

# **9th International Conference on Language Resources and Evaluation**

**(LREC-2014)**

**Reykjavik, Iceland  
26-31 May 2014**

**Volume 1 of 5**

ISBN: 978-1-63266-621-5

**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© (2014) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2014)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

## Volume 1

<b>CLIPS STYLOMETRY INVESTIGATION (CSI) CORPUS: A DUTCH CORPUS FOR THE DETECTION OF AGE, GENDER, PERSONALITY, SENTIMENT AND DECEPTION IN TEXT</b> .....	1
<i>Ben Verhoeven, Walter Daelemans</i>	
<b>ON PARAPHRASE IDENTIFICATION CORPORA</b> .....	6
<i>Vasile Rus, Rajendra Banjade, Mihai Lintean</i>	
<b>A CORPUS OF COMPARISONS IN PRODUCT REVIEWS</b> .....	14
<i>Wiltrud Kessler, Jonas Kuhn</i>	
<b>GENERATING AND USING PROBABILISTIC MORPHOLOGICAL RESOURCES FOR THE BIOMEDICAL DOMAIN</b> .....	21
<i>Vincent Claveau, Ewa Kijak</i>	
<b>A SYSTEM FOR EXPERIMENTS WITH DEPENDENCY PARSERS</b> .....	28
<i>Kiril Simov, Iliana Simova, Ginka Ivanova, Maria Mateva, Petya Osenova</i>	
<b>SOCKPUPPET DETECTION IN WIKIPEDIA: A CORPUS OF REAL-WORLD DECEPTIVE WRITING FOR LINKING IDENTITIES</b> .....	33
<i>Thamar Solorio, Ragib Hasan, Mainul Mizan</i>	
<b>TEXTUAL EMIGRATION ANALYSIS (TEA)</b> .....	37
<i>Andre Blessing, Jonas Kuhn</i>	
<b>PHONEME SET DESIGN USING ENGLISH SPEECH DATABASE BY JAPANESE FOR DIALOGUE-BASED ENGLISH CALL SYSTEMS</b> .....	42
<i>Xiaoyun Wang, Jinsong Zhang, Masafumi Nishida, Seiichi Yamamoto</i>	
<b>TOWARD A UNIFYING MODEL FOR OPINION, SENTIMENT AND EMOTION INFORMATION EXTRACTION</b> .....	46
<i>Amel Fraisse, Patrick Paroubek</i>	
<b>TOWARDS AUTOMATIC QUALITY ASSESSMENT OF COMPONENT METADATA</b> .....	52
<i>Thorsten Trippel, Daan Broeder, Matej Durco, Oddrun Ohren</i>	
<b>PROPBANK: SEMANTICS OF NEW PREDICATE TYPES</b> .....	58
<i>Claire Bonial, Julia Bonn, Kathryn Conger, Jena D. Hwang, Martha Palmer</i>	
<b>RESTFUL ANNOTATION AND EFFICIENT COLLABORATION</b> .....	65
<i>Jonathan Wright</i>	
<b>ALICO: A MULTIMODAL CORPUS FOR THE STUDY OF ACTIVE LISTENING</b> .....	72
<i>Hendrik Buschmeier, Zofia Malisz, Joanna Skubisz, Marcin Wlodarczak, Ipke Wachsmuth, Stefan Kopp, Petra Wagner</i>	
<b>POLITA: A MULTITAGGER FOR POLISH</b> .....	78
<i>Lukasz Kobylinski</i>	
<b>A CORPUS OF PARTICIPANT ROLES IN CONTENTIOUS DISCUSSIONS</b> .....	84
<i>Siddharth Jain, Archana Bhatia, Angeliqne Rein, Eduard Hovy</i>	
<b>BILINGUAL DICTIONARY CONSTRUCTION WITH TRANSLITERATION FILTERING</b> .....	90
<i>John Richardson, Toshiaki Nakazawa, Sadao Kurohashi</i>	
<b>REVISING THE ANNOTATION OF A BROADCAST NEWS CORPUS: A LINGUISTIC APPROACH</b> .....	95
<i>Vera Cabarrão, Helena Moniz, Fernando Batista, Ricardo Ribeiro, Nuno Mamede, Hugo Meinedo, Isabel Trancoso, Ana Isabel Mata, David Martins De Matos</i>	
<b>GOLD-STANDARD FOR TOPIC-SPECIFIC SENTIMENT ANALYSIS OF ECONOMIC TEXTS</b> .....	101
<i>Pyry Takala, Pekka Malo, Ankur Sinha, Oskar Ahlgren</i>	
<b>VISUALIZATION OF LANGUAGE RELATIONS AND FAMILIES: MULTITREE</b> .....	107
<i>Damir Cavar, Malgorzata Cavar</i>	
<b>HIEVE: A CORPUS FOR EXTRACTING EVENT HIERARCHIES FROM NEWS STORIES</b> .....	111
<i>Goran Glavas, Jan Snajder, Marie-Francine Moens, Parisa Kordjamshidi</i>	
<b>IMAGE ANNOTATION WITH ISO-SPACE: DISTINGUISHING CONTENT FROM STRUCTURE</b> .....	117
<i>James Pustejovsky, Zachary Yocum</i>	
<b>THE ETAPE SPEECH PROCESSING EVALUATION</b> .....	123
<i>Olivier Galibert, Jeremy Leixa, Gilles Adda, Khalid Choukri, Guillaume Gravier</i>	
<b>PANLEX: BUILDING A RESOURCE FOR PANLINGUAL LEXICAL TRANSLATION</b> .....	128
<i>David Kamholz, Jonathan Pool, Susan Colowick</i>	

<b>LARGE SMT DATA-SETS EXTRACTED FROM WIKIPEDIA</b> .....	134
<i>Dan Tufi</i>	
<b>NOMLEX-PT: A LEXICON OF PORTUGUESE NOMINALIZATIONS</b> .....	142
<i>Valeria De Paiva, Livy Real, Alexandre Rademaker, Gerard De Melo</i>	
<b>VERTA: FACING A MULTILINGUAL EXPERIENCE OF A LINGUISTICALLY-BASED MT EVALUATION</b> .....	150
<i>Elisabet Comelles, Jordi Atserias, Victoria Arranz, Irene Castellon, Jordi Sese</i>	
<b>CORPUS AND METHOD FOR IDENTIFYING CITATIONS IN NON-ACADEMIC TEXT</b> .....	157
<i>Yifan He, Adam Meyers</i>	
<b>BUILDING A DATASET FOR SUMMARIZATION AND KEYWORD EXTRACTION FROM EMAILS</b> .....	161
<i>Vanessa Loza, Shibamouli Lahiri, Rada Mihalcea, Po-Hsiang Lai</i>	
<b>IMPROVING OPEN RELATION EXTRACTION VIA SENTENCE RE-STRUCTURING</b> .....	167
<i>Jordan Schmeidek, Denilson Barbosa</i>	
<b>HOW TO USE LESS FEATURES AND REACH BETTER PERFORMANCE IN AUTHOR GENDER IDENTIFICATION</b> .....	171
<i>Juan Soler, Leo Wanner</i>	
<b>SEMI-SUPERVISED METHODS FOR EXPANDING PSYCHOLINGUISTICS NORMS BY INTEGRATING DISTRIBUTIONAL SIMILARITY WITH THE STRUCTURE OF WORDNET</b> .....	176
<i>Michael Mohler, Marc Tomlinson, David Bracewell, Bryan Rink</i>	
<b>LANGUAGESINDANGER.EU - INCLUDING MULTIMEDIA LANGUAGE RESOURCES TO DISSEMINATE KNOWLEDGE AND CREATE EDUCATIONAL MATERIAL ON LESS-RESOURCED LANGUAGES</b> .....	183
<i>Dagmar Jung, Katarzyna Klessa, Zsuzsa Duray, Beatrix Oszkó, Mária Sipos, Sándor Szeverényi, Zsuzsa Várnai, Trilsbeek Paul, Tamás Váradi</i>	
<b>A CROSS-LANGUAGE CORPUS FOR STUDYING THE PHONETICS AND PHONOLOGY OF PROMINENCE</b> .....	189
<i>Bistra Andreeva, William Barry, Jacques Koreman</i>	
<b>DELEX, A FREELY-AVAIBLE, LARGE-SCALE AND LINGUISTICALLY GROUNDED MORPHOLOGICAL LEXICON FOR GERMAN</b> .....	194
<i>Benoît Sagot</i>	
<b>USING RESOURCE-RICH LANGUAGES TO IMPROVE MORPHOLOGICAL ANALYSIS OF UNDER-RESOURCED LANGUAGES</b> .....	201
<i>Peter Baumann, Janet Pierrehumbert</i>	
<b>ACCOMMODATIONS IN TUSCANY AS LINKED DATA</b> .....	206
<i>Clara Bacciu, Angelica Lo Duca, Andrea Marchetti, Maurizio Tesconi</i>	
<b>THE DWAN FRAMEWORK: APPLICATION OF A WEB ANNOTATION FRAMEWORK FOR THE GENERAL HUMANITIES TO THE DOMAIN OF LANGUAGE RESOURCES</b> .....	210
<i>Przemyslaw Lenkiewicz, Olha Shkaravska, Twan Goosen, Daan Broeder, Menzo Windhouwer, Stephanie Roth, Olof Olsson</i>	
<b>CROATIAN MEMORIES</b> .....	216
<i>Arjan Van Hessen, Franciska De Jong, Stef Scagliola, Tanja Petrovic</i>	
<b>COMBINING ELICITED IMITATION AND FLUENCY FEATURES FOR ORAL PROFICIENCY MEASUREMENT</b> .....	223
<i>Deryle Lonsdale, Carl Christensen</i>	
<b>SEMANTIC SEARCH IN DOCUMENTS ENRICHED BY LOD-BASED ANNOTATIONS</b> .....	229
<i>Pavel Smrz, Jan Kouril</i>	
<b>A META-DATA DRIVEN PLATFORM FOR SEMI-AUTOMATIC CONFIGURATION OF ONTOLOGY MEDIATORS</b> .....	233
<i>Manuel Fiorelli, Maria Teresa Pazienza, Armando Stellato</i>	
<b>SEMANTIC APPROACHES TO SOFTWARE COMPONENT RETRIEVAL WITH ENGLISH QUERIES</b> .....	239
<i>Huijing Deng, Grzegorz Chrupala</i>	
<b>THE ELLOGON PATTERN ENGINE: CONTEXT-FREE GRAMMARS OVER ANNOTATIONS</b> .....	244
<i>Georgios Petasis</i>	
<b>MISSED OPPORTUNITIES IN TRANSLATION MEMORY MATCHING</b> .....	250
<i>Friedel Wolff, Laurette Pretorius, Paul Buitelaar</i>	
<b>UNIVERSAL STANFORD DEPENDENCIES: A CROSS-LINGUISTIC TYPOLOGY</b> .....	256
<i>Marie-Catherine De Marneffe, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, Christopher D. Manning</i>	
<b>GETTING RELIABLE ANNOTATIONS FOR SARCASM IN ONLINE DIALOGUES</b> .....	264
<i>Reid Swanson, Stephanie Lukin, Luke Eisenberg, Thomas Corcoran, Marilyn Walker</i>	

<b>COLLABORATIVELY ANNOTATING MULTILINGUAL PARALLEL CORPORA IN THE BIOMEDICAL DOMAIN?SOME MANTRAS</b> .....	272
<i>Johannes Hellrich, Simon Clematide, Udo Hahn, Dietrich Rebholz-Schuhmann</i>	
<b>ON THE IMPORTANCE OF TEXT ANALYSIS FOR STOCK PRICE PREDICTION</b> .....	280
<i>Heeyoung Lee, Mihai Surdeanu, Bill Maccartney, Dan Jurafsky</i>	
<b>ON THE USE OF A FUZZY CLASSIFIER TO SPEED UP THE SP TOBI LABELING OF THE GLISSANDO SPANISH CORPUS</b> .....	286
<i>David Escudero, Aguilar-Cuevas Lourdes, González-Ferreras César, Gutiérrez-González Yurena, Valentín Cardeñoso-Payo</i>	
<b>DEFINITION PATTERNS FOR PREDICATIVE TERMS IN SPECIALIZED LEXICAL RESOURCES</b> .....	294
<i>Antonio San Martín, Marie-Claude L'Homme</i>	
<b>NATIVE LANGUAGE IDENTIFICATION USING LARGE, LONGITUDINAL DATA</b> .....	302
<i>Xiao Jiang, Yufan Guo, Jeroen Geertzen, Dora Alexopoulou, Lin Sun, Anna Korhonen</i>	
<b>PRODUCTION OF PHRASE TABLES IN 11 EUROPEAN LANGUAGES USING AN IMPROVED SUB-SENTENTIAL ALIGNER</b> .....	306
<i>Juan Luo, Yves Lepage</i>	
<b>CONSTRUCTION AND ANNOTATION OF A FRENCH FOLKSTALE CORPUS</b> .....	312
<i>Anne Garcia-Fernandez, Anne-Laure Ligozat, Anne Vilnat</i>	
<b>THE MAKING OF ANCIENT GREEK WORDNET</b> .....	318
<i>Yuri Bizzoni, Federico Boschetti, Harry Diakoff, Riccardo Del Gratta, Monica Monachini, Gregory Crane</i>	
<b>ENRICHING ODIN</b> .....	326
<i>Fei Xia, William Lewis, Michael Wayne Goodman, Joshua Crowgey, Emily M. Bender</i>	
<b>TURKISH TREEBANK AS A GOLD STANDARD FOR MORPHOLOGICAL DISAMBIGUATION AND ITS INFLUENCE ON PARSING</b> .....	333
<i>Ozlem Cetinoglu</i>	
<b>CRODERIV: A NEW RESOURCE FOR PROCESSING CROATIAN MORPHOLOGY</b> .....	339
<i>Kresimir Sojat, Matea Srebacic, Marko Tadic, Tin Pavelic</i>	
<b>BUILDING A DATABASE OF JAPANESE ADJECTIVE EXAMPLES FROM SPECIAL PURPOSE WEB CORPORA</b> .....	344
<i>Masaya Yamaguchi</i>	
<b>PRAALINE: INTEGRATING TOOLS FOR SPEECH CORPUS RESEARCH</b> .....	349
<i>George Christodoulides</i>	
<b>EXTRACTING A BILINGUAL SEMANTIC GRAMMAR FROM FRAMENET-ANNOTATED CORPORA</b> .....	353
<i>Dana Dannells, Normunds Gruzitis</i>	
<b>MAPPING BETWEEN ENGLISH STRINGS AND REENTRANT SEMANTIC GRAPHS</b> .....	361
<i>Fabienne Braune, Daniel Bauer, Kevin Knight</i>	
<b>THE KIEZDEUTSCH KORPUS (KIDKO) RELEASE 1.0</b> .....	367
<i>Ines Rehbein, Sören Schalowski, Heike Wiese</i>	
<b>ETYMOLOGICAL WORDNET: TRACING THE HISTORY OF WORDS</b> .....	375
<i>Gerard De Melo</i>	
<b>THE INTERPLAY BETWEEN LEXICAL AND SYNTACTIC RESOURCES IN INCREMENTAL PARSEBANKING</b> .....	382
<i>Victoria Rosén, Petter Haugereid, Martha Thunes, Gyri S. Losnegaard, Helge Dyvik</i>	
<b>INTEROPERABILITY AND CUSTOMISATION OF ANNOTATION SCHEMATA IN ARGO</b> .....	390
<i>Rafal Rak, Jacob Carter, Andrew Rowley, Riza Theresa Batista-Navarro, Sophia Ananiadou</i>	
<b>POLISH COREFERENCE CORPUS IN NUMBERS</b> .....	396
<i>Maciej Ogrodniczuk, Mateusz Kopec, Agata Savary</i>	
<b>A GOLD STANDARD DEPENDENCY CORPUS FOR ENGLISH</b> .....	401
<i>Natalia Silveira, Timothy Dozat, Marie-Catherine De Marneffe, Samuel Bowman, Miriam Connor, John Bauer, Chris Manning</i>	
<b>DERIVBASE.HR: A HIGH-COVERAGE DERIVATIONAL MORPHOLOGY RESOURCE FOR CROATIAN</b> .....	409
<i>Jan Šnajder</i>	
<b>EXTRACTING INFORMATION FOR CONTEXT-AWARE MEETING PREPARATION</b> .....	416
<i>Simon Scerri, Behrang Q. Zadeh, Maciej Dabrowski, Ismael Rivera</i>	
<b>A REPOSITORY OF STATE OF THE ART AND COMPETITIVE BASELINE SUMMARIES FOR GENERIC NEWS SUMMARIZATION</b> .....	421
<i>Kai Hong, John Conroy, Benoit Favre, Alex Kulesza, Hui Lin, Ani Nenkova</i>	

