Janda
16:00 – 16:30	Nikola Ljubešić, Tanja Samardžić, Tomaž Erjavec, Darja Fišer, Maja Miličević and Simon Krek. "Kad se mnogo malih složi": Collaborative development of gold resources for Slovene, Croatian and Serbian	Katarzyna Dziwirek. To taste is to live and love: Verbs of taste in Polish and English	Thomas Samuelsson. The Russian adjectives antirossijskij , antiruskij and antisovetskij in Russian media: a corpus study
16:30 – 17:00	Vladimír Benko and Radovan Garabík. Ensemble Tagging Slovak Web Data	Magdalena Adamczyk. A contrastive look at discursive uses of English 'now' and Polish 'teraz'	Martina Berrocal. A corpus-assisted discourse study of the Presidential interviews of Milos Zeman
17:00 – 17:30	Witold Kieraś, Łukasz Kobyliński and Maciej Ogrodniczuk. Korpusomat --- new functionalities and future development	Juho Härme. Last year but not yesterday? Explaining differences in the locations of Finnish and Russian time adverbials using comparable corpora	[special talk] Darja Fišer. CLARIN for researchers of Slavic languages

17:30 – 18:30 Welcome drink

TUESDAY 25 SEP 2018

Chair: Alexandr Rosen (room no. 1)

9:30 – 10:30 **Plenary:** Björn Hansen & Edyta Jurkiewicz-Rohrbacher: Detecting constraints on clitic climbing with the help of Web corpora

10:30 – 11:00 Coffee break

	Quantitative approaches (room no. 1)	Corpus building 1 (room no. 2)	Syntax (room no. 3)
	Chair: Ruprecht von Waldenfels	Chair: Marek Łaziński	Chair: Jan Patrick Zeller
11:00 – 11:30	Laura Janda and Francis Tyers. Parts Give More Than Wholes: Paradigms from the Perspective of Corpus Data	Anna Kryvenko. A reference corpus for discourse dynamics analysis in Ukrainian	Jakob Horsch. A construction grammar account of the Slovak comparative correlative construction
11:30 – 12:00	Miroslav Kubát, Jan Hůla, Radek Čech, David Číž and Kateřina Pelegrinová. Context Specificity of Lemma. Diachronic analysis	Natalia Grabar, Olga Kanishcheva and Thierry Hamon. Multilingual aligned corpus with Ukrainian as the target language	Svatava Škodová. Sebrat se a a construction between coordination and subordination in contemporary Czech
12:00 – 12:30	David Lukeš, Zuzana Komrsková, Marie Kopřivová and Petra Poukarová. Pronunciation of casual spoken Czech: A quantitative survey	Simon Krek. Gigafida, Kres, ccGigafida and ccKres	Matea Filko, Krešimir Šojat and Marko Tadić. Construction za + infinitive – evidence from the Croatian corpora
12:30 – 13:00	Jiří Milička and Alžběta Růžičková. Slovak Vowel Phonotactics: Slavic Origins vs. Hungarian Influences	Ranka Stanković, Miloš Utvić, Aleksandra Tomašević, Ivan Obradović and Biljana Lazić. Development and application of a domain specific corpus for mining engineering	Petar Vuković. The second future tense in contemporary Croatian: A corpus-driven study in grammatical semantics
13:00 – 14:30	Lunch		

TUESDAY 25 SEP 2018

	Language change (room no. 1)	Lexicon and multiword expressions (room no. 2)	SLAVICORP ROUND TABLE (room no. 3)
	Chair: Mikhail Kopotev	Chair: Radek Čech	
14:30 – 15:00	Tore Nasset. Cascading S-curves: What corpus linguistics tells us about language change	Tomáš Káňa. Terminology in and around Diminutives	
15:00 – 15:30	Stefan Heck and Eugen Kravchenko. Polish być w trakcie + verbal noun – a progressive periphrastic construction?	Milena Hnátková, Tomáš Jelínek, Marie Kopřivová, Vladimír Petkevič, Alexandr Rosen, Hana Skoumalová and Pavel Vondříčka. Multiword Expressions in Czech: Typology and Lexicon	
15:30 – 16:00	Anna Řehořková. Czech conditional verb forms in complement clauses	Lucie Lukešová (Chlumská) and Dominika Kovářiková. Extracting Multi-word Expressions for the Czech Academic Phrase List	
16:00 – 16:30	Coffee break		
	Morphology (room no. 1)	Corpus building 2 (room no. 2)	
	Chair: Thierry Hamon	Chair: Hanne Martine Eckhoff	
16:30 – 17:00	Lucie Kopáčková. Oprahin or Opražin? How to Correctly Form Possessive Adjective from Female First Name or Surname of Foreign Origin in Contemporary Written Czech Language?	Dorota Adamiec, Renata Bronikowska, Włodzimierz Gruszczyński, Emanuel Modrzejewski and Aleksandra Wieczorek. The Electronic Corpus of the 17th and 18th c. Polish Texts (up to 1772). The final result	
17:00 – 17:30	Adrian Jan Zasina and Michal Škrabal. Morfio.pl – the possibilities for the application of Czech corpus tools to other languages	Moulay Zaidan Lahjouji. The Corpus of Spoken Rusyn – A user friendly resource for research on Rusyn dialects	
17:30 – 18:00	Neil Bermel and Luděk Knittl. The fate of variant forms in historical corpora: Tracing locative exponents in DIAKON	Marcin Szczepański. Recent challenges and advances in the development of Lower Sorbian corpus resources	

