

# Contents

## Speech

Improved Syllable-Based Text to Speech Synthesis for Tone Language Systems . . . . .	3
<i>Moses Ekpenyong, EmemObong Udoh, Escor Udosen, and Eno-Abasi Urua</i>	
Slovak Automatic Dictation System for Judicial Domain. . . . .	16
<i>Milan Rusko, Jozef Juhár, Marián Trnka, Ján Staš, Sakhia Darjaa, Daniel Hládek, Róbert Sabo, Matúš Pleva, Marián Ritomský, and Martin Lojka</i>	
The Intonation of Backchannel Tokens in Italian Collaborative Dialogues. . .	28
<i>Michelina Savino</i>	

## Parsing

Improved Parsing for Arabic by Combining Diverse Dependency Parsers . . .	43
<i>Maytham Alabbas and Allan Ramsay</i>	
Parsing Coordination Extragrammatically . . . . .	55
<i>Valmi Dufour-Lussier, Bruno Guillaume, and Guy Perrier</i>	
Lexical Disambiguation in LTAG Using Left Context. . . . .	67
<i>Claire Gardent, Yannick Parmentier, Guy Perrier, and Sylvain Schmitz</i>	

## Computational Semantics

Resolving Anaphors in Sanskrit . . . . .	83
<i>Madhav Gopal and Girish Nath Jha</i>	
Unsupervised Coreference Resolution Using a Graph Labeling Approach . . .	93
<i>Nafise Sadat Moosavi and GholamReza GhassemSani</i>	
The XCDC Relations as a Spatio-Temporal Ontology . . . . .	104
<i>Jędrzej Osiński</i>	
Normalization of Term Weighting Scheme for Sentiment Analysis. . . . .	116
<i>Alexander Pak, Patrick Paroubek, Amel Fraisse, and Gil Francopoulo</i>	
Thel, a Language for Utterance Generation in the Thetos System. . . . .	129
<i>Julia Romaniuk, Nina Suszczańska, and Przemysław Szmaj</i>	

An Algorithm for Inconsistency Management in Spatial Knowledge Integration . . . . .	141
<i>Marcin Walas and Krzysztof Jassem</i>	

## Text Analysis

Information Extraction for Czech Based on Syntactic Analysis . . . . .	155
<i>Vít Baisa and Vojtěch Kovář</i>	
Applying Rule-Based Normalization to Different Types of Historical Texts — An Evaluation . . . . .	166
<i>Marcel Bollmann, Florian Petran, and Stefanie Dipper</i>	
A Rule Based Method for the Identification of TAM Features in a PoS Tagged Corpus . . . . .	178
<i>Narayan Choudhary, Pramod Pandey, and Girish Nath Jha</i>	
Digging for Names in the Mountains: Combined Person Name Recognition and Reference Resolution for German Alpine Texts . . . . .	189
<i>Sarah Ebling, Rico Sennrich, and David Klaper</i>	
Enhancing Labeled Data Using Unlabeled Data for Topic Tracking . . . . .	201
<i>Fumiyo Fukumoto, Yoshimi Suzuki, and Takeshi Yamamoto</i>	
Temporal Expression Recognition Using Dependency Trees . . . . .	213
<i>Paweł Mazur and Robert Dale</i>	
Pattern Mining for Named Entity Recognition . . . . .	226
<i>Damien Nouvel, Jean-Yves Antoine, and Nathalie Friburger</i>	
Lexical Bundles in Swedish Secondary School Textbooks . . . . .	238
<i>Judy Ribeck and Lars Borin</i>	

## Text Annotation

Active Learning to Speed-Up the Training Process for Dialogue Act Labelling . . . . .	253
<i>Fabrizio Ghigi, Carlos-D. Martínez-Hinarejos, and José-Miguel Benedí</i>	
Direct and Wordgraph-Based Confidence Measures in Dialogue Annotation with N-Gram Transducers . . . . .	264
<i>Carlos-D. Martínez-Hinarejos, Vicent Tamarit, and José-Miguel Benedí</i>	
Orwell's 1984 — From Simple to Multi-word Units . . . . .	276
<i>Cvetana Krstev, Duško Vitas, and Aleksandra Trtovac</i>	

