

# **Serbian-Slovakian bilateral project proposal for period 2017-2018.**

Research area:

Applied mathematics

Title:

Quantitative analysis of syllables in Slavic languages (Russian, Slovak, Serbian)

Keywords:

Quantitative linguistics, syllables modelling, Slavic languages

## **SERBIAN INSTITUTION:**

Institution:

University of Belgrade

Department:

Faculty of Mining and Geology

Address:

Djušina 7

Phone:

+381113219101

Fax:

+381113235539

Public institution:

Yes

## **SERBIAN PRINCIPAL INVESTIGATOR:**

First name:

Ivan

Last name:

Obradović

Position:

Professor (full)

JMBG:

1106951710276

Address:

Djušina 7

Phone:

+381113219259

Fax:

+381113235539

e-mail 1:

ivan.obradovic@rgf.bg.ac.rs

e-mail 2:  
ivanobradovic@gmail.com

**Biography:**

Ivan Obradović is a full professor of mathematics and informatics at the University of Belgrade Faculty of Mining and Geology and the director of the University of Belgrade Center for e-learning and distance education. He has published four books and more than 130 scientific papers related to artificial intelligence methods and tools, language resources and tools, quantitative linguistics and e-learning applications in higher education. He is one of the key persons in several international and national scientific projects related to human language technologies and e-learning.

**Relevant publications:**

1. I. Obradović, A. Obuljen, D. Vitas, C. Krstev, V. Radulović, "Distribution of canonical syllable types in Serbian", in: Text and Language, Structures · Functions · Interrelations. Quantitative Perspectives, P. Grzybek, E. Kelih, J. Mačutek (eds.), Praesens Verlag, Wien 2010, pp.145157, ISBN 9783706906258. 2. I. Obradović, C. Krstev, D. Vitas, "Measuring semantic relevance of words in synsets", in: Text and Language, Structures · Functions · Interrelations, Quantitative Perspectives, P. Grzybek, E. Kelih, J. Mačutek (eds.), Praesens Verlag, Wien 2010, pp. 133144, ISBN 9783706906258. 3. I. Obradović, "A Method for Extracting Translational Equivalents from Aligned Texts", in: Methods and Applications of Quantitative Linguistics, Ivan Obradović, Emmerich Kelih, Reinhard Köhler (Eds.), University of Belgrade, Academic Mind, Belgrade, 2013, pp. 119129, ISBN 978-86-7466-465-0. 4. Cvetana Krstev, Ivan Obradović, Miloš Utvić, and Duško Vitas, "A System for Named Entity Recognition Based on Local Grammars", Journal of Logic and Computation, 2014, Vol. 24, No. 2, pp. 473489, ISSN 0955792X. 5. Ranka Stanković, Cvetana Krstev, Ivan Obradović, Olivera Kitanović, "Indexing of Textual Databases Based on Lexical Resources: A Case Study for Serbian", in: Semantic Keywordbased Search on Structured Data Sources, Cardoso, J., Guerra, F., Houben, G.J., Pinto, A.M., Velegrakis, Y. (Eds.), Springer International Publishing, 2015, pp. 167181, ISBN: 9783319279312, DOI: 10.1007/9783319279329.

**SERBIAN RESEARCH GROUP:**

Marija Radojičić, Teaching assistant, PhD student

Biography: Marija Radojičić is a PhD student at the University of Belgrade Faculty of Mathematics. Her research focus is language processing of mathematical content. In the last 3 years she has been working as a teaching assistant at the University of Belgrade Faculty of Mining and Geology. She has recently been involved in the Tempus project BAEKTEL, a project for technology enhanced learning, blending academic and entrepreneurial knowledge. Marija Radojičić has also has been involved in various projects linked to the Ministry of Trade, Telecommunications and Tourism. She has presented papers at different recognised national and international conferences. As a consequence of her proactive engagement in the diverse projects, she has gained experience in using different tools for language processing and valuable insights into international cooperation. Her PhD thesis aims to further explore language processing with a specific focus on the processing of mathematical content for the Serbian language.

Biljana Lazić, Librarian, PhD student

Biography: Biljana Lazić is a PhD student at the University of Belgrade Faculty of Philology. In previous work she dealt with Computational Linguistics topics. Her thesis is focused on Electronic Lexicography. The main purpose is to take advantage of already developed language resources for Serbian language (morphological dictionaries, WordNet, corpora etc.) to create a dictionary on the web. Her participation in the project will provide her the opportunity to take chance in quantitative linguistics. The second benefit is the possibility of applying results in her dissertation thesis in purpose of syllabification of headwords which could be one of the services in final online dictionary. It is anticipated that experience gained during this project will be helpful for her academic career in sense of incorporation into the international scientific community.