WEDNESDAY 26 SEP 2018

Chair: Lucie Lukešová (Chlumská) (room no. 1)

9:30 – 10:30 **Plenary:** Alexandr Rosen: The merits of a parallel corpus and how to get the most out of it

10:30 – 11:00 Coffee break

	Natural language processing 2 (room no. 1)	Contrastive research 2 (room no. 2)	Lexicology and lexicography (room no. 3)
	Chair: Simon Krek	Chair: Björn Hansen	Chair: Magda Ševčíková
11:00 – 11:30	Witold Kieraś and Marcin Woliński. Basic natural language processing toolkit for 19th century Polish	Marek Łaziński. Actional Interpretation of Verbal Aspect in Legal Texts - Corpus Analysis	Jana Nová, Vít Michalec, Zdeňka Opavská and Renáta Neprašová. Frequency (not) sacred: The headword list of a contemporary Czech monolingual dictionary and corpora
11:30 – 12:00	Tomaž Erjavec, Nikola Ljubešić and Darja Fišer. Training data and tools for processing user-generated content in Slovene	Tora Hedin. The Swedish particle alltså and the Czech particles tedy and takže: a corpus study	Olga Goritskaya and Mikita Suprunchuk. Frequency Dictionary of Belarusian Borrowings in the Belarusian Variety of the Russian Language
12:00 – 12:30	Natalia Kotsyba and Bohdan Moskalevskiy. An essential infrastructure of Ukrainian language resources and its possible applications	Katja Brankačkec. Productivity and Meaning of the prefix nad- in the Word-Formation of Upper Sorbian, Lower Sorbian and Czech in a diachronic perspective: Evidence from Corpora	Jan Patrick Zeller. Syntagmatic corpus analyses of mixed speech: code-shifting in Belarusian trasyanka and Ukrainian suržyk
12:30 – 13:00	Anastasiia Baranchikova, Anna Dmitrieva, Mariia Fedorova, Aleksandr Klimov, Olesya Kisselev, Mikhail Kopotev, Svetlana Toldova and Natalia Zevakhina. CAT&kittens: a corpus-based text-analytic tool for Russian academic writing	Tatiana Perevozchikova. Pronominal expression of possession in noun phrases in Russian, Czech, and Bulgarian	Valeriia Kolosova, Ksenia Zaytseva and Kira Kovalenko. PhytoLex – the Database of Russian Phytonyms: from Idea to Implementation
13:00 – 14:30	Lunch		
	Language resources (room no. 1)	Aspect (room no. 2)	Second language acquisition (room no. 3)
	Chair: Dmitri Sitchinava	Chair: Tore Nesset	Chair: Vladimír Benko
14:30 – 15:00	Adam Przepiórkowski and Agnieszka Patejuk. An Enhanced Universal Dependencies Treebank of Polish	Alexander Piperski. Aspect-Specific Keywords in Russian	Tomáš Jelínek. New error annotation of Czech learner corpora
15:00 – 15:30	Magda Ševčíková, Adéla Kalužová and Zdeněk Žabokrtský. A language resource specialized in Czech word-formation: Recent achievements in developing the DeriNet database	Ilona Starý Kořánová. Aspectual homonymy and polysemy in Czech	Adrian Jan Zasina. Evaluating a corpus-driven approach in L2 classroom on the example of Czech
15:30 – 16:00	Hanne Martine Eckhoff, Aleksandrs Berdičevskis and Marius Jøhndal. From diachronic treebank to dictionary resource: the Varangian Rus project		Jane Hacking, Erin Schnur and Fernando Rubio. MuSSeL: Designing and building a corpus of multilingual second language speech
16:00 – 16:10	Conference closing		