Application of Audio and Video Processing Methods for Language Research and Documentation: The AVATeCH Project. . . . .	288
<i>Przemysław Lenkiewicz, Sebastian Drude, Anna Lenkiewicz, Binyam Gebrekidan Gebre, Stefano Masneri, Oliver Schreer, Jochen Schwenninger, and Rolf Bardeli</i>	

## Language Resources: General Issues

Crowdsourcing for Language Resource Development: Criticisms About Amazon Mechanical Turk Overpowering Use . . . . .	303
<i>Karĕn Fort, Gilles Adda, Benoît Sagot, Joseph Mariani, and Alain Couillault</i>	
Extending a Tool Resource Framework with U-Compare . . . . .	315
<i>Michael Rosner, Andrew Attard, Paul Thompson, Albert Gatt, and Sophia Ananiadou</i>	

## Language Resources: Ontologies and Wordnets

Aligning GermaNet Senses with Wiktionary Sense Definitions . . . . .	329
<i>Verena Henrich, Erhard Hinrichs, and Tatiana Vodolazova</i>	
A Tool for Transforming WordNet-Like Databases. . . . .	343
<i>Marek Kubis</i>	
KABA Subject Heading Language as the Main Resource Subject Organization Tool in a Semantic Knowledge Base . . . . .	356
<i>Cezary Mazurek, Krzysztof Sielski, Justyna Walkowska, and Marcin Werla</i>	
Enhancing Tagging Systems by Wordnet Based Ontologies. . . . .	367
<i>Jacek Marciniak</i>	
Natural Language Ontology of Action: A Gap with Huge Consequences for Natural Language Understanding and Machine Translation. . . . .	379
<i>Massimo Moneglia</i>	
Classification-Based Extension of Wordnets from Heterogeneous Resources . . .	396
<i>Benoît Sagot and Darja Fišer</i>	
PolNet – Polish WordNet . . . . .	408
<i>Zygmunt Vetulani</i>	

## Machine Translation

Improving the Distribution of N-Grams in Phrase Tables Obtained by the Sampling-Based Method . . . . .	419
<i>Juan Luo, Adrien Lardilleux, and Yves Lepage</i>	

Marker-Based Chunking in Eleven European Languages for Analogy-Based Translation . . . . .	432
<i>Kota Takeya and Yves Lepage</i>	
Comparing CBMT Approaches for German-Romanian . . . . .	445
<i>Monica Gavrilă and Natalia Elita</i>	
Text Genre – An Unexplored Parameter in Statistical Machine Translation . . .	456
<i>Monica Gavrilă and Cristina Vertan</i>	
<b>Problems Concerning Less Resourced Languages</b>	
Detecting Gaps in Language Resources and Tools in the Project CESAR . . .	471
<i>Marko Tadić, Tamás Váradi, Radovan Garabík, Svetla Koeva, Maciej Ogrodniczuk, and Duško Vitas</i>	
A First LVCSR System for Luxembourgish, a Low-Resourced European Language. . . . .	479
<i>Martine Adda-Decker, Lori Lamel, Gilles Adda, and Thomas Lavergne</i>	
Developing LRs for Non-scheduled Indian Languages . . . . .	491
<i>Ritesh Kumar, Bornini Lahiri, and Deepak Alok</i>	
Quizzes on Tap: Exporting a Test Generation System from One Less-Resourced Language to Another . . . . .	502
<i>Montse Maritxalar, Elaine Uí Donnchadha, Jennifer Foster, and Monica Ward</i>	
A Multilingual Text Normalization Approach . . . . .	515
<i>Brigitte Bigi</i>	
Creating Multilingual Parallel Corpora in Indian Languages . . . . .	527
<i>Narayan Choudhary and Girish Nath Jha</i>	
Inducing Grammars from IGT . . . . .	538
<i>Lars Hellan and Dorothee Beermann</i>	
<b>Author Index</b> . . . . .	549

Human Language Technology Challenges for Computer  
Science and Linguistics

5th Language and Technology Conference, LTC 2011,  
Poznań, Poland, November 25--27, 2011, Revised  
Selected Papers

Vetulani, Z.; Mariani, J. (Eds.)

2014, XVI, 550 p. 109 illus., Softcover

ISBN: 978-3-319-08957-7