,  
Biography:

,  
Biography:

,  
Biography:

**SLOVAKIAN INSTITUTION:**

Institution:

Comenius University in Bratislava

Department:

Faculty of Mathematics Physics and Informatics

Address:

Šafárikovo námestie 6, 81499 Bratislava, Slovakia

Phone:

+421260295182

Fax:

+421265425882

Public institution:

Yes

Fiscal number:

00397865

Bank Account:

IBAN SK28 8180 0000 0070 00072962

**SLOVAKIAN PRINCIPAL INVESTIGATOR:**

First name:

Ján

Last name:

Mačutek

Position:

Associate professor

Address:

Department of Applied Mathematics and Statistics, Comenius University, Mlynská dolina, 84248 Bratisl

Phone:

+421260295717

Fax:

+421265425882

e-mail 1:

macutek@fmph.uniba.sk

e-mail 2:

jmacutek@yahoo.com

Biography:

Ján Mačutek is an associate professor at Faculty of Mathematics Physics and Informatics of Comenius University in Bratislava. J. Mačutek has been active in quantitative linguistics research for more than a decade. He published extensively in this area (more than 60 publications, see [http://www.iam.fmph.uniba.sk/ospm/Macutek/jm\\_publications.pdf](http://www.iam.fmph.uniba.sk/ospm/Macutek/jm_publications.pdf) for the list of publications). He is an active member of the International Quantitative Linguistics Association (IQLA). J. Mačutek has been the Treasurer of IQLA since 2009. He is a member of editorial boards of Journal of Quantitative Linguistics, Journal of Language Modelling, Glottotheory, Mathematical Linguistics, and Glottometrics, and also a member of the

editorial board of the book series Studies in Quantitative Linguistics (published by RAM-Verlag in Germany).

Relevant publications:

1. Čech, R., Mačutek, J., Liu, H. (2016). Syntactic complex networks and their applications. In: Mehler, A., Lücking, A., Banisch, S., Blanchard, P., Frank-Job, B. (eds.), Towards a Theoretical Framework for Analyzing Complex Linguistic Networks: 167-186. Berlin, Heidelberg: Springer. Kelih, E., Mačutek, J. (2013). 2. Number of canonical syllable types: a continuous bivariate model. Journal of Quantitative Linguistics 20, 241-251. 3. Mačutek, J., Rovenchak, A. (2011). Canonical word forms: Menzerath-Altmann law, phonemic length and syllabic length. In: Kelih, E., Levickij, V., Matskulyak, Y. (eds.), Issues in Quantitative Linguistics 2: 136-147. Lüdenscheid: RAM-Verlag. 4. Popescu, I.-I., Altmann, G., Grzybek, P., Jayaram, B.D., Köhler, R., Krupa, V., Mačutek, J., Mehler, A., Pustet, R., Uhlířová, L., Vidya, M.N. (2009). Word Frequency Studies. Berlin, New York: de Gruyter. 278 pp. 5. Popescu, I.-I., Mačutek, J., Altmann, G. (2009). Aspects of Word Frequencies. Lüdenscheid: RAM-Verlag. 193 pp.

**SLOVAKIAN RESEARCH GROUP:**

Michaela Koščová, PhD student

Biography: M. Koščová is a PhD student in applied mathematics, with J. Mačutek serving as her supervisor. Already her diploma thesis dealt with statistical analyses of linguistic data (results from the diploma thesis will be published in the Journal of Quantitative Linguistics; a draft of the paper, of which J. Mačutek and E. Kelih are co-authors, can be found at <http://arxiv.org/pdf/1504.03608.pdf>). Her dissertation thesis is focused on discrete probability distributions, which are one of the most commonly used mathematical models in quantitative linguistics. Her participation in the project will give her a twofold advantage – first, working with two experienced researchers she will have a unique opportunity to gain insight into mathematical modelling in linguistics; second, a mathematical/statistical analysis of (not only linguistic) data often opens up new vistas on mathematical models and methods themselves, and thus provides an impetus towards further mathematical research on a theoretical level. Moreover, experience she will obtain from an international cooperation, her integration into the international scientific community, and publications from the project will significantly increase her chances to start an academic career after her PhD study.

,  
Biography:

,  
Biography:

,  
Biography:

,  
Biography:

**Date:**

03.07.2016

**Signature of principal investigator:**



Professor Dušan Polomčič

Dean of University of Belgrade Faculty of Mining and Geology