<b>COLLECTING NATURAL SMS AND CHAT CONVERSATIONS IN MULTIPLE LANGUAGES: THE BOLT PHASE 2 CORPUS .....</b>	<b>430</b>
<i>Zhiyi Song, Stephanie Strassel, Haejoong Lee, Kevin Walker, Jonathan Wright, Jennifer Garland, Dana Fore, Brian Gainor, Preston Cabe, Thomas Thomas, Brendan Callahan, Ann Sawyer</i>	
<b>COMPARING THE QUALITY OF FOCUSED CRAWLERS AND OF THE TRANSLATION RESOURCES OBTAINED FROM THEM.....</b>	<b>436</b>
<i>Bruno Laranjeira, Viviane Moreira, Aline Villavicencio, Carlos Ramisch, Maria José Finatto</i>	
<b>AUGMENTING ENGLISH ADJECTIVE SENSES WITH SUPERSENSES .....</b>	<b>443</b>
<i>Yulia Tsvetkov, Nathan Schneider, Dirk Hovy, Archana Bhatia, Manaal Faruqui, Chris Dyer</i>	
<b>N-GRAM COUNTS AND LANGUAGE MODELS FROM THE COMMON CRAWL.....</b>	<b>450</b>
<i>Christian Buck, Kenneth Heafield, Bas Van Ooyen</i>	
<b>COLLEX.EN: AUTOMATICALLY GENERATING AND EVALUATING A FULL-FORM LEXICON FOR ENGLISH .....</b>	<b>456</b>
<i>Tim Vor Der Brück, Alexander Mehler, Zahurul Islam</i>	
<b>EVALUATION OF SIMPLE DISTRIBUTIONAL COMPOSITIONAL OPERATIONS ON LONGER TEXTS.....</b>	<b>461</b>
<i>Tamara Polajnar, Laura Rimell, Stephen Clark</i>	
<b>CREATING SUMMARIZATION SYSTEMS WITH SUMMA .....</b>	<b>465</b>
<i>Horacio Saggion</i>	
<b>BIOGRAPHYNET: METHODOLOGICAL ISSUES WHEN NLP SUPPORTS HISTORICAL RESEARCH.....</b>	<b>472</b>
<i>Antske Fokkens, Serge Ter Braake, Niels Ockeloen, Piek Vossen, Susan Legene, Guus Schreiber</i>	
<b>ENHANCING THE TED-LIUM CORPUS WITH SELECTED DATA FOR LANGUAGE MODELING AND MORE TED TALKS .....</b>	<b>480</b>
<i>Anthony Rousseau, Paul Deléglise, Yannick Estève</i>	
<b>ENRICHMENT OF BILINGUAL DICTIONARY THROUGH NEWS STREAM DATA .....</b>	<b>485</b>
<i>Ajay Dubey, Parth Gupta, Vasudeva Varma, Paolo Rosso</i>	
<b>SLOWCROWD: A CROWDSOURCING TOOL FOR LEXICOGRAPHIC TASKS.....</b>	<b>490</b>
<i>Darja Fiser, Ales Tavcar, Tomaz Erjavec</i>	
<b>A LARGE SCALE DATABASE OF STRONGLY-RELATED EVENTS IN JAPANESE.....</b>	<b>495</b>
<i>Tomohide Shibata, Shotaro Kohama, Sadao Kurohashi</i>	
<b>FLELEX: A GRADED LEXICAL RESOURCE FOR FRENCH FOREIGN LEARNERS .....</b>	<b>501</b>
<i>Thomas Francois, Nùria Gala, Patrick Watrin, Cédric Faron</i>	
<b>BUILDING DOMAIN SPECIFIC BILINGUAL DICTIONARIES.....</b>	<b>509</b>
<i>Lucas Hilgert, Lucelene Lopes, Artur Freitas, Renata Vieira, Denise Hogetop, Aline Vamin</i>	
<b>A CORPUS OF MACHINE TRANSLATION ERRORS EXTRACTED FROM TRANSLATION STUDENTS EXERCISES.....</b>	<b>515</b>
<i>Guillaume Wisniewski, Natalie Kubler, François Yvon</i>	
<b>FINDING ROMANIZED ARABIC DIALECT IN CODE-MIXED TWEETS.....</b>	<b>519</b>
<i>Clare Voss, Stephen Tratz, Jamal Laoudi, Douglas Briesch</i>	
<b>CO-TRAINING FOR CLASSIFICATION OF LIVE OR STUDIO MUSIC RECORDINGS .....</b>	<b>524</b>
<i>Nicolas Auguin, Pascale Fung</i>	
<b>#MYGOAL: FINDING MOTIVATIONS ON TWITTER.....</b>	<b>528</b>
<i>Marc Tomlinson, David Bracewell, Wayne Krug, David Hinote</i>	
<b>THE RATS COLLECTION: SUPPORTING HLT RESEARCH WITH DEGRADED AUDIO DATA.....</b>	<b>534</b>
<i>David Graff, Kevin Walker, Stephanie Strassel, Xiaoyi Ma, Karen Jones, Ann Sawyer</i>	
<b>MODELING LANGUAGE PROFICIENCY USING IMPLICIT FEEDBACK.....</b>	<b>542</b>
<i>Chris Hokamp, Rada Mihalcea, Peter Schuelke</i>	
<b>EVENT EXTRACTION USING DISTANT SUPERVISION .....</b>	<b>546</b>
<i>Kevin Reschke, Martin Jankowiak, Mihai Surdeanu, Christopher Manning, Daniel Jurafsky</i>	
<b>FIRST INSIGHT INTO QUALITY-ADAPTIVE DIALOGUE .....</b>	<b>551</b>
<i>Stefan Ultes, Hüseyin Dikme, Wolfgang Minker</i>	
<b>TLXCALA: A MULTILINGUAL CORPUS OF INDEPENDENT NEWS.....</b>	<b>557</b>
<i>Antonio Toral</i>	
<b>CREATING AND USING LARGE MONOLINGUAL PARALLEL CORPORA FOR SENTENTIAL PARAPHRASE GENERATION.....</b>	<b>561</b>
<i>Sander Wubben, Antal Van Den Bosch, Emiel Kraemer</i>	
<b>BENCHMARKING OF ENGLISH-HINDI PARALLEL CORPORA.....</b>	<b>569</b>
<i>Jayendra Rakesh Yeka, Prasanth Kolachina, Dipti Misra Sharma</i>	
<b>A NEW FRAMEWORK FOR SIGN LANGUAGE RECOGNITION BASED ON 3D HANDSHAPE IDENTIFICATION AND LINGUISTIC MODELING .....</b>	<b>576</b>
<i>Mark Dilsizian, Polina Yanovich, Shu Wang, Carol Neidle, Dimitris Metaxas</i>	

<b>EVALUATING IMPROVISED HIP HOP LYRICS - CHALLENGES AND OBSERVATIONS</b> .....	582
<i>Karteek Addanki, Dekai Wu</i>	
<b>VOTTER CORPUS: A CORPUS OF SOCIAL POLLING LANGUAGE</b> .....	590
<i>Nathan Green, Septina Dian Larasati</i>	
<b>SINOCOREFERENCER: AN END-TO-END CHINESE EVENT COREFERENCE RESOLVER</b> .....	595
<i>Chen Chen, Vincent Ng</i>	
<b>DEVELOPING AN EGYPTIAN ARABIC TREEBANK: IMPACT OF DIALECTAL MORPHOLOGY ON ANNOTATION AND TOOL DEVELOPMENT</b> .....	602
<i>Mohamed Maamouri, Ann Bies, Seth Kulick, Michael Ciul, Nizar Habash, Ramy Eskander</i>	
<b>A GERMAN TWITTER SNAPSHOT</b> .....	609
<i>Tatjana Scheffler</i>	
<b>A COMPARATIVE EVALUATION METHODOLOGY FOR NLG IN INTERACTIVE SYSTEMS</b> .....	615
<i>Helen Hastie, Anja Belz</i>	
<b>RELATING FRAMES AND CONSTRUCTIONS IN JAPANESE FRAMENET</b> .....	623
<i>Kyoko Ohara</i>	
<b>TMO ? THE FEDERATED ONTOLOGY OF THE TRENDMINER PROJECT</b> .....	627
<i>Hans-Ulrich Krieger, Thierry Declerck</i>	
<b>A GRAPH-BASED APPROACH FOR COMPUTING FREE WORD ASSOCIATIONS</b> .....	635
<i>Gemma Bel Enguix, Reinhard Rapp, Michael Zock</i>	
<b>DEVELOPING TEXT RESOURCES FOR TEN SOUTH AFRICAN LANGUAGES</b> .....	642
<i>Roald Eiselen, Martin Puttkammer</i>	
<b>MOMRESP: A BAYESIAN MODEL FOR MULTI-ANNOTATOR DOCUMENT LABELING</b> .....	648
<i>Paul Felt, Robbie Haertel, Eric Ringger, Kevin Seppi</i>	
<b>TOWARDS AUTOMATIC DETECTION OF NARRATIVE STRUCTURE</b> .....	656
<i>Jessica Ouyang, Kathy McKeown</i>	
<b>OPENLOGOS SEMANTICO-SYNTACTIC KNOWLEDGE-RICH BILINGUAL DICTIONARIES</b> .....	664
<i>Anabela Barreiro, Fernando Batista, Ricardo Ribeiro, Helena Moniz, Isabel Trancoso</i>	
<b>USING LARGE BIOMEDICAL DATABASES AS GOLD ANNOTATIONS FOR AUTOMATIC RELATION EXTRACTION</b> .....	672
<i>Tilia Ellendorff, Fabio Rinaldi, Simon Clematide</i>	
<b>CROWDSOURCING FOR THE IDENTIFICATION OF EVENT NOMINALS: AN EXPERIMENT</b> .....	678
<i>Rachele Sprugnoli, Alessandro Lenci</i>	
<b>AUTOMATIC REFINEMENT OF SYNTACTIC CATEGORIES IN CHINESE WORD STRUCTURES</b> .....	685
<i>Jianqiang Ma</i>	
<b>INCORPORATING ALTERNATE TRANSLATIONS INTO ENGLISH TRANSLATION TREEBANK</b> .....	691
<i>Ann Bies, Justin Mott, Seth Kulick, Jennifer Garland, Colin Warner</i>	
<b>ZMORGE: A GERMAN MORPHOLOGICAL LEXICON EXTRACTED FROM WIKTIONARY</b> .....	697
<i>Rico Sennrich, Beat Kunz</i>	
<b>THARWA: A LARGE SCALE DIALECTAL ARABIC - STANDARD ARABIC - ENGLISH LEXICON</b> .....	702
<i>Mona Diab, Mohamed Albadrashiny, Maryam Aminian, Mohammed Attia, Heba Elfardy, Nizar Habash, Abdelati Hawwari, Wael Salloum, Pradeep Dasigi, Ramy Eskander</i>	
<b>A SKOS-BASED SCHEMA FOR TEI ENCODED DICTIONARIES AT ICLTT</b> .....	710
<i>Thierry Declerck, Karlheinz Mörth, Eveline Wandl-Vogt</i>	
<b>SEMANTIC TECHNOLOGIES FOR QUERYING LINGUISTIC ANNOTATIONS: AN EXPERIMENT FOCUSING ON GRAPH-STRUCTURED DATA</b> .....	714
<i>Milen Kouylekov, Stephan Oepen</i>	
<b>ELICITING AND ANNOTATING UNCERTAINTY IN SPOKEN LANGUAGE</b> .....	720
<i>Heather Pon-Barry, Stuart Shieber, Nicholas Longenbaugh</i>	
<b>A HIERARCHICAL TAXONOMY FOR CLASSIFYING HARDNESS OF INFERENCE TASKS</b> .....	726
<i>Martin Gleize, Brigitte Grau</i>	
<b>AUTOMATIC METHODS FOR THE EXTENSION OF A BILINGUAL DICTIONARY USING COMPARABLE CORPORA</b> .....	733
<i>Michael Rosner, Kurt Sultana</i>	
<b>A METHOD FOR BUILDING BURST-ANNOTATED CO-OCCURRENCE NETWORKS FOR ANALYSING TRENDS IN TEXTUAL DATA</b> .....	741
<i>Yutaka Mitsuishi, Vit Novacek, Pierre-Yves Vandenbussche</i>	

