# Second Student Research Workshop 2011

# (RANLPStud 2011)

**Associated with the 8th International Conference on Recent Advances in Natural Language Processing (RANLP 2011)**

## Hissar, Bulgaria
## 13 September 2011

**ISBN: 978-1-62276-512-6**

**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© (2011) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2012)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

**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

*Domain-Dependent Detection of Light Verb Constructions*
István T. Nagy, Gábor Berend, György Móra and Veronika Vincze . . . . . 1

*A Weighted Lexicon of French Event Names*
Béatrice Arnulphy . . . . . 9

*Towards a Better Exploitation of the Brown 'Family' Corpora in Diachronic Studies of British and American English Language Varieties*
Sanja Štajner . . . . . 17

*Projecting Farsi POS Data To Tag Pashto*
Mohammad Khan, Eric Baucom, Anthony Meyer and Lwin Moe . . . . . 25

*Enriching Phrase-Based Statistical Machine Translation with POS Information*
Miriam Kaeshammer and Dominikus Wetzel . . . . . 33

*Inter-domain Opinion Phrase Extraction Based on Feature Augmentation*
Gábor Berend, István T. Nagy, György Móra and Veronika Vincze . . . . . 41

*ArbTE: Arabic Textual Entailment*
Maytham Alabbas . . . . . 48

*RDFa Editor for Ontological Annotation*
Melania Duma . . . . . 54

*Extracting Protein-Protein Interactions with Language Modelling*
Ali Reza Ebadat . . . . . 60

*Experiments with Small-size Corpora in CBMT*
Monica Gavrila and Natalia Elita . . . . . 67

*Question Parsing for QA in Spanish*
Iria Gayo . . . . . 73

*Incremental Semantics Driven Natural Language Generation with Self-Repairing Capability*
Julian Hough . . . . . 79

*Initial Experiments with Multilingual Extraction of Rhetoric Figures by means of PERL-compatible Regular Expressions*
Daniel Devatman Hromada . . . . . 85

*Is Three the Optimal Context Window for Memory-Based Word Sense Disambiguation?*
Rodrigo de Oliveira, Lucas Hausmann and Desislava Zhekova . . . . . 91

*Heterogeneous Natural Language Processing Tools via Language Processing Chains*
Diman Karagiozov . . . . . 97

*Pattern-Based Ontology Construction from Selected Wikipedia Pages*
Carmen Klaussner and Desislava Zhekova . . . . . 103

*Lexico-Syntactic Patterns for Automatic Ontology Building*
Carmen Klaussner and Desislava Zhekova . . . . . 109

*Towards a Grounded Model for Ontological Metaphors*
Sushobhan Nayak . . . . . 115

*Automatic Acquisition of Possible Contexts for Low-Frequent Words*
Silvia Necsulescu . . . . . 121

*Adapting Standard Open-Source Resources To Tagging A Morphologically Rich Language: A Case Study With Arabic*
Hajder S. Rabiee . . . . . 127

*Towards Cross-Language Word Sense Disambiguation for Quechua*
Alex Rudnick . . . . . 133

*Annotating Negation and Speculation: the Case of the Review Domain*
Natalia Konstantinova and Sheila C. M. de Sousa . . . . . 139

*N-gram Based Text Classification According To Authorship*
Anđelka Zečević . . . . . 145

*Instance Sampling for Multilingual Coreference Resolution*
Desislava Zhekova . . . . . 150