

Zygmunt Vetulani · Hans Uszkoreit
Marek Kubis (Eds.)

Human Language Technology

Challenges for Computer Science and Linguistics

6th Language and Technology Conference, LTC 2013
Poznań, Poland, December 7–9, 2013
Revised Selected Papers

Contents

Speech Processing

- Boundary Markers in Spontaneous Hungarian Speech 3
András Beke, Mária Gósy, and Viktória Horváth
- Adaptive Prosody Modelling for Improved Synthetic Speech Quality. 16
Moses E. Ekpenyong, Udoinyang G. Inyang, and EmemObong O. Udoh
- Diacritics Restoration in the Slovak Texts Using Hidden Markov Model 29
Daniel Hládek, Ján Staš, and Jozef Juhár
- Temporal and Lexical Context of Diachronic Text Documents for
Automatic Out-Of-Vocabulary Proper Name Retrieval 41
Irina Illina, Dominique Fohr, Georges Linarès, and Imane Nkairi
- Advances in the Slovak Judicial Domain Dictation System. 55
*Milan Rusko, Jozef Juhár, Marian Trnka, Ján Staš, Sakhia Darjaa,
Daniel Hládek, Róbert Sabo, Matúš Pleva, Marian Ritomský,
and Stanislav Ondáš*
- A Revised Comparison of Polish Taggers in the Application for Automatic
Speech Recognition. 68
Aleksander Smywiński-Pohl and Bartosz Ziółko

Morphology

- Automatic Morpheme Slot Identification Using Genetic Algorithm 85
Wondwossen Mulugeta, Michael Gasser, and Baye Yimam
- From Morphology to Lexical Hierarchies and Back. 98
Krešimir Šojat and Matea Srebačić

Parsing Related Issues

- System for Generating Questions Automatically from Given Punjabi Text . . . 115
Vishal Goyal, Shikha Garg, and Umrinderpal Singh
- Hierarchical Amharic Base Phrase Chunking Using HMM
with Error Pruning 126
Abeba Ibrahim and Yaregal Assabie

A Hybrid Approach to Parsing Natural Languages	136
<i>Sardar Jaf and Allan Ramsay</i>	

Experiments in PCFG-like Disambiguation of Constituency Parse Forests for Polish.	146
<i>Marcin Woliński and Dominika Rogozińska</i>	

Computational Semantics

A Method for Measuring Similarity of Books: A Step Towards an Objective Recommender System for Readers	161
<i>Adam Wojciechowski and Krzysztof Gorzynski</i>	

Digital Language Resources

MCBF: Multimodal Corpora Building Framework.	177
<i>Maria Chiara Caschera, Arianna D'Ulizia, Fernando Ferri, and Patrizia Grifoni</i>	

Syntactic Enrichment of LMF Normalized Dictionaries Based on the Context-Field Corpus.	191
<i>Imen Elleuch, Bilel Gargouri, and Abdelmajid Ben Hamadou</i>	

An Example of a Compatible NLP Toolkit.	205
<i>Krzysztof Jassem and Roman Grundkiewicz</i>	

Polish Coreference Corpus	215
<i>Maciej Ogrodniczuk, Katarzyna Głowińska, Mateusz Kopeć, Agata Savary, and Magdalena Zawisławska</i>	

Ontologies and Wordnets

GeoDomainWordNet: Linking the Geonames Ontology to WordNet	229
<i>Francesca Frontini, Riccardo Del Gratta, and Monica Monachini</i>	

Building Wordnet Based Ontologies with Expert Knowledge	243
<i>Jacek Marciniak</i>	

Diagnostic Tools in plWordNet Development Process	255
<i>Maciej Piasecki, Łukasz Burdka, Marek Maziarz, and Michał Kaliński</i>	

Written Text and Document Processing

Simile or Not Simile?: Automatic Detection of Metonymic Relations in Japanese Literal Comparisons	277
<i>Paweł Dybala, Rafał Rzepka, Kenji Araki, and Kohichi Sayama</i>	

Spanish Diacritic Error Detection and Restoration—A Survey. 290
Mans Hulden and Jerid Francom

Identification of Event and Topic for Multi-document Summarization 304
*Fumiyo Fukumoto, Yoshimi Suzuki, Atsuhiko Takasu,
and Suguru Matsuyoshi*

Itemsets-Based Amharic Document Categorization Using an Extended
A Priori Algorithm 317
Abraham Hailu and Yaregal Assabie

NERosetta for the Named Entity Multi-lingual Space 327
Cvetana Krstev, Anđelka Zečević, Duško Vitas, and Tita Kyriacopoulou

A Hybrid Approach to Statistical Machine Translation Between Standard
and Dialectal Varieties. 341
*Friedrich Neubarth, Barry Haddow, Adolfo Hernández Huerta,
and Harald Trost*

Evaluation of Uryupina’s Coreference Resolution Features for Polish 354
Bartłomiej Nitoń

Information and Data Extraction

Aspect-Based Restaurant Information Extraction for the Recommendation
System 371
Ekaterina Pronoza, Elena Yagunova, and Svetlana Volskaya

A Study on Turkish Meronym Extraction Using a Variety
of Lexico-Syntactic Patterns 386
Tuğba Yıldız, Savaş Yıldırım, and Banu Diri

Less-Resourced Languages

A Phonetization Approach for the Forced-Alignment Task in SPPAS 397
Brigitte Bigi

POS Tagging and Less Resources Languages Individuated Features
in CorpusWiki 411
Maarten Janssen

Author Index 421