<b>CASA DE LA LHÉNGUA: A SET OF LANGUAGE RESOURCES AND NATURAL LANGUAGE PROCESSING TOOLS FOR MIRANDESE .....</b>	<b>747</b>
<i>José Pedro Ferreira, Cristiano Chesi, Daan Baldewijns, Fernando Miguel Pinto, Margarita Correia, Daniela Braga, Hyongsil Cho, Amadeu Ferreira, Miguel Dias</i>	
<b>UNTRAINED FORCED ALIGNMENT OF TRANSCRIPTIONS AND AUDIO FOR LANGUAGE DOCUMENTATION CORPORA USING WEBMAUS .....</b>	<b>752</b>
<i>Jan Strunk, Florian Schiel, Frank Seifart</i>	
<b>MULTIVAL - TOWARDS A MULTILINGUAL VALENCE LEXICON .....</b>	<b>760</b>
<i>Lars Hellan, Dorothee Beermann, Tore Bruland, Mary Esther Kropp Dakubu, Montserrat Marimon</i>	
<b>THE SWEET-HOME SPEECH AND MULTIMODAL CORPUS FOR HOME AUTOMATION INTERACTION .....</b>	<b>768</b>
<i>Michel Vacher, Benjamin Lecouteux, Pedro Chahuara, François Portet, Brigitte Meillon, Nicolas Bonnefond</i>	
<b>GLOBAL INTELLIGENT CONTENT: ACTIVE CURATION OF LANGUAGE RESOURCES USING LINKED DATA .....</b>	<b>776</b>
<i>David Lewis, Rob Brennan, Leroy Finn, Dominic Jones, Alan Meehan, Declan O'Sullivan, Sebastian Hellmann, Felix Sasaki</i>	
<b>ON THE ROMANCE LANGUAGES MUTUAL INTELLIGIBILITY .....</b>	<b>780</b>
<i>Liviu Dinu, Alina Maria Ciobanu</i>	
<b>AGGREGATION METHODS FOR EFFICIENT COLLOCATION DETECTION .....</b>	<b>786</b>
<i>Anca Dinu, Liviu Dinu, Ionut Sorodoc</i>	
<b>ANNOTATING INTER-SENTENCE TEMPORAL RELATIONS IN CLINICAL NOTES .....</b>	<b>791</b>
<i>Jennifer D'Souza, Vincent Ng</i>	
<b>TERMINOLOGY LOCALIZATION GUIDELINES FOR THE NATIONAL SCENARIO .....</b>	<b>799</b>
<i>Juris Borzovs, Ilze Ilzina, Iveta Keiša, Marcis Pinnis, Andrejs Vasiljevs</i>	
<b>TOOLS FOR ARABIC NATURAL LANGUAGE PROCESSING: A CASE STUDY IN QALQALAH PROSODY .....</b>	<b>805</b>
<i>Claire Brierley, Majdi Sawalha, Eric Atwell</i>	
<b>TEENAGE AND ADULT SPEECH IN SCHOOL CONTEXT: BUILDING AND PROCESSING A CORPUS OF EUROPEAN PORTUGUESE .....</b>	<b>810</b>
<i>Ana Isabel Mata, Helena Moniz, Fernando Batista, Julia Hirschberg</i>	
<b>A UNIFIED ANNOTATION SCHEME FOR THE SEMANTIC/PRAGMATIC COMPONENTS OF DEFINITENESS .....</b>	<b>816</b>
<i>Archana Bhatia, Mandy Simons, Lori Levin, Yulia Tsvetkov, Chris Dyer, Jordan Bender</i>	
<b>ALIGNING PREDICATE-ARGUMENT STRUCTURES FOR PARAPHRASE FRAGMENT EXTRACTION .....</b>	<b>823</b>
<i>Michaela Regneri, Rui Wang, Manfred Pinkal</i>	
<b>AN EVALUATION OF THE ROLE OF STATISTICAL MEASURES AND FREQUENCY FOR MWE IDENTIFICATION .....</b>	<b>831</b>
<i>Sandra Antunes, Amália Mendes</i>	
<b>ON THE RELIABILITY AND INTER-ANNOTATOR AGREEMENT OF HUMAN SEMANTIC MT EVALUATION VIA HMEANT .....</b>	<b>837</b>
<i>Chi-Kiu Lo, Dekai Wu</i>	
<b>DUAL SUBTITLES AS PARALLEL CORPORA .....</b>	<b>843</b>
<i>Shikun Zhang, Wang Ling, Chris Dyer</i>	
<b>REFRACTIVE: AN OPEN SOURCE TOOL TO EXTRACT KNOWLEDGE FROM SYNTACTIC AND SEMANTIC RELATIONS .....</b>	<b>849</b>
<i>Peter Exner, Pierre Nugues</i>	
<b>VARIATIONS ON QUANTITATIVE COMPARABILITY MEASURES AND THEIR EVALUATIONS ON SYNTHETIC FRENCH-ENGLISH COMPARABLE CORPORA .....</b>	<b>855</b>
<i>Guiyao Ke, Pierre-Francois Marteau, Gildas Menier</i>	
<b>USING A MACHINE LEARNING MODEL TO ASSESS THE COMPLEXITY OF STRESS SYSTEMS .....</b>	<b>862</b>
<i>Liviu Dinu, Alina Maria Ciobanu, Ioana Chitoran, Vlad Niculae</i>	
<b>PROSODIC, SYNTACTIC, SEMANTIC GUIDELINES FOR TOPIC STRUCTURES ACROSS DOMAINS AND CORPORA .....</b>	<b>868</b>
<i>Ana Isabel Mata, Helena Moniz, Telmo Móia, Anabela Gonçalves, Fátima Silva, Fernando Batista, Inês Duarte, Fátima Oliveira, Isabel Falé</i>	
<b>EVALUATING LEMMATIZATION MODELS FOR MACHINE-ASSISTED CORPUS-DICTIONARY LINKAGE .....</b>	<b>874</b>
<i>Kevin Black, Eric Ringger, Paul Felt, Kevin Seppi, Kristian Heal, Deryle Lonsdale</i>	
<b>FINITE-STATE MORPHOLOGICAL TRANSDUCERS FOR THREE KYPCHAK LANGUAGES .....</b>	<b>882</b>
<i>Jonathan Washington, Innar Salimzyanov, Francis Tyers</i>	



<b>AUTOMATIC CREATION OF WORDNETS FROM PARALLEL CORPORA</b> .....	890
<i>Antoni Oliver, Salvador Climent</i>	
<b>THE POLISH SUMMARIES CORPUS</b> .....	895
<i>Maciej Ogrodniczuk, Mateusz Kopec</i>	
<b>GLOBALPHONE: PRONUNCIATION DICTIONARIES IN 20 LANGUAGES</b> .....	899
<i>Tanja Schultz, Tim Schlippe</i>	
<b>PRE-ORDERING OF PHRASE-BASED MACHINE TRANSLATION INPUT IN TRANSLATION WORKFLOW</b> .....	904
<i>Alexandru Ceausu, Sabine Hunsicker</i>	
<b>A VECTOR SPACE MODEL FOR SYNTACTIC DISTANCES BETWEEN DIALECTS</b> .....	908
<i>Emanuele Di Buccio, Giorgio Maria Di Nunzio, Gianmaria Silvello</i>	
<b>A FINITE-STATE MORPHOLOGICAL ANALYZER FOR A LAKOTA HPSG GRAMMAR</b> .....	912
<i>Christian Curtis</i>	
<b>A WIKIPEDIA-BASED CORPUS FOR CONTEXTUALIZED MACHINE TRANSLATION</b> .....	916
<i>Jennifer Drexler, Pushpendre Rastogi, Jacqueline Aguilar, Benjamin Van Durme, Matt Post</i>	
<b>MAPPING WORDNET DOMAINS, WORDNET TOPICS AND WIKIPEDIA CATEGORIES TO GENERATE MULTILINGUAL DOMAIN SPECIFIC RESOURCES</b> .....	920
<i>Spandana Gella, Carlo Strapparava, Vivi Nastase</i>	
<b>ANNOTATING EVENTS IN AN EMOTION CORPUS</b> .....	925
<i>Sophia Lee, Shoushan Li, Chu-Ren Huang</i>	

Volume 2

<b>STATISTICAL ANALYSIS OF MULTILINGUAL TEXT CORPUS AND DEVELOPMENT OF LANGUAGE MODELS</b> .....	931
<i>Shyam Sundar Agrawal, Abhimanue, Shweta Bansal, Minakshi Mahajan</i>	
<b>NEW DIRECTIONS FOR LANGUAGE RESOURCE DEVELOPMENT AND DISTRIBUTION</b> .....	936
<i>Christopher Cieri, Denise Dipersio, Mark Liberman, Andrea Mazzucchi, Stephanie Strassel, Jonathan Wright</i>	
<b>REDISCOVERING 15 YEARS OF DISCOVERIES IN LANGUAGE RESOURCES AND EVALUATION: THE LREC ANTHOLOGY ANALYSIS</b> .....	N/A
<i>: Joseph Mariani, Patrick Paroubek, Gil Francopoulo, Olivier Hamon</i>	
<b>ANNOTATING QUESTION DECOMPOSITION ON COMPLEX MEDICAL QUESTIONS</b> .....	944
<i>Kirk Roberts, Kate Masterton, Marcelo Fiszman, Halil Kilicoglu, Dina Demner-Fushman</i>	
<b>BOOSTING OPEN INFORMATION EXTRACTION WITH NOUN-BASED RELATIONS</b> .....	949
<i>Clarissa Xavier, Vera Lima</i>	
<b>BOOTSTRAPPING OPEN-SOURCE ENGLISH-BULGARIAN COMPUTATIONAL DICTIONARY</b> .....	954
<i>Krasimir Angelov</i>	
<b>MOTÀMOT PROJECT: CONVERSION OF A FRENCH-KHMER PUBLISHED DICTIONARY FOR BUILDING A MULTILINGUAL LEXICAL SYSTEM</b> .....	960
<i>Mathieu Mangeot</i>	
<b>JUST.ASK, A QA SYSTEM THAT LEARNS TO ANSWER NEW QUESTIONS FROM PREVIOUS INTERACTIONS</b> .....	968
<i>Sérgio Curto, Ana C. Mendes, Pedro Curto, Luísa Coheur, Angela Costa</i>	
<b>DESIGN AND DEVELOPMENT OF AN ONLINE COMPUTATIONAL FRAMEWORK TO FACILITATE LANGUAGE COMPREHENSION RESEARCH ON INDIAN LANGUAGES</b> .....	973
<i>Manjira Sinha, Tirthankar Dasgupta, Anupam Basu</i>	
<b>THE NIJMEGEN CORPUS OF CASUAL CZECH</b> .....	981
<i>Mirjam Ernestus, Lucie Kocková-Amortová, Petr Pollak</i>	
<b>MODERN CHINESE HELPS ARCHAIC CHINESE PROCESSING: FINDING AND EXPLOITING THE SHARED PROPERTIES</b> .....	987
<i>Yan Song, Fei Xia</i>	
<b>DIGITAL LIBRARY 2.0: SOURCE OF KNOWLEDGE AND RESEARCH COLLABORATION PLATFORM</b> .....	995
<i>Włodzimierz Gruszczynski, Maciej Ogrodniczuk</i>	
<b>OPEN-DOMAIN INTERACTION AND ONLINE CONTENT IN THE SAMI LANGUAGE</b> .....	1000
<i>Kristiina Jokinen</i>	
<b>A CHARACTER-BASED APPROACH TO DISTRIBUTIONAL SEMANTIC MODELS: EXPLOITING KANJI CHARACTERS FOR CONSTRUCTING JAPANESEWORD VECTORS</b> .....	1006
<i>Akira Utsumi</i>	

<b>CROSS-LANGUAGE AUTHORSHIP ATTRIBUTION</b> .....	1013
<i>Dasha Bogdanova, Angeliki Lazaridou</i>	
<b>USING TRANSFER LEARNING TO ASSIST EXPLORATORY CORPUS ANNOTATION</b> .....	1019
<i>Paul Felt, Eric Ringger, Kevin Seppi, Kristian Heal</i>	
<b>ANCOR CENTRE, A LARGE FREE SPOKEN FRENCH COREFERENCE CORPUS: DESCRIPTION OF THE RESOURCE AND RELIABILITY MEASURES</b> .....	1025
<i>Judith Muzerelle, Anaïs Lefewvre, Emmanuel Schang, Jean-Yves Antoine, Aurore Pelletier, Denis Maurel, Iris Eshkol, Jeanne Villaneau</i>	
<b>GUAMPA: A TOOLKIT FOR COLLABORATIVE TRANSLATION</b> .....	1030
<i>Alex Rudnick, Taylor Skidmore, Alberto Samaniego, Michael Gasser</i>	
<b>EXPERIENCES WITH THE ISOCAT DATA CATEGORY REGISTRY</b> .....	1035
<i>Daan Broeder, Ineke Schuurman, Menzo Windhouwer</i>	
<b>RELISH LMF: UNLOCKING THE FULL POWER OF THE LEXICAL MARKUP FRAMEWORK</b> .....	1039
<i>Menzo Windhouwer, Justin Petro, Shakila Shayan</i>	
<b>BUILDING A CORPUS OF MANUALLY REVISED TEXTS FROM DISCOURSE PERSPECTIVE</b> .....	1045
<i>Ryu Iida, Takenobu Tokunaga</i>	
<b>THE CMD CLOUD</b> .....	1051
<i>Matej Durco, Menzo Windhouwer</i>	
<b>LINGUISTIC LANDSCAPING OF SOUTH ASIA USING DIGITAL LANGUAGE RESOURCES: GENETIC VS. AREAL LINGUISTICS</b> .....	1055
<i>Lars Borin, Anju Saxena, Taraka Rama, Bernard Comrie</i>	
<b>LINGUISTIC EVALUATION OF SUPPORT VERB CONSTRUCTIONS BY OPENLOGOS AND GOOGLE TRANSLATE</b> .....	1063
<i>Anabela Barreiro, Johanna Monti, Brigitte Orliac, Susanne Preuß, Kutz Arrieta, Wang Ling, Fernando Batista, Isabel Trancoso</i>	
<b>SINGLE-PERSON AND MULTI-PARTY 3D VISUALIZATIONS FOR NONVERBAL COMMUNICATION ANALYSIS</b> .....	1069
<i>Michael Kipp, Levin Freiherr Von Hollen, Michael Christopher Hrstka, Franziska Zamponi</i>	
<b>COLLECTION OF A SIMULTANEOUS TRANSLATION CORPUS FOR COMPARATIVE ANALYSIS</b> .....	1074
<i>Hiroaki Shimizu, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura</i>	
<b>THE AV-LASYN DATABASE : A SYNCHRONOUS CORPUS OF AUDIO AND 3D FACIAL MARKER DATA FOR AUDIO-VISUAL LAUGHTER SYNTHESIS</b> .....	1078
<i>Huseyin Cakmak, Jerome Urbain, Thierry Dutoit, Joelle Tilmanne</i>	
<b>INTEROPERABILITY OF DIALOGUE CORPORA THROUGH ISO 24617-2-BASED QUERYING</b> .....	1084
<i>Volha Petukhova, Andrei Malchanau, Harry Bunt</i>	
<b>ASR-BASED CALL SYSTEMS AND LEARNER SPEECH DATA: NEW RESOURCES AND OPPORTUNITIES FOR RESEARCH AND DEVELOPMENT IN SECOND LANGUAGE LEARNING</b> .....	1092
<i>Catia Cucchiari, Steve Bodnar, Bart Penning De Vries, Roeland Van Hout, Helmer Strik</i>	
<b>THE DBOX CORPUS COLLECTION OF SPOKEN HUMAN-HUMAN AND HUMAN-MACHINE DIALOGUES</b> .....	1099
<i>Volha Petukhova, Martin Gropp, Dietrich Klakow, Gregor Eigner, Mario Topf, Stefan Srb, Petr Motlicek, Blaise Potard, John Dines, Olivier Deroo, Ronny Egeler, Uwe Meinz, Steffen Liersch, Anna Schmidt</i>	
<b>THE DATABASE FOR SPOKEN GERMAN - DGD2</b> .....	1106
<i>Thomas Schmidt</i>	
<b>BUILDING A DATASET OF MULTILINGUAL COGNATES FOR THE ROMANIAN LEXICON</b> .....	1113
<i>Liviu Dinu, Alina Maria Ciobanu</i>	
<b>BENCHMARKING THE EXTRACTION AND DISAMBIGUATION OF NAMED ENTITIES ON THE SEMANTIC WEB</b> .....	1119
<i>Giuseppe Rizzo, Marieke Van Erp, Raphaël Troncy</i>	
<b>AUTOMATIC EXPANSION OF THE MRC PSYCHOLINGUISTIC DATABASE IMAGEABILITY RATINGS</b> .....	1127
<i>Ting Liu, Kit Cho, G. Aaron Broadwell, Samira Shaikh, Tomek Strzalkowski, John Lien, Sarah Taylor, Laurie Feldman, Boris Yamrom, Nick Webb, Umit Boz, Ignacio Cases, Ching-Sheng Lin</i>	
<b>TOWARDS BUILDING A KASHMIRI TREEBANK: SETTING UP THE ANNOTATION PIPELINE</b> .....	1133
<i>Riyaz Ahmad Bhat, Shahid Musjtaq Bhat, Dipti Misra Sharma</i>	
<b>SENTUBE: A CORPUS FOR SENTIMENT ANALYSIS ON YOUTUBE SOCIAL MEDIA</b> .....	1138
<i>Olga Uryupina, Barbara Plank, Aliaksei Severyn, Agata Rotondi, Alessandro Moschitti</i>	
<b>CIEMPIESS: A NEW OPEN-SOURCED MEXICAN SPANISH RADIO CORPUS</b> .....	1144
<i>Carlos Daniel Hernandez Mena, Abel Herrera Camacho</i>	

