

7th Workshop on Recent Advances in Slavonic Natural Language Processing

(RASLAN 2013)

**Karlova Studanka, Czech Republic
6 – 8 December 2013**

Editors:

**A. Horak
P. Rychly**

ISBN: 978-1-62993-833-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2013) by the Masaryk University Faculty of Informatics
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the Masaryk University Faculty of Informatics
at the address below.

Masaryk University Faculty of Informatics
c/o Dana Hlaváčková
Centre for Natural Language Processing
B206 Botanická 554/68a, Ponava, Brno
Czech Republic

Phone: +420 549 49 1864, 3329

ydana@aurora.fi.muni.cz

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

I Syntax, Morphology and Lexicon

Preparing VerbaLex Printed Edition	3
<i>Dana Hlaváčková, Aleš Horák, and Karel Pala</i>	
Web Application for Semantic Network Editing	13
<i>Adam Rambousek and Tomáš Hrušo</i>	
Portable Lexical Analysis for Parsing of Morphologically-Rich Languages	21
<i>Marek Medved' and Miloš Jakubíček</i>	

II Semantics

Acquiring Data for Textual Entailment Recognition	29
<i>Zuzana Nevěřilová</i>	
Semi-automatic Theme-Rheme Identification	39
<i>Karel Pala and Ondřej Svoboda</i>	

III Text Corpora

Intrinsic Methods for Comparison of Corpora	51
<i>Vít Baisa and Vít Suchomel</i>	
Typos in Czech Corpora	59
<i>Marek Grác</i>	
Fast Construction of a Word \leftrightarrow Number Index for Large Data	63
<i>Miloš Jakubíček, Pavel Rychlý, and Pavel Šmerk</i>	

IV Language Modelling and Machine Translation

Expanding Translation Memories: Proposal and Evaluation of Several Methods	71
<i>Vít Baisa, Josef Bušta, and Aleš Horák</i>	
Methods for Detection of Word Usage over Time	79
<i>Ondřej Herman and Vojtěch Kovář</i>	

Towards the Realistic Natural Language Representations	87
<i>Petr Sojka</i>	
Type-based Search of Idiomatic Expression	93
<i>Jan Bušta</i>	
Author Index	97