<b>SAVAS: COLLECTING, ANNOTATING AND SHARING AUDIOVISUAL LANGUAGE RESOURCES FOR AUTOMATIC SUBTITLING</b> .....	1149
<i>Arantza Del Pozo, Carlo Aliprandi, Aitor Alvarez, Carlos Mendes, Joao P. Neto, Sérgio Paulo, Nicola Piccinini, Matteo Raffaelli</i>	
<b>EXPLORING AND VISUALIZING VARIATION IN LANGUAGE RESOURCES</b> .....	1154
<i>Peter Fankhauser, Jörg Knappen, Elke Teich</i>	
<b>SIMPLE EFFECTIVE MICROBLOG NAMED ENTITY RECOGNITION: ARABIC AS AN EXAMPLE</b> .....	1158
<i>Kareem Darwish, Wei Gao</i>	
<b>PRIBERAM COMPRESSIVE SUMMARIZATION CORPUS: A NEW MULTI-DOCUMENT SUMMARIZATION CORPUS FOR EUROPEAN PORTUGUESE</b> .....	1163
<i>Miguel B. Almeida, Mariana S. C. Almeida, André F. T. Martins, Helena Figueira, Pedro Mendes, Cláudia Pinto</i>	
<b>HOPE AND FEAR: HOW OPINIONS INFLUENCE FACTUALITY</b> .....	1170
<i>Chantal Van Son, Marieke Van Erp, Antske Fokkens, Piek Vossen</i>	
<b>LINKING PICTOGRAPHS TO SYNSETS: SCLERA2CORNETTO</b> .....	1178
<i>Vincent Vandeghinste, Ineke Schuurman</i>	
<b>CHARACTERIZING AND PREDICTING BURSTY EVENTS: THE BUZZ CASE STUDY ON TWITTER</b> .....	1185
<i>Mohamed Morchid, Georges Linares, Richard Dufour</i>	
<b>INFORMATION EXTRACTION FROM GERMAN PATIENT RECORDS VIA HYBRID PARSING AND RELATION EXTRACTION STRATEGIES</b> .....	1191
<i>Hans-Ulrich Krieger, Christian Spurk, Hans Uszkoreit, Feiyu Xu, Yi Zhang, Frank Müller, Thomas Tolxdorff</i>	
<b>THE MMASCS MULTI-MODAL ANNOTATED SYNCHRONOUS CORPUS OF AUDIO, VIDEO, FACIAL MOTION AND TONGUE MOTION DATA OF NORMAL, FAST AND SLOW SPEECH</b> .....	1197
<i>Dietmar Schabus, Michael Pucher, Phil Hoole</i>	
<b>GENRES IN THE PRAGUE DISCOURSE TREEBANK</b> .....	1203
<i>Lucie Polakova, Pavlína Jinova, Jiri Mirovsky</i>	
<b>SPEECH RECOGNITION WEB SERVICES FOR DUTCH</b> .....	1210
<i>Joris Pelemans, Kris Demuyne, Hugo Van Hamme, Patrick Wambacq</i>	
<b>TRANSLATION ERRORS FROM ENGLISH TO PORTUGUESE: AN ANNOTATED CORPUS</b> .....	1214
<i>Angela Costa, Tiago Luis, Luisa Coheur</i>	
<b>AMAZIGH VERB CONJUGATOR</b> .....	1218
<i>Fadoua Ataa Allah, Siham Boulaknadel</i>	
<b>THE LIABILITY OF SERVICE PROVIDERS IN E-RESEARCH INFRASTRUCTURES: KILLING THE MESSENGER?</b> .....	1223
<i>Pawel Kamocki</i>	
<b>ADAPTING VERBNET TO FRENCH USING EXISTING RESOURCES</b> .....	1228
<i>Quentin Pradet, Laurence Danlos, Gaël De Chalendar</i>	
<b>ENGLISH-FRENCH VERB PHRASE ALIGNMENT IN EUROPARL FOR TENSE TRANSLATION MODELING</b> .....	1233
<i>Sharid Loaiciga, Thomas Meyer, Andrei Popescu-Belis</i>	
<b>THE EVOLVING INFRASTRUCTURE FOR LANGUAGE RESOURCES AND THE ROLE FOR DATA SCIENTISTS</b> .....	1241
<i>Nelleke Oostdijk, Henk Van Den Heuvel</i>	
<b>A NEW FORM OF HUMOR ? MAPPING CONSTRAINT-BASED COMPUTATIONAL MORPHOLOGIES TO A FINITE-STATE REPRESENTATION</b> .....	1246
<i>Attila Novák</i>	
<b>SLMOTION - AN EXTENSIBLE SIGN LANGUAGE ORIENTED VIDEO ANALYSIS TOOL</b> .....	1252
<i>Matti Karppa, Ville Viitaniemi, Marcos Luzardo, Jorma Laaksonen, Tommi Jantunen</i>	
<b>CONSTRUCTING A CHINESE/JAPANESE PARALLEL CORPUS FROM WIKIPEDIA</b> .....	1258
<i>Chenhui Chu, Toshiaki Nakazawa, Sadao Kurohashi</i>	
<b>CFT13: A RESOURCE FOR RESEARCH INTO THE POST-EDITING PROCESS</b> .....	1264
<i>Michael Carl, Mercedes Martínez García, Bartolomé Mesa-Lao</i>	
<b>MÖRKUM NJÁLU. AN ANNOTATED CORPUS TO ANALYSE AND EXPLAIN GRAMMATICAL DIVERGENCES BETWEEN 14TH-CENTURY MANUSCRIPTS OF NJÁL'S SAGA</b> .....	1272
<i>Ludger Zeevaert</i>	
<b>A CROWDSOURCING SMARTPHONE APPLICATION FOR SWISS GERMAN: PUTTING LANGUAGE DOCUMENTATION IN THE HANDS OF THE USERS</b> .....	1279
<i>Jean-Philippe Goldman, Adrian Leeman, Marie-José Kolly, Ingrid Hove, Ibrahim Almajai, Volker Dellwo, Steven Moran</i>	

<b>CLEARTK 2.0: DESIGN PATTERNS FOR MACHINE LEARNING IN UIMA .....</b>	<b>1283</b>
<i>Steven Bethard, Philip Ogren, Lee Becker</i>	
<b>A CONVENTIONAL ORTHOGRAPHY FOR TUNISIAN ARABIC .....</b>	<b>1288</b>
<i>Inès Zribi, Rahma Boujelbane, Abir Masmoudi, Mariem Ellouze, Lamia Belguith, Nizar Habash</i>	
<b>CREATING A MASSIVELY PARALLEL BIBLE CORPUS.....</b>	<b>1295</b>
<i>Thomas Mayer, Michael Cysouw</i>	
<b>CORPUS-BASED COMPUTATION OF REVERSE ASSOCIATIONS.....</b>	<b>1301</b>
<i>Reinhard Rapp</i>	
<b>LEXTEC - A RICH LANGUAGE RESOURCE FOR TECHNICAL DOMAINS IN PORTUGUESE.....</b>	<b>1308</b>
<i>Palmira Marrafa, Raquel Amaro, Sara Mendes</i>	
<b>BOOSTING THE CREATION OF A TREEBANK.....</b>	<b>1315</b>
<i>Blanca Arias, Nuria Bel, Mercè Lorente, Montserrat Marimón, Alba Milà, Jorge Vivaldi, Muntsa Padró, Marina Fomicheva, Imanol Larrea</i>	
<b>THE DUTCH LESLLA CORPUS .....</b>	<b>1322</b>
<i>Eric Sanders, Ineke Van De Craats, Vanja De Lint</i>	
<b>IMPROVEMENTS TO DEPENDENCY PARSING USING AUTOMATIC SIMPLIFICATION OF DATA .....</b>	<b>1326</b>
<i>Tomáš Jelínek</i>	
<b>THE META-KNOWLEDGE OF CAUSALITY IN BIOMEDICAL SCIENTIFIC DISCOURSE .....</b>	<b>1331</b>
<i>Claudiu Mihaila, Sophia Ananiadou</i>	
<b>DISCOSUITE - A PARSER TEST SUITE FOR GERMAN DISCONTINUOUS STRUCTURES .....</b>	<b>1339</b>
<i>Wolfgang Maier, Miriam Kaeshammer, Peter Baumann, Sandra Kübler</i>	
<b>MODELLING IRONY IN TWITTER: FEATURE ANALYSIS AND EVALUATION .....</b>	<b>1347</b>
<i>Francesco Barbieri, Horacio Saggion</i>	
<b>DBPEDIA DOMAINS: AUGMENTING DBPEDIA WITH DOMAIN INFORMATION .....</b>	<b>1354</b>
<i>Gregor Titze, Volha Bryl, Căcilia Zirn, Simone Paolo Ponzetto</i>	
<b>MINING A MULTIMODAL CORPUS FOR NON-VERBAL BEHAVIOR SEQUENCES CONVEYING ATTITUDES .....</b>	<b>1360</b>
<i>Mathieu Chollet, Magalie Ochs, Catherine Pelachaud</i>	
<b>BIOMEDICAL ENTITY EXTRACTION USING MACHINE-LEARNING BASED APPROACHES.....</b>	<b>1368</b>
<i>Cyril Grouin</i>	
<b>PARSING HETEROGENEOUS CORPORA WITH A RICH DEPENDENCY GRAMMAR .....</b>	<b>1374</b>
<i>Achim Stein</i>	
<b>ML-OPTIMIZATION OF PORTED CONSTRAINT GRAMMARS .....</b>	<b>1382</b>
<i>Eckhard Bick</i>	
<b>A MULTI-CULTURAL REPOSITORY OF AUTOMATICALLY DISCOVERED LINGUISTIC AND CONCEPTUAL METAPHORS .....</b>	<b>1387</b>
<i>Samira Shaikh, Tomek Strzalkowski, Ting Liu, George Aaron Broadwell, Boris Yamrom, Sarah Taylor, Laurie Feldman, Kit Cho, Umit Boz, Ignacio Cases, Yuliya Peshkova, Ching-Sheng Lin</i>	
<b>FIRST APPROACH TOWARD SEMANTIC ROLE LABELING FOR BASQUE.....</b>	<b>1393</b>
<i>Haritz Salaberri, Olatz Arregi, Beñat Zapirain</i>	
<b>GENERATING A LEXICON OF ERRORS IN PORTUGUESE TO SUPPORT AN ERROR IDENTIFICATION SYSTEM FOR SPANISH NATIVE LEARNERS.....</b>	<b>1400</b>
<i>Lianet Sepulveda Torres, Magali Sanches Duran, Sandra Aluísio</i>	
<b>XLID-LEXICA: CROSS-LINGUAL LINKED DATA LEXICA .....</b>	<b>1406</b>
<i>Lei Zhang, Michael Färber, Achim Rettinger</i>	
<b>A STUDY ON EXPERT SOURCING ENTERPRISE QUESTION COLLECTION AND CLASSIFICATION.....</b>	<b>1411</b>
<i>Yuan Luo, Thomas Boucher, Tolga Oral, David Osofsky, Sara Weber</i>	
<b>ANNOTATING RELATION MENTIONS IN TABLOID PRESS .....</b>	<b>1419</b>
<i>Hong Li, Sebastian Krause, Feiyu Xu, Hans Uszkoreit, Robert Hummel, Veselina Mironova</i>	
<b>EFFICIENT REUSE OF STRUCTURED AND UNSTRUCTURED RESOURCES FOR ONTOLOGY POPULATION .....</b>	<b>1424</b>
<i>Chetana Gavankar, Ashish Kulkarni, Ganesh Ramakrishnan</i>	
<b>MAPPING DIATOPIC AND DIACHRONIC VARIATION IN SPOKEN CZECH: THE ORTOFON AND DIALEKT CORPORA .....</b>	<b>1431</b>
<i>Marie Koprivová, Hana Golanova, Petra Klimesova, David Lukes</i>	
<b>CORPUS AND EVALUATION OF HANDWRITING RECOGNITION OF HISTORICAL GENEALOGICAL RECORDS.....</b>	<b>1438</b>
<i>Patrick Schone, Heath Nielson, Mark Ward</i>	
<b>REUSING SWEDISH FRAMENET FOR TRAINING SEMANTIC ROLES.....</b>	<b>1445</b>
<i>Ildikó Pilán, Elena Volodina</i>	

<b>A DATABASE OF FREELY WRITTEN TEXTS OF GERMAN SCHOOL STUDENTS FOR THE PURPOSE OF AUTOMATIC SPELLING ERROR CLASSIFICATION</b> .....	1450
<i>Kay Berkling, Johanna Fay, Masood Ghayoomi, Katrin Hein, Rémi Lavalley, Ludwig Linhuber, Sebastian Stüker</i>	
<b>PACE CORPUS: A MULTILINGUAL CORPUS OF POLARITY-ANNOTATED TEXTUAL DATA FROM THE DOMAINS AUTOMOTIVE AND CELLPHONE</b> .....	1456
<i>Christian Haenig, Andreas Niekler, Carsten Wuensch</i>	
<b>SZEGED CORPUS 2.5: MORPHOLOGICAL MODIFICATIONS IN A MANUALLY POSTAGGED HUNGARIAN CORPUS</b> .....	1462
<i>Veronika Vincze, Viktor Varga, Katalin Ilona Simkó, János Zsibrita, Ágoston Nagy, Richárd Farkas, János Csirik</i>	
<b>LINKED OPEN DATA AND WEB CORPUS DATA FOR NOUN COMPOUND BRACKETING</b> .....	1467
<i>Pierre André Ménard, Caroline Barriere</i>	
<b>MULTIMODAL CORPORA FOR SILENT SPEECH INTERACTION</b> .....	1475
<i>João Freitas, António Teixeira, Miguel Dias</i>	
<b>CONSTRUCTING A CORPUS OF JAPANESE PREDICATE PHRASES FOR SYNONYM/ANTONYM RELATIONS</b> .....	1480
<i>Tomoko Izumi, Tomohide Shibata, Hisako Asano, Yoshihiro Matsuo, Sadao Kurohashi</i>	
<b>THE IMPACT OF COHESION ERRORS IN EXTRACTION BASED SUMMARIES</b> .....	1487
<i>Evelina Rennes, Arne Jonsson</i>	
<b>THE CUHK DISCOURSE TREEBANK FOR CHINESE: ANNOTATING EXPLICIT DISCOURSE CONNECTIVES FOR THE CHINESE TREEBANK</b> .....	1495
<i>Lanjuan Zhou, Binyang Li, Zhongyu Wei, Kam-Fai Wong</i>	
<b>EXTRACTION OF DAILY CHANGING WORDS FOR QUESTION ANSWERING</b> .....	1503
<i>Kugatsu Sadamitsu, Ryuichiro Higashinaka, Yoshihiro Matsuo</i>	
<b>MTWATCH: A TOOL FOR THE ANALYSIS OF NOISY PARALLEL DATA</b> .....	1508
<i>Sandipan Dandapat, Declan Groves</i>	
<b>DISTRIBUTED DISTRIBUTIONAL SIMILARITIES OF GOOGLE BOOKS OVER THE CENTURIES</b> .....	1513
<i>Martin Riedl, Richard Steuer, Chris Biemann</i>	
<b>THE CLE URDU POS TAGSET</b> .....	1518
<i>Saba Urooj, Sarmad Hussain, Asad Mustafa, Rahila Parveen, Farah Adeeba, Tafseer Ahmed Khan, Miriam Butt, Annette Hautli</i>	
<b>NOSTA-D NAMED ENTITY ANNOTATION FOR GERMAN: GUIDELINES AND DATASET</b> .....	1524
<i>Darina Benikova, Chris Biemann, Marc Reznicek</i>	
<b>PHONE BOUNDARY ANNOTATION IN CONVERSATIONAL SPEECH</b> .....	1532
<i>Yi-Fen Liu, Shu-Chuan Tseng, J.-S Roger Jang</i>	
<b>A COLLOQUIAL CORPUS OF JAPANESE SIGN LANGUAGE: LINGUISTIC RESOURCES FOR OBSERVING SIGN LANGUAGE CONVERSATIONS</b> .....	1538
<i>Mayumi Bono, Kouhei Kikuchi, Paul Cibulka, Yutaka Osugi</i>	
<b>VALENTY: TOWARDS A COMPREHENSIVE VALENCE DICTIONARY OF POLISH</b> .....	1545
<i>Adam Przepiórkowski, Elzbieta Hajnicz, Agnieszka Patejuk, Marcin Wolinski, Filip Skwarski, Marek Swidzinski</i>	
<b>CAN THE CROWD BE CONTROLLED?: A CASE STUDY ON CROWD SOURCING AND AUTOMATIC VALIDATION OF COMPLETED TASKS BASED ON USER MODELING</b> .....	1553
<i>A. R. Balamurali</i>	
<b>COMPUTATIONAL NARRATOLOGY: EXTRACTING TENSE CLUSTERS FROM NARRATIVE TEXTS</b> .....	1560
<i>Thomas Bögel, Jannik Strötgen, Michael Gertz</i>	
<b>DESIGNING THE LATVIAN SPEECH RECOGNITION CORPUS</b> .....	1566
<i>Marcis Pinnis, Ilze Auzina, Karlis Goba</i>	
<b>ALIGNING PARALLEL TEXTS WITH INTERTEXT</b> .....	1573
<i>Pavel Vondricka</i>	
<b>CORPUS FOR COREFERENCE RESOLUTION ON SCIENTIFIC PAPERS</b> .....	1578
<i>Panot Chaimongkol, Akiko Atzawa, Yuka Tateisi</i>	
<b>FROM NON WORD TO NEW WORD: AUTOMATICALLY IDENTIFYING NEOLOGISMS IN FRENCH NEWSPAPERS</b> .....	1582
<i>Ingrid Falk, Delphine Bernhard, Christophe Gérard</i>	
<b>EVALUATING THE EFFECTS OF INTERACTIVITY IN A POST-EDITING WORKBENCH</b> .....	1590
<i>Nancy Underwood, Bartolome Mesa-Lao, Mercedes García Martínez, Michael Carl, Vicent Alabau, Jesús González-Rubio, Luis A. Leiva, Germán Sanchis-Trilles, Daniel Ortíz-Martínez, Francisco Casacuberta</i>	
<b>USING TWITTER AND SENTIMENT ANALYSIS FOR EVENT DETECTION</b> .....	N/A
<i>Georgios Paltoglou</i>	
<b>THE RESEARCH AND TEACHING CORPUS OF SPOKEN GERMAN - FOLK</b> .....	1597
<i>Thomas Schmidt</i>	

<b>DATA MINING WITH SHALLOW VS. LINGUISTIC FEATURES TO STUDY DIVERSIFICATION OF SCIENTIFIC REGISTERS.....</b>	1602
<i>Stefania Degaetano-Ortlieb, Peter Fankhauser, Hannah Kermes, Ekaterina Lapshinova-Koltunski, Noam Ordan, Elke Teich</i>	
<b>ON STOPWORDS, FILTERING AND DATA SPARSITY FOR SENTIMENT ANALYSIS OF TWITTER.....</b>	1610
<i>Hassan Saif, Miriam Fernandez, Yulan He, Harith Alani</i>	
<b>ADAPTING FREELY AVAILABLE RESOURCES TO BUILD AN OPINION MINING PIPELINE IN PORTUGUESE .....</b>	1618
<i>Patrik Lambert, Carlos Rodriguez-Penagos</i>	
<b>THE SYN-SERIES CORPORA OF WRITTEN CZECH.....</b>	1622
<i>Milena Hnatkova, Michal Kren, Pavel Prochazka, Hana Skoumalova</i>	
<b>PARCOR 1.0: A PARALLEL PRONOUN-COREFERENCE CORPUS TO SUPPORT STATISTICAL MT.....</b>	1627
<i>Liane Guillou, Christian Hardmeier, Aaron Smith, Jörg Tiedemann, Bonnie Webber</i>	
<b>A BENCHMARK DATABASE OF PHONETIC ALIGNMENTS IN HISTORICAL LINGUISTICS AND DIALECTOLOGY.....</b>	1635
<i>Johann-Mattis List, Jelena Prokic</i>	
<b>EXTRACTING NEWS WEB PAGE CREATION TIME WITH DCTFINDER.....</b>	1642
<i>Xavier Tannier</i>	
<b>CORPUS OF 19<sup>TH</sup>-CENTURY CZECH TEXTS: PROBLEMS AND SOLUTIONS.....</b>	1648
<i>Karel Kucera, Martin Stluka</i>	
<b>INVESTIGATING THE IMAGE OF ENTITIES IN SOCIAL MEDIA: DATASET DESIGN AND FIRST RESULTS .....</b>	1652
<i>Julien Velcin, Young-Min Kim, Caroline Brun, Jean-Yves Dormagen, Eric Sanjuan, Leila Khouas, Anne Peradotto, Stéphane Bonnevey, Claude Roux, Julien Boyadjian, Alejandro Molina, Marie Neihouser</i>	
<b>THE NORWEGIAN DEPENDENCY TREEBANK.....</b>	1657
<i>Per Erik Solberg, Arne Skjaeholt, Lilja Ovreliid, Kristin Hagen, Janne Bondi Johannessen</i>	
<b>EXTENDING STANDOFF ANNOTATION .....</b>	1664
<i>Maik Stuhrenberg</i>	
<b>AN EFFICIENT LANGUAGE INDEPENDENT TOOLKIT FOR COMPLETE MORPHOLOGICAL DISAMBIGUATION.....</b>	1670
<i>Laszlo Laki, Gyorgy Orosz</i>	
<b>A DECADE OF HLT AGENCY ACTIVITIES IN THE LOW COUNTRIES: FROM RESOURCE MAINTENANCE (BLARK) TO SERVICE OFFERINGS (BLAISE).....</b>	1676
<i>Peter Spyns, Remco Van Veenendaal</i>	
<b>A CORPUS OF SPONTANEOUS SPEECH IN LECTURES: THE KIT LECTURE CORPUS FOR SPOKEN LANGUAGE PROCESSING AND TRANSLATION.....</b>	1684
<i>Eunah Cho, Sarah Fünfer, Sebastian Stüker, Alex Waibel</i>	
<b>FREE ACOUSTIC AND LANGUAGE MODELS FOR LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION IN SWEDISH .....</b>	1690
<i>Niklas Vanhainen, Giampiero Salvi</i>	
<b>TURKISH RESOURCES FOR VISUAL WORD RECOGNITION.....</b>	1695
<i>Begun Erten, Cem Bozsahin, Deniz Zeyrek</i>	
<b>AN ARABIC TWITTER CORPUS FOR SUBJECTIVITY AND SENTIMENT ANALYSIS.....</b>	1700
<i>Eshrag Refaee, Verena Rieser</i>	
<b>THE IMAGACT VISUAL ONTOLOGY. AN EXTENDABLE MULTILINGUAL INFRASTRUCTURE FOR THE REPRESENTATION OF LEXICAL ENCODING OF ACTION .....</b>	1706
<i>Massimo Moneglia, Susan Brown, Francesca Frontini, Gloria Gagliardi, Fahad Khan, Monica Monachini, Alessandro Panunzi</i>	
<b>COLLABORATION IN THE PRODUCTION OF A MASSIVELY MULTILINGUAL LEXICON.....</b>	1714
<i>Martin Benjamin</i>	
<b>AN EFFORTLESS WAY TO CREATE LARGE-SCALE DATASETS FOR FAMOUS SPEAKERS .....</b>	1719
<i>Francois Salmon, Félicien Vallet</i>	
<b>BRIDGING THE GAP BETWEEN SPEECH TECHNOLOGY AND NATURAL LANGUAGE PROCESSING: AN EVALUATION TOOLBOX FOR TERM DISCOVERY SYSTEMS .....</b>	1724
<i>Bogdan Ludusan, Maarten Versteegh, Aren Jansen, Guillaume Gravier, Xuan-Nga Cao, Mark Johnson, Emmanuel Dupoux</i>	
<b>MODELING AND EVALUATING DIALOG SUCCESS IN THE LAST MINUTE CORPUS.....</b>	1732
<i>Dietmar Rösner, Rafael Friesen, Stephan Günther, Rico Andrich</i>	
<b>COMPARISON OF GENDER- AND SPEAKER-ADAPTIVE EMOTION RECOGNITION .....</b>	1739
<i>Maxim Sidorov, Stefan Ultes, Alexander Schmitt</i>	

<b>THE PRAGMATIC ANNOTATION OF A CORPUS OF ACADEMIC LECTURES</b> .....	1744
<i>Sian Alsop, Hilary Nesi</i>	
<b>USING TEL, CMDI AND ISOCAT IN CLARIN-DK</b> .....	1748
<i>Dorte Haltrup Hansen, Lene Offersgaard, Sussi Olsen</i>	
<b>COMPARATIVE ANALYSIS OF VERBAL ALIGNMENT IN HUMAN-HUMAN AND HUMAN-AGENT INTERACTIONS</b> .....	1754
<i>Sabrina Campano, Jessica Durand, Chloé Clavel</i>	
<b>TRANSLITERATION AND ALIGNMENT OF PARALLEL TEXTS FROM CYRILLIC TO LATIN</b> .....	1762
<i>Petic Mircea, Daniela Gifu</i>	
<b>HOW TO TELL A SCHNEEMANN FROM A MILCHMANN: AN ANNOTATION SCHEME FOR COMPOUND-INTERNAL RELATIONS</b> .....	1767
<i>Corina Dima, Verena Henrich, Erhard Hinrichs, Christina Hoppermann</i>	
<b>CAN NUMERICAL EXPRESSIONS BE SIMPLER? IMPLEMENTATION AND DEMONSTRATION OF A NUMERICAL SIMPLIFICATION SYSTEM FOR SPANISH</b> .....	1775
<i>Susana Bautista, Horacio Saggion</i>	
<b>4FX: LIGHT VERB CONSTRUCTIONS IN A MULTILINGUAL PARALLEL CORPUS</b> .....	1782
<i>Anita Rácz, T. István Nagy, Veronika Vincze</i>	
<b>THE EIDENTITY TEXT EXPLORATION WORKBENCH</b> .....	1788
<i>Fritz Kliche, Andre Blessing, Ulrich Heid, Jonathan Sonntag</i>	
<b>EMILYA: EMOTIONAL BODY EXPRESSION IN DAILY ACTIONS DATABASE</b> .....	1795
<i>Nesrine Fourati, Catherine Pelachaud</i>	
<b>USING STEM-TEMPLATES TO IMPROVE ARABIC POS AND GENDER/NUMBER TAGGING</b> .....	1803
<i>Kareem Darwish, Ahmed Abdelali, Hamdy Mubarak</i>	
<b>CONSTRUCTION OF DIACHRONIC ONTOLOGIES FROM PEOPLE'S DAILY OF FIFTY YEARS</b> .....	1809
<i>Shaoda He, Xiaojun Zou, Liumingjing Xiao, Junfeng Hu</i>	
<b>ROOTS: A TOOLKIT FOR EASY, FAST AND CONSISTENT PROCESSING OF LARGE SEQUENTIAL ANNOTATED DATA COLLECTIONS</b> .....	1815
<i>Jonathan Chevelu, Gwénoél Lecorvé, Damien Lolive</i>	
<b>COMPUTER-AIDED QUALITY ASSURANCE OF AN ICELANDIC PRONUNCIATION DICTIONARY</b> .....	1823
<i>Martin Jansche</i>	
<b>DISAMBIGUATING VERBS BY COLLOCATION: CORPUS LEXICOGRAPHY MEETS NATURAL LANGUAGE PROCESSING</b> .....	1827
<i>Ismail El Maarouf, Jane Bradbury, Vít Baisa, Patrick Hanks</i>	
<b>AUTOMATIC ERROR DETECTION CONCERNING THE DEFINITE AND INDEFINITE CONJUGATION IN THE HUNLEARNER CORPUS</b> .....	1833
<i>Veronika Vincze, János Zsibrita, Péter Durst, Martina Katalin Szabó</i>	
<b>SPEECH-BASED EMOTION RECOGNITION: FEATURE SELECTION BY SELF-ADAPTIVE MULTI-CRITERIA GENETIC ALGORITHM</b> .....	1838
<i>Maxim Sidorov, Christina Brester, Wolfgang Minker, Eugene Semenkina</i>	
<b>CONSTRUCTING AND EXPLOITING AN AUTOMATICALLY ANNOTATED RESOURCE OF LEGISLATIVE TEXTS</b> .....	1843
<i>Stefan Höfler, Kyoko Sugisaki</i>	

Volume 3

<b>GENITIVDB - A CORPUS-GENERATED DATABASE FOR GERMAN GENITIVE CLASSIFICATION</b> .....	1849
<i>Roman Schneider</i>	
<b>RESOURCES IN CONFLICT: A BILINGUAL VALENCY LEXICON VS. A BILINGUAL TREEBANK VS. A LINGUISTIC THEORY</b> .....	1856
<i>Jana Sindlerova, Zdenka Uresova, Eva Fucikova</i>	
<b>THE NEWSOME CORPUS: A UNIFYING OPINION ANNOTATION FRAMEWORK ACROSS GENRES AND IN MULTIPLE LANGUAGES</b> .....	1861
<i>Roser Saurí, Judith Domingo, Toni Badia</i>	
<b>BUY ONE GET ONE FREE: DISTANT ANNOTATION OF CHINESE TENSE, EVENT TYPE AND MODALITY</b> .....	1869
<i>Nianwen Xue, Yuchen Zhang</i>	
<b>MULTIMODAL DIALOGUE SEGMENTATION WITH GESTURE POST-PROCESSING</b> .....	1874
<i>Kodai Takahashi, Masashi Inoue</i>	

<b>THE DANGEROUS MYTH OF THE STAR SYSTEM.....</b>	1879
<i>André Bittar, Dini Luca, Sigrid Maurel, Mathieu Ruhlmann</i>	
<b>COMPARISON OF THE IMPACT OF WORD SEGMENTATION ON NAME TAGGING FOR CHINESE AND JAPANESE .....</b>	1884
<i>Haibo Li, Masato Hagiwara, Qi Li, Heng Ji</i>	
<b>COROLA - THE REFERENCE CORPUS OF CONTEMPORARY ROMANIAN LANGUAGE.....</b>	1889
<i>Verginica Barbu Mititelu, Elena Irimia, Dan Tuft</i>	
<b>BUILDING A REFERENCE LEXICON FOR COUNTABILITY IN ENGLISH.....</b>	1894
<i>Tibor Kiss, Francis Jeffry Pelletier, Tobias Stadfeld</i>	
<b>THE LIMA MULTILINGUAL ANALYZER MADE FREE: FLOSS RESOURCES ADAPTATION AND CORRECTION .....</b>	1900
<i>Gael De Chalendar</i>	
<b>A SICK CURE FOR THE EVALUATION OF COMPOSITIONAL DISTRIBUTIONAL SEMANTIC MODELS .....</b>	1906
<i>Marco Marelli, Stefano Menini, Marco Baroni, Luisa Bentivogli, Raffaella Bernardi, Roberto Zamparelli</i>	
<b>TWENTE DEBATE CORPUS ? A MULTIMODAL CORPUS FOR HEAD MOVEMENT ANALYSIS .....</b>	1914
<i>Bayu Rahayudi, Ronald Poppe, Dirk Heylen</i>	
<b>THE EASR CORPORA OF EUROPEAN PORTUGUESE, FRENCH, HUNGARIAN AND POLISH ELDERLY SPEECH .....</b>	1919
<i>Annika Hämäläinen, Jairo Avelar, Silvia Rodrigues, Miguel Sales Dias, Artur Kolesinski, Tibor Fegyó, Géza Németh, Petra Csobánka, Karine Lan, David Hewson</i>	
<b>SHARING CULTURAL HERITAGE: THE CLAVIUS ON THE WEB PROJECT .....</b>	1926
<i>Matteo Abrate, Angelo Mario Del Grosso, Emiliano Giovannetti, Angelica Lo Duca, Damiana Luzzi, Lorenzo Mancini, Andrea Marchetti, Irene Pedretti, Silvia Piccini</i>	
<b>3D FACE TRACKING AND MULTI-SCALE, SPATIO-TEMPORAL ANALYSIS OF LINGUISTICALLY SIGNIFICANT FACIAL EXPRESSIONS AND HEAD POSITIONS IN ASL .....</b>	1934
<i>Bo Liu, Jingjing Liu, Xiang Yu, Dimitris Metaxas, Carol Neidle</i>	
<b>EXPLORING FACTORS THAT CONTRIBUTE TO SUCCESSFUL FINGERSPELLING COMPREHENSION .....</b>	1941
<i>Leah Geer, Jonathan Keane</i>	
<b>THE DARE CORPUS: A RESOURCE FOR ANAPHORA RESOLUTION IN DIALOGUE BASED INTELLIGENT TUTORING SYSTEMS .....</b>	1947
<i>Nobal Niraula, Vasile Rus, Rajendra Banjade, Dan Stefanescu, William Baggett, Brent Morgan</i>	
<b>ON THE ANNOTATION OF TMX TRANSLATION MEMORIES FOR ADVANCED LEVERAGING IN COMPUTER-AIDED TRANSLATION.....</b>	1952
<i>Mikel Forcada</i>	
<b>THE D-ANS CORPUS: THE DUBLIN-AUTONOMOUS NERVOUS SYSTEM CORPUS OF BIOSIGNAL AND MULTIMODAL RECORDINGS OF CONVERSATIONAL SPEECH.....</b>	1957
<i>Shannon Hennig, Ryad Chellali, Nick Campbell</i>	
<b>ANNOTATING THE MASC CORPUS WITH BABELNET.....</b>	1963
<i>Andrea Moro, Roberto Navigli, Francesco Maria Tucci, Rebecca J. Passonneau</i>	
<b>ALL FRAGMENTS COUNT IN PARSER EVALUATION .....</b>	1969
<i>Joost Bastings, Khalil Sima'An</i>	
<b>TEXAFON 2.0: A TEXT PROCESSING TOOL FOR THE GENERATION OF EXPRESSIVE SPEECH IN TTS APPLICATIONS .....</b>	1974
<i>Juan-María Garrido, Yesika Laplaza, Benjamin Kolz, Miquel Cornudella</i>	
<b>A PERSIAN TREEBANK WITH STANFORD TYPED DEPENDENCIES .....</b>	1981
<i>Mojgan Seraji, Carina Jahani, Beáta Megyesi, Joakim Nivre</i>	
<b>A LANGUAGE-INDEPENDENT APPROACH TO EXTRACTING DERIVATIONAL RELATIONS FROM AN INFLECTIONAL LEXICON .....</b>	1987
<i>Marion Baranes, Benoît Sagot</i>	
<b>NAMED ENTITY RECOGNITION ON TURKISH TWEETS .....</b>	1994
<i>Dilek Kucuk, Guillaume Jacquet, Ralf Steinberger</i>	
<b>RHAPSODIE: A PROSODIC-SYNTACTIC TREEBANK FOR SPOKEN FRENCH.....</b>	1999
<i>Anne Lacheret, Sylvain Kahane, Julie Beliao, Anne Dister, Kim Gerdes, Jean-Philippe Goldman, Nicolas Obin, Paola Pietrandrea, Atanas Tchobanov</i>	
<b>THE IULA SPANISH LSP TREEBANK .....</b>	2006
<i>Montserrat Marimon, Núria Bel, Beatriz Fisas, Blanca Arias, Silvia Vázquez, Jorge Vivaldi, Carlos Morell, Mercè Lorente</i>	
<b>MORPHO-SYNTACTIC STUDY OF ERRORS FROM SPEECH RECOGNITION SYSTEM.....</b>	2013
<i>Maria Goryainova, Cyril Grouin, Sophie Rosset, Ioana Vasilescu</i>	



<b>NOT AN INTERLINGUA, BUT CLOSE: COMPARISON OF ENGLISH AMRS TO CHINESE AND CZECH</b> .....	2018
<i>Nianwen Xue, Ondrej Bojar, Jan Hajic, Martha Palmer, Zdenka Uresova, Xiuhong Zhang</i>	
<b>ANNOTATING RELATING IN SCIENTIFIC ARTICLES</b> .....	2026
<i>Adam Meyers, Giancarlo Lee, Angus Grieve-Smith, Yifan He, Harriet Taber</i>	
<b>ANNOTATING CLINICAL EVENTS IN TEXT SNIPPETS FOR PHENOTYPE DETECTION</b> .....	2034
<i>Prescott Klassen, Fei Xia, Lucy Vanderwende, Meliha Yetisgen</i>	
<b>PHONEME SIMILARITY MATRICES TO IMPROVE LONG AUDIO ALIGNMENT FOR AUTOMATIC SUBTITLING</b> .....	2039
<i>Pablo Ruiz, Aitor Álvarez, Haritz Arzelus</i>	
<b>USE OF UNSUPERVISED WORD CLASSES FOR ENTITY RECOGNITION: APPLICATION TO THE DETECTION OF DISORDERS IN CLINICAL REPORTS</b> .....	2045
<i>Maria Evangelia Chatzimina, Cyril Grouin, Pierre Zweigenbaum</i>	
<b>THREE DIMENSIONS OF THE SO-CALLED "INTEROPERABILITY" OF ANNOTATION SCHEMES</b> .....	2053
<i>Eva Hajicova, Maria Evangelia Chatzimina, Cyril Grouin, Pierre Zweigenbaum</i>	
<b>ON COMPLEX WORD ALIGNMENT CONFIGURATIONS</b> .....	2059
<i>Miriam Kaeshammer, Anika Westburg</i>	
<b>HFST-SWENER - A NEW NER RESOURCE FOR SWEDISH</b> .....	2067
<i>Dimitrios Kokkinakis, Jyrki Niemi, Sam Hardwick, Krister Lindén, Lars Borin</i>	
<b>BUILDING AND MODELLING MULTILINGUAL SUBJECTIVE CORPORA</b> .....	2074
<i>Motaz Saad, David Langlois, Kamel Smaili</i>	
<b>GRASS: THE GRAZ CORPUS OF READ AND SPONTANEOUS SPEECH</b> .....	2080
<i>Barbara Schuppler, Martin Hagmueller, Juan A. Morales-Cordovilla, Hannes Pessentheiner</i>	
<b>LINGUISTIC RESOURCES AND CATS: HOW TO USE ISOCAT, RELCAT AND SCHEMACAT</b> .....	2086
<i>Menzo Windhouwer, Ineke Schuurman</i>	
<b>BRING VS. MTROGET: EVALUATING AUTOMATIC THESAURUS TRANSLATION</b> .....	2091
<i>Lars Borin, Jens Allwood, Gerard De Melo</i>	
<b>INTRODUCING A FRAMEWORK FOR THE EVALUATION OF MUSIC DETECTION TOOLS</b> .....	2098
<i>Paula Lopez-Otero, Laura Docio-Fernandez, Carmen Garcia-Mateo</i>	
<b>WORDNET?WIKIPEDIA?WIKTIONARY: CONSTRUCTION OF A THREE-WAY ALIGNMENT</b> .....	2103
<i>Tristan Miller, Iryna Gurevych</i>	
<b>CROSS-LINGUISTIC ANNOTATION OF NARRATIVITY FOR ENGLISH/FRENCH VERB TENSE DISAMBIGUATION</b> .....	2110
<i>Cristina Grisot, Thomas Meyer</i>	
<b>THE TARAXÛ CORPUS OF HUMAN-ANNOTATED MACHINE TRANSLATIONS</b> .....	2114
<i>Eleftherios Avramidis, Aljoscha Burchardt, Sabine Hunsicker, Maja Popovic, Cindy Tscherwinka, David Vilar, Hans Uszkoreit</i>	
<b>DETECTING DOCUMENT STRUCTURE IN A VERY LARGE CORPUS OF UK FINANCIAL REPORTS</b> .....	2118
<i>Mahmoud El-Haj, Paul Rayson, Steve Young, Martin Walker</i>	
<b>LATENT SEMANTIC ANALYSIS MODELS ON WIKIPEDIA AND TASA</b> .....	2122
<i>Dan Stefanescu, Rajendra Banjade, Vasile Rus</i>	
<b>THE STRATEGIC IMPACT OF META-NET ON THE REGIONAL, NATIONAL AND INTERNATIONAL LEVEL</b> .....	2128
<i>Georg Rehm, Hans Uszkoreit, Sophia Ananiadou, Núria Bel, Audroné Bieleviciéné, Lars Borin, António Branco, Gerhard Budin, Nicoletta Calzolari, Walter Daelemans, Radovan Garabik, Marko Grobelnik, Carmen Garcia-Mateo, Josef Van Genabith, Jan Hajic, Inma Hernaez, John Judge, Svetla Koeva, Simon Krek, Cvetana Krstev, Krister Linden, Bernardo Magnini, Joseph Mariani, John McNaught, Maite Melero, Monica Monachini, Asuncion Moreno, Jan Odijk, Maciej Ogrodniczuk, Piotr Pezik, Stelios Piperidis, Adam Przepiórkowski, Eirikur Rögnvaldsson, Michael Rosner, Bolette Pedersen, Inguna Skadina, Koenraad De Smedt, Marko Tadic, Paul Thompson, Dan Tufi, Tamás Váradi, Andrejs Vasiljevs, Kadri Vider, Jolanta Zabarskaite</i>	
<b>THE WELTMODELL: A DATA-DRIVEN COMMONSENSE KNOWLEDGE BASE</b> .....	2136
<i>Alan Akbik, Thilo Michael</i>	
<b>GERMAN ALCOHOL LANGUAGE CORPUS - THE QUESTION OF DIALECT</b> .....	2141
<i>Florian Schiel, Thomas Kisler</i>	
<b>CLARA: A NEW GENERATION OF RESEARCHERS IN COMMON LANGUAGE RESOURCES AND THEIR APPLICATIONS</b> .....	2145
<i>Koenraad De Smedt, Erhard Hinrichs, Detmar Meurers, Inguna Skadina, Bolette Pedersen, Costanza Navarretta, Núria Bel, Krister Linden, Marketa Lopatkova, Jan Hajic, Gisle Andersen, Przemyslaw Lenkiewicz</i>	

<b>A LARGE CORPUS OF PRODUCT REVIEWS IN PORTUGUESE: TACKLING OUT-OF-VOCABULARY WORDS</b> .....	2154
<i>Nathan Hartmann, Lucas Avanco, Pedro Balage, Magali Duran, Maria Das Gracias Volpe Nunes, Thiago Pardo, Sandra Aluísio</i>	
<b>SHATA-ANUVADAK: TACKLING MULTIWAY TRANSLATION OF INDIAN LANGUAGES</b> .....	2161
<i>Anoop Kunchukuttan, Abhijit Mishra, Rajen Chatterjee, Ritesh Shah, Pushpak Bhattacharyya</i>	
<b>THE CLARIN RESEARCH INFRASTRUCTURE: RESOURCES AND TOOLS FOR EHUMANITIES SCHOLARS</b> .....	2168
<i>Erhard Hinrichs, Steven Krauwer</i>	
<b>SPRINTER: LANGUAGE TECHNOLOGIES FOR INTERACTIVE AND MULTIMEDIA LANGUAGE LEARNING</b> .....	2175
<i>Renlong Ai, Marcela Charfuelan, Walter Kasper, Tina Kluwer, Hans Uszkoreit, Feiyu Xu, Sandra Gasber, Philip Gienandt</i>	
<b>BILINGUAL DICTIONARY INDUCTION AS AN OPTIMIZATION PROBLEM</b> .....	2181
<i>Wushouer Mairidan, Toru Ishida, Donghui Lin, Katsutoshi Hirayama</i>	
<b>TWO APPROACHES TO METAPHOR DETECTION</b> .....	2189
<i>Brian Macwhinney, Davida Fromm</i>	
<b>A JAPANESE WORD DEPENDENCY CORPUS</b> .....	2195
<i>Shinsuke Mori, Hideki Ogura, Tetsuro Sasada</i>	
<b>CROWDSOURCING AND ANNOTATING NER FOR TWITTER #DRIFT</b> .....	2201
<i>Hege Fromreide, Dirk Hovy, Anders Sogaard</i>	
<b>LANGUAGE RESOURCES AND ANNOTATION TOOLS FOR CROSS-SENTENCE RELATION EXTRACTION</b> .....	2205
<i>Sebastian Krause, Hong Li, Feiyu Xu, Hans Uszkoreit, Robert Hummel, Luise Spielhagen</i>	
<b>EVALUATING CORPORA DOCUMENTATION WITH REGARDS TO THE ETHICS AND BIG DATA CHARTER</b> .....	2211
<i>Alain Couillaud, Karèn Fort, Gilles Adda, Hugues Mazancourt</i>	
<b>BOOTSTRAPPING TERM EXTRACTORS FOR MULTIPLE LANGUAGES</b> .....	2216
<i>Ahmet Aker, Monica Paramita, Emma Barker, Robert Gaizauskas</i>	
<b>EVALUATION OF AUTOMATIC HYPERNYM EXTRACTION FROM TECHNICAL CORPORA IN ENGLISH AND DUTCH</b> .....	2223
<i>Els Lefever, Marjan Van De Kauter, Veronique Hoste</i>	
<b>MEASURING READABILITY OF POLISH TEXTS: BASELINE EXPERIMENTS</b> .....	2231
<i>Bartosz Broda, Bartłomiej Niton, Włodzimierz Gruszczyński, Maciej Ogródniczuk</i>	
<b>INTRODUCING A WEB APPLICATION FOR LABELING, VISUALIZING SPEECH AND CORRECTING DERIVED SPEECH SIGNALS</b> .....	2239
<i>Raphael Winkelmann, Georg Raess</i>	
<b>AN ITERATIVE APPROACH FOR MINING PARALLEL SENTENCES IN A COMPARABLE CORPUS</b> .....	2244
<i>Lise Rebut, Phillippe Langlais</i>	
<b>DEVELOPMENT OF A TV BROADCASTS SPEECH RECOGNITION SYSTEM FOR QATARI ARABIC</b> .....	2252
<i>Mohamed Elmahdy, Mark Hasegawa-Johnson, Eiman Mustafawi</i>	
<b>CAN CROWDSOURCING BE USED FOR EFFECTIVE ANNOTATION OF ARABIC?</b> .....	2257
<i>Wajdi Zaghouni, Kais Dukes</i>	
<b>DESIGN AND DEVELOPMENT OF AN RDB VERSION OF THE CORPUS OF SPONTANEOUS JAPANESE</b> .....	2262
<i>Hanae Koiso, Yasuharu Den, Ken'ya Nishikawa, Kikuo Maekawa</i>	
<b>AUTOMATIC LONG AUDIO ALIGNMENT AND CONFIDENCE SCORING FOR CONVERSATIONAL ARABIC SPEECH</b> .....	2268
<i>Mohamed Elmahdy, Mark Hasegawa-Johnson, Eiman Mustafawi</i>	
<b>VOCABULARY-BASED LANGUAGE SIMILARITY USING WEB CORPORA</b> .....	2273
<i>Dirk Goldhahn, Uwe Quasthoff</i>	
<b>NEWSREADER: RECORDING HISTORY FROM DAILY NEWS STREAMS</b> .....	2279
<i>Piek Vossen, German Rigau, Luciano Serafini, Pim Stouten, Francis Irving, Willem Van Hage</i>	
<b>A SET OF OPEN SOURCE TOOLS FOR TURKISH NATURAL LANGUAGE PROCESSING</b> .....	2287
<i>Çagri Coltekin</i>	
<b>THE GULF OF GUINEA CREOLE CORPORA</b> .....	2295
<i>Tjerk Hagemeijer, Michel Génèreux, Iris Hendrickx, Amália Mendes, Abigail Tiny, Armando Zamora</i>	
<b>S-POT - A BENCHMARK IN SPOTTING SIGNS WITHIN CONTINUOUS SIGNING</b> .....	2302
<i>Ville Viitaniemi, Tommi Jantunen, Leena Savolainen, Matti Karppa, Jorma Laaksonen</i>	

<b>CONVERTING AN HPSG-BASED TREEBANK INTO ITS PARALLEL DEPENDENCY-BASED TREEBANK</b> .....	2308
<i>Masood Ghayoomi, Jonas Kuhn</i>	
<b>TWEETNORM_ES: AN ANNOTATED CORPUS FOR SPANISH MICROTEXT NORMALIZATION</b> .....	2316
<i>Iñaki Alegria, Nora Aranberri, Pere Comas, Victor Fresno, Pablo Gamallo, Lluís Padró, Inaki San Vicente, Jordi Turmo, Arkaitz Zubiaga</i>	
<b>THE PROCEDURE OF LEXICO-SEMANTIC ANNOTATION OF SKLADNICA TREEBANK</b> .....	2321
<i>Elzbieta Hajnicz</i>	
<b>MEDIA MONITORING AND INFORMATION EXTRACTION FOR THE HIGHLY INFLECTED AGGLUTINATIVE LANGUAGE HUNGARIAN</b> .....	2329
<i>Júlia Pajzs, Ralf Steinberger, Maud Ehrmann, Mohamed Ebrahim, Leonida Della Rocca, Stefano Bucci, Eszter Simon, Tamás Váradi</i>	
<b>FRENCH RESOURCES FOR EXTRACTION AND NORMALIZATION OF TEMPORAL EXPRESSIONS WITH HEIDELTIME</b> .....	2337
<i>Véronique Moriceau, Xavier Tannier</i>	
<b>LEGAL ASPECTS OF TEXT MINING</b> .....	2342
<i>Maarten Truyens, Patrick Van Eecke</i>	
<b>TREELET PROBABILITIES FOR HPSG PARSING AND ERROR CORRECTION</b> .....	2347
<i>Angelina Ivanova, Gertjan Van Noord</i>	
<b>A CORPUS AND PHONETIC DICTIONARY FOR TUNISIAN ARABIC SPEECH RECOGNITION</b> .....	2353
<i>Abir Masmoudi, Mariem Ellouze Khmekhem, Yannick Esteve, Lamia Hadrich Belguith, Nizar Habash</i>	
<b>DISCOVERING FRAMES IN SPECIALIZED DOMAINS</b> .....	2358
<i>Marie-Claude L'Homme, Benoît Robichaud, Carlos Subirats Rüggeberg</i>	
<b>RESOURCES FOR THE DETECTION OF CONVENTIONALIZED METAPHORS IN FOUR LANGUAGES</b> .....	2366
<i>Lori Levin, Teruko Mitamura, Brian Macwhinney, Davida Fromm, Jaime Carbonell, Weston Feely, Robert Frederking, Anatole Gershman, Carlos Ramirez</i>	
<b>CLARIN-NL: MAJOR RESULTS</b> .....	2370
<i>Jan Odijk</i>	
<b>EXPLOITING PORTUGUESE LEXICAL KNOWLEDGE BASES FOR ANSWERING OPEN DOMAIN CLOZE QUESTIONS AUTOMATICALLY</b> .....	2377
<i>Hugo Gonçalo Oliveira, Inês Coelho, Paulo Gomes</i>	
<b>ANNOTATION OF COMPUTER SCIENCE PAPERS FOR SEMANTIC RELATION EXTRACTION</b> .....	2385
<i>Yuka Tateisi, Yo Shidahara, Yusuke Miyao, Akiko Aizawa</i>	
<b>IDENTIFYING IDIOMS IN CHINESE TRANSLATIONS</b> .....	2392
<i>Wan Yu Ho, Christine Kng, Shan Wang, Francis Bond</i>	
<b>MACHINE TRANSLATION FOR SUBTITLING: A LARGE-SCALE EVALUATION</b> .....	2398
<i>Thierry Etchegoyhen, Lindsay Bywood, Mark Fishel, Panayota Georgakopoulou, Jie Jiang, Gerard Van Loenhout, Arantza Del Pozo, Mirjam Sepesy Maucec, Anja Turner, Martin Volk</i>	
<b>T-PAS; A RESOURCE OF TYPED PREDICATE ARGUMENT STRUCTURES FOR LINGUISTIC ANALYSIS AND SEMANTIC PROCESSING</b> .....	2406
<i>Elisabetta Jezek, Bernardo Magnini, Anna Feltracco, Alessia Bianchini, Octavian Popescu</i>	
<b>NARROWING THE GAP BETWEEN TERMBASES AND CORPORA IN COMMERCIAL ENVIRONMENTS</b> .....	2412
<i>Kara Warburton</i>	
<b>AUTHOR-SPECIFIC SENTIMENT AGGREGATION FOR POLARITY PREDICTION OF REVIEWS</b> .....	2418
<i>Subhabrata Mukherjee, Sachindra Joshi</i>	
<b>CLUSTERING OF MULTI-WORD NAMED ENTITY VARIANTS: MULTILINGUAL EVALUATION</b> .....	2426
<i>Guillaume Jacquet, Maud Ehrmann, Ralf Steinberger</i>	
<b>A DATABASE FOR MEASURING LINGUISTIC INFORMATION CONTENT</b> .....	2432
<i>Richard Sproat, Bruno Cartoni, Hyunjeong Choe, David Huynh, Linne Ha, Ravindran Rajakumar, Evelyn Wenzel-Grondie</i>	
<b>DESIGNING AND EVALUATING A RELIABLE CORPUS OF WEB GENRES VIA CROWD-SOURCING</b> .....	2440
<i>Noushin Rezapour Asheghi, Serge Sharoff, Katja Markert</i>	
<b>CROWDSOURCING AS A PREPROCESSING FOR COMPLEX SEMANTIC ANNOTATION TASKS</b> .....	2448
<i>Héctor Martínez Alonso, Lauren Romeo</i>	

<b>AUTOMATIC ANNOTATION OF MACHINE TRANSLATION DATASETS WITH BINARY QUALITY JUDGEMENTS</b> .....	2454
<i>Marco Turchi, Matteo Negri</i>	
<b>SEMANTIC CLUSTERING OF PIVOT PARAPHRASES</b> .....	2459
<i>Marianna Apidianaki, Emilia Verzeni, Diana McCarthy</i>	
<b>WHEN POS DATA SETS DON'T ADD UP: COMBATTING SAMPLE BIAS</b> .....	2465
<i>Dirk Hovy, Barbara Plank, Anders Søgaard</i>	
<b>OUT IN THE OPEN: FINDING AND CATEGORISING ERRORS IN THE LEXICAL SIMPLIFICATION PIPELINE</b> .....	2469
<i>Matthew Shardlow</i>	
<b>THE N2 CORPUS: A SEMANTICALLY ANNOTATED COLLECTION OF ISLAMIST EXTREMIST STORIES</b> .....	2477
<i>Mark Finlayson, Jeffrey Halverson, Steven Corman</i>	
<b>LEARNING FROM DOMAIN COMPLEXITY</b> .....	2484
<i>Robert Remus, Dominique Ziegelmayer</i>	
<b>BENCHMARKING TWITTER SENTIMENT ANALYSIS TOOLS</b> .....	2492
<i>Ahmed Abbasi, Ammar Hassan, Milan Dhar</i>	
<b>DESIGNING A BILINGUAL SPEECH CORPUS FOR FRENCH AND GERMAN LANGUAGE LEARNERS: A TWO-STEP PROCESS</b> .....	2499
<i>Camille Fauth, Anne Bonneau, Frank Zimmerer, Juergen Trouvain, Bistra Andreeva, Vincent Colotte, Dominique Fohr, Denis Jouvet, Jeanin Jugler, Yves Laprie, Odile Mella, Bernd Möbius</i>	
<b>CREATIVE LANGUAGE EXPLORATIONS THROUGH A HIGH-EXPRESSIVITY N-GRAMS QUERY LANGUAGE</b> .....	2505
<i>Carlo Strapparava, Lorenzo Gatti, Marco Guerini, Oliviero Stock</i>	
<b>USING WORD FAMILIARITIES AND WORD ASSOCIATIONS TO MEASURE CORPUS REPRESENTATIVENESS</b> .....	2510
<i>Reinhard Rapp</i>	
<b>DEEP SYNTAX ANNOTATION OF THE SEQUOIA FRENCH TREEBANK</b> .....	2518
<i>Marie Candito, Guy Perrier, Bruno Guillaume, Corentin Ribeyre, Karèn Fort, Djamé Seddah, Eric De La Clergerie</i>	
<b>DEVELOPING A FRENCH FRAMENET: METHODOLOGY AND FIRST RESULTS</b> .....	2526
<i>Marie Candito, Pascal Amsili, Lucie Barque, Farah Benamara, Gaël De Chalendar, Marianne Djemaa, Pauline Haas, Richard Huyghe, Yvette Yannick Mathieu, Philippe Muller, Benoit Sagot, Laure Vieu</i>	
<b>CORPUS ANNOTATION THROUGH CROWDSOURCING: TOWARDS BEST PRACTICE GUIDELINES</b> .....	2534
<i>Marta Sabou, Kalina Bontcheva, Leon Derczynski, Arno Scharl</i>	
<b>LOCATING REQUESTS AMONG OPEN SOURCE SOFTWARE COMMUNICATION MESSAGES</b> .....	2542
<i>Ioannis Korkontzelos, Sophia Ananiadou</i>	
<b>VALENCY AND WORD ORDER IN CZECH - A CORPUS PROBE</b> .....	2550
<i>Katerina Rysova, Jiri Mirovský</i>	
<b>HARMONIZATION OF GERMAN LEXICAL RESOURCES FOR OPINION MINING</b> .....	2556
<i>Thierry Declerck, Hans-Ulrich Krieger</i>	
<b>WORD-FORMATION NETWORK FOR CZECH</b> .....	2561
<i>Magda Sevcikova, Zdenek Zabokrtsky</i>	
<b>AN ANALYSIS OF OLDER USERS' INTERACTIONS WITH SPOKEN DIALOGUE SYSTEMS</b> .....	2568
<i>Jamie Bost, Johanna Moore</i>	
<b>INNOVATIONS IN PARALLEL CORPUS SEARCH TOOLS</b> .....	2574
<i>Martin Volk, Johannes Graën, Elena Callegaro</i>	
<b>MACHINE TRANSLATIONNESS: MACHINE-LIKENESS IN MACHINE TRANSLATION EVALUATION</b> .....	2581
<i>Joaquim Moré, Salvador Climent</i>	
<b>TOWARDS AN ENVIRONMENT FOR THE PRODUCTION AND THE VALIDATION OF LEXICAL SEMANTIC RESOURCES</b> .....	2589
<i>Mikaël Morardo, Eric De La Clergerie</i>	
<b>THE DISTRESS ANALYSIS INTERVIEW CORPUS OF HUMAN AND COMPUTER INTERVIEWS</b> .....	2597
<i>Jonathan Gratch, Ron Artstein, Gale Lucas, Giota Stratou, Stefan Scherer, Angela Nazarian, Rachel Wood, Jill Boberg, David Devault, Stacy Marsella, David Traum, Albert Rizzo, Louis-Philippe Morency</i>	
<b>REPRESENTING MULTIMODAL LINGUISTIC ANNOTATED DATA</b> .....	2603
<i>Brigitte Bigi, Tatsuya Watanabe, Laurent Prévot</i>	

<b>SWIFT ALIGNER, A MULTIFUNCTIONAL TOOL FOR PARALLEL CORPORA: VISUALIZATION, WORD ALIGNMENT, AND (MORPHO)-SYNTACTIC CROSS-LANGUAGE TRANSFER.....</b>	2610
<i>Timur Gilmanov, Olga Scrivner, Sandra Kübler</i>	
<b>SEMI-AUTOMATIC ANNOTATION OF THE UCU ACCENTS SPEECH CORPUS .....</b>	2617
<i>Rosemary Orr, Marijn Huijbregts, Roeland Van Beek, Lisa Teunissen, Kate Backhouse, David Van Leeuwen</i>	
<b>COMPARATIVE ANALYSIS OF PORTUGUESE NAMED ENTITIES RECOGNITION TOOLS.....</b>	2622
<i>Daniela Amaral, Evandro Fonseca, Lucelene Lopes, Renata Vieira</i>	
<b>A CORPUS OF EUROPEAN PORTUGUESE CHILD AND CHILD-DIRECTED SPEECH.....</b>	2627
<i>Ana Lúcia Santos, Michel Génèreux, Aida Cardoso, Celina Agostinho, Silvana Abalada</i>	
<b>USING C5.0 AND EXHAUSTIVE SEARCH FOR BOOSTING FRAME-SEMANTIC PARSING ACCURACY .....</b>	2631
<i>Guntis Barzdins, Didzis Gosko, Laura Rituma, Peteris Paikens</i>	
<b>INTERHIST' Ĩ AN INTERACTIVE VISUAL INTERFACE FOR CORPUS EXPLORATION.....</b>	2638
<i>Verena Lyding, Lionel Nicolas, Egon Stemle</i>	
<b>IDENTIFICATION OF MULTIWORD EXPRESSIONS IN THE BRWAC .....</b>	2645
<i>Rodrigo Boos, Kassius Prestes, Aline Villavicencio</i>	
<b>COLLOCATION OR FREE COMBINATION? ? APPLYING MACHINE TRANSLATION TECHNIQUES TO IDENTIFY COLLOCATIONS IN JAPANESE.....</b>	2653
<i>Lis Pereira, Elga Strafella, Yuji Matsumoto</i>	
<b>EXTRINSIC CORPUS EVALUATION WITH A COLLOCATION DICTIONARY TASK .....</b>	2657
<i>Adam Kilgarriř, Pavel Rychly, Milos Jakubicek, Vojtech Kovar, Vit Baisa, Lucia Kocincov</i>	
<b>AUSTALK: AN AUDIO-VISUAL CORPUS OF AUSTRALIAN ENGLISH.....</b>	2665
<i>Dominique Estival, Steve Cassidy, Felicity Cox, Denis Burnham</i>	
<b>COMPREHENSIVE ANNOTATION OF MULTIWORD EXPRESSIONS IN A SOCIAL WEB CORPUS .....</b>	2670
<i>Nathan Schneider, Spencer Onuffer, Nora Kazour, Emily Danchik, Michael T. Mordowanec, Henrietta Conrad, Noah A. Smith</i>	
<b>AUTOMATIC SEMANTIC RELATION EXTRACTION FROM PORTUGUESE TEXTS.....</b>	2677
<i>Leonardo Sameshima Taba, Helena Caseli</i>	
<b>A MULTIDIALECTAL PARALLEL CORPUS OF ARABIC .....</b>	2685
<i>Houda Bouamor, Nizar Habash, Kemal Oflazer</i>	
<b>TRANSFER LEARNING OF FEEDBACK HEAD EXPRESSIONS IN DANISH AND POLISH COMPARABLE MULTIMODAL CORPORA .....</b>	2691
<i>Costanza Navarretta, Magdalena Lis</i>	
<b>COMPARING TWO ACQUISITION SYSTEMS FOR AUTOMATICALLY BUILDING AN ENGLISH?CROATIAN PARALLEL CORPUS FROM MULTILINGUAL WEBSITES .....</b>	2698
<i>Miquel Espl-Gomis, Filip Klubicka, Nikola Ljubešic, Sergio Ortiz-Rojas, Vasilis Papavassiliou, Prokopis Prokopidis</i>	
<b>HASHTAG OCCURRENCES, LAYOUT AND TRANSLATION: A CORPUS-DRIVEN ANALYSIS OF TWEETS PUBLISHED BY THE CANADIAN GOVERNMENT .....</b>	2705
<i>Fabrizio Gotti, Phillippe Langlais, Atefeh Farzindar</i>	
<b>TOWARDS AN INTEGRATION OF SYNTACTIC AND TEMPORAL ANNOTATIONS IN ESTONIAN .....</b>	2713
<i>Siim Orasmaa</i>	
<b>HURIC: A HUMAN ROBOT INTERACTION CORPUS .....</b>	2721
<i>Emanuele Bastianelli, Giuseppe Castellucci, Danilo Croce, Luca Iocchi, Roberto Basili, Daniele Nardi</i>	
<b>ON THE ORIGIN OF ERRORS: A FINE-GRAINED ANALYSIS OF MT AND PE ERRORS AND THEIR RELATIONSHIP .....</b>	2729
<i>Joke Daems, Lieve Macken, Sonia Vandepitte</i>	
<b>THE WAVESURFER AUTOMATIC SPEECH RECOGNITION PLUGIN .....</b>	2734
<i>Giampiero Salvi, Niklas Vanhainen</i>	
<b>FREE ENGLISH AND CZECH TELEPHONE SPEECH CORPUS SHARED UNDER THE CC-BY-SA 3.0 LICENSE.....</b>	2739
<i>Matej Korvas, Ondrej Pltek, Ondrej Dusek, Lukas Zilka, Filip Jurcicek</i>	
<b>CREATING A GOLD STANDARD CORPUS FOR THE EXTRACTION OF CHEMISTRY-DISEASE RELATIONS FROM PATENT TEXTS.....</b>	2745
<i>Antje Schlaf, Claudia Bobach, Matthias Irmer</i>	
<b>THE SSPNET-MOBILE CORPUS: SOCIAL SIGNAL PROCESSING OVER MOBILE PHONES.....</b>	2750
<i>Anna Polychroniou, Hugues Salamin, Alessandro Vinciarelli</i>	
<b>PROJECTION-BASED ANNOTATION OF A POLISH DEPENDENCY TREEBANK.....</b>	2757
<i>Alina Wrblewska, Adam Przepirkowski</i>	

<b>BOOTSTRAPPING AN ITALIAN VERBNET: DATA-DRIVEN ANALYSIS OF VERB ALTERNATIONS</b> .....	2764
<i>Gianluca Lebani, Veronica Viola, Alessandro Lenci</i>	
<b>SELF-TRAINING A CONSTITUENCY PARSER USING N-GRAM TREES</b> .....	2772
<i>Arda Celebi, Arzucan Ozgur</i>	
<b>A TAGGED CORPUS AND A TAGGER FOR URDU</b> .....	2776
<i>Bushra Jawaid, Amir Kamran, Ondrej Bojar</i>	

Volume 4

<b>LEXICAL SUBSTITUTION DATASET FOR GERMAN</b> .....	2782
<i>Kostadin Cholakov, Chris Biemann, Judith Eckle-Kohler, Iryna Gurevych</i>	
<b>A CASCADE APPROACH FOR COMPLEX-TYPE CLASSIFICATION</b> .....	2788
<i>Lauren Romeo, Sara Mendes, N�ria Bel</i>	
<b>GENERATING A RESOURCE FOR PRODUCTS AND BRANDNAMES RECOGNITION. APPLICATION TO THE COSMETIC DOMAIN.</b> .....	2796
<i>C�dric Lopez, Fr�d�rique Segond, Olivier Hondermarck, Paolo Curtoni, Luca Dini</i>	
<b>ANNOTATION OF SPECIALIZED CORPORA USING A COMPREHENSIVE ENTITY AND RELATION SCHEME</b> .....	2802
<i>Louise Deleger, Anne-Laure Ligozat, Cyril Grouin, Pierre Zweigenbaum, Aurelie Neveol</i>	
<b>ANNOTATION PRO + TGA: AUTOMATION OF SPEECH TIMING ANALYSIS</b> .....	2810
<i>Katarzyna Klessa, Dafydd Gibbon</i>	
<b>POLYSEMY INDEX FOR NOUNS: AN EXPERIMENT ON ITALIAN USING THE PAROLE SIMPLE CLIPS LEXICAL DATABASE</b> .....	2817
<i>Francesca Frontini, Valeria Quochi, Sebastian Pad�, Monica Monachini, Jason Utt</i>	
<b>YODACC: THE YOUTUBE DIALECTAL ARABIC COMMENT CORPUS</b> .....	2826
<i>Ahmed Salama, Houda Bouamor, Behrang Mohit, Kemal Oflazer</i>	
<b>TOWARDS AN ENCYCLOPEDIA OF COMPOSITIONAL SEMANTICS: DOCUMENTING THE INTERFACE OF THE ENGLISH RESOURCE GRAMMAR</b> .....	2832
<i>Dan Flickinger, Emily M. Bender, Stephan Oepen</i>	
<b>ENRICHING THE "SENSO COMUNE" PLATFORM WITH AUTOMATICALLY ACQUIRED DATA</b> .....	2839
<i>Tommaso Caselli, Laure Vieu, Carlo Strapparava, Guido Vetere</i>	
<b>ONLINE EXPERIMENTS WITH THE PERCY SOFTWARE FRAMEWORK - EXPERIENCES AND SOME EARLY RESULTS</b> .....	2847
<i>Christoph Draxler</i>	
<b>IMPROVING THE EXPLOITATION OF LINGUISTIC ANNOTATIONS IN ELAN</b> .....	2853
<i>Onno Crasborn, Han Sloetjes</i>	
<b>A DEEP CONTEXT GRAMMATICAL MODEL FOR AUTHORSHIP ATTRIBUTION</b> .....	2858
<i>Simon Fuller, Phil Maguire, Philippe Moser</i>	
<b>MANUAL ANALYSIS OF STRUCTURALLY INFORMED REORDERING IN GERMAN-ENGLISH MACHINE TRANSLATION</b> .....	2863
<i>Teresa Herrmann, Jan Niehues, Alex Waibel</i>	
<b>PRESENTING A SYSTEM OF HUMAN-MACHINE INTERACTION FOR PERFORMING MAP TASKS.</b> .....	2871
<i>Gabriele Pallotti, Francesca Frontini, Fabio Aff�, Monica Monachini, Stefania Ferrari</i>	
<b>AUTOMATIC EXTRACTION OF SYNONYMS FOR GERMAN PARTICLE VERBS FROM PARALLEL DATA WITH DISTRIBUTIONAL SIMILARITY AS A RE-RANKING FEATURE</b> .....	2875
<i>Moritz Wittmann, Marion Weller, Sabine Schulte Im Walde</i>	
<b>NASTIA: NEGOTIATING APPOINTMENT SETTING INTERFACE</b> .....	2883
<i>Layla El Asri, R�mi Lemonnier, Romain Laroche, Olivier Pietquin, Hatim Khouzaimi</i>	
<b>DINASTI: DIALOGUES WITH A NEGOTIATING APPOINTMENT SETTING INTERFACE</b> .....	2889
<i>Layla El Asri, Romain Laroche, Olivier Pietquin</i>	
<b>LQVSUMM: A CORPUS OF LINGUISTIC QUALITY VIOLATIONS IN MULTI-DOCUMENT SUMMARIZATION</b> .....	2896
<i>Annemarie Friedrich, Marina Valeeva, Alexis Palmer</i>	
<b>POTSDAM COMMENTARY CORPUS 2.0: ANNOTATION FOR DISCOURSE RESEARCH</b> .....	2905
<i>Manfred Stede, Arne Neumann</i>	
<b>GL�FF, A LARGE VERSATILE FRENCH LEXICON</b> .....	2910
<i>Nabil Hathout, Franck Sajous, Basilio Calderone</i>	

<b>DENSE COMPONENTS IN THE STRUCTURE OF WORDNET</b> .....	2916
<i>Ahti Lohk, Kaarel Allik, Heili Orav, Leo Võhandu</i>	
<b>CHOOSING WHICH TO USE? A STUDY OF DISTRIBUTIONAL MODELS FOR NOMINAL LEXICAL SEMANTIC CLASSIFICATION</b> .....	2921
<i>Lauren Romeo, Gianluca Lebani, Núria Bel, Alessandro Lenci</i>	
<b>EXTENSIONS OF THE SIGN LANGUAGE RECOGNITION AND TRANSLATION CORPUS RWTH-PHOENIX-WEATHER</b> .....	2929
<i>Jens Forster, Christoph Schmidt, Oscar Koller, Martin Bellgardt, Hermann Ney</i>	
<b>HESITA(TE) IN PORTUGUESE</b> .....	2935
<i>Sara Candeias, Dirce Celorico, Jorge Proença, Arlindo Veiga, Carla Lopes, Fernando Perdigao</i>	
<b>MUHIT: A MULTILINGUAL HARMONIZED DICTIONARY</b> .....	2939
<i>Sameh Alansary</i>	
<b>PREDICATE MATRIX: EXTENDING SEMLINK THROUGH WORDNET MAPPINGS</b> .....	2947
<i>Maddalen Lopez De Lacalle, Egoitz Laparra, German Rigau</i>	
<b>TALAPI - A THAI LINGUISTICALLY ANNOTATED CORPUS FOR LANGUAGE PROCESSING</b> .....	2954
<i>Aiti Aw, Sharifah Mahani Aljunied, Nattadaporn Lertcheva, Sasiwimon Kalunsima</i>	
<b>T2K<sup>2</sup>: A SYSTEM FOR AUTOMATICALLY EXTRACTING AND ORGANIZING KNOWLEDGE FROM TEXTS</b> .....	2962
<i>Felice Dell'Orletta, Giulia Venturi, Andrea Cimino, Simonetta Montemagni</i>	
<b>EMOVO CORPUS: AN ITALIAN EMOTIONAL SPEECH DATABASE</b> .....	2971
<i>Giovanni Costantini, Iacopo Iaderola, Andrea Paoloni, Massimiliano Todisco</i>	
<b>MADAMIRA: A FAST, COMPREHENSIVE TOOL FOR MORPHOLOGICAL ANALYSIS AND DISAMBIGUATION OF ARABIC</b> .....	2975
<i>Arfath Pasha, Mohamed Al-Badrashiny, Mona Diab, Ahmed El Kholy, Ramy Eskander, Nizar Habash, Manoj Pooleery, Owen Rambow, Ryan Roth</i>	
<b>DEVELOPING POLITENESS ANNOTATED CORPUS OF HINDI BLOGS</b> .....	2983
<i>Ritesh Kumar</i>	
<b>THE CMU METAL FARSI NLP APPROACH</b> .....	2989
<i>Weston Feely, Mehdi Manshadi, Robert Frederking, Lori Levin</i>	
<b>RECOGNISING SUICIDAL MESSAGES IN DUTCH SOCIAL MEDIA</b> .....	2993
<i>Bart Desmet, Véronique Hoste</i>	
<b>A RANK-BASED DISTANCE MEASURE TO DETECT POLYSEMY AND TO DETERMINE SALIENT VECTOR-SPACE FEATURES FOR GERMAN PREPOSITIONS</b> .....	2999
<i>Maximilian Köper, Sabine Schulte Im Walde</i>	
<b>EXPANDING N-GRAM ANALYTICS IN ELAN AND A CASE STUDY FOR SIGN SYNTHESIS</b> .....	3007
<i>Rosalee Wolfe, John McDonald, Larwan Berke, Marie Stumbo</i>	
<b>SENTENCE REPHRASING FOR PARSING SENTENCES WITH OOV WORDS</b> .....	3013
<i>Hen-Hsen Huang, Huan-Yuan Chen, Chang-Sheng Yu, Hsin-Hsi Chen, Po-Ching Lee, Chun-Hsun Chen</i>	
<b>MAPPING THE LEXIQUE DES VERBES DU FRANCAIS (LEXICON OF FRENCH VERBS) TO A NLP LEXICON USING EXAMPLES</b> .....	3017
<i>Bruno Guillaume, Karèn Fort, Guy Perrier, Paul Bédaride</i>	
<b>LANGUAGE RESOURCES FOR FRENCH IN THE BIOMEDICAL DOMAIN</b> .....	3022
<i>Aurelie Neveol, Julien Grosjean, Stéfan Darmoni, Pierre Zweigenbaum</i>	
<b>THE MERLIN CORPUS: LEARNER LANGUAGE AND THE CEFR</b> .....	3028
<i>Adriane Boyd, Jirka Hana, Lionel Nicolas, Detmar Meurers, Katrin Wisniewski, Andrea Abel, Karin Schöne, Barbora Štindlová, Chiara Vettori</i>	
<b>COMPUTER-AIDED MORPHOLOGY EXPANSION FOR OLD SWEDISH</b> .....	3036
<i>Yvonne Adesam, Malin Ahlberg, Peter Andersson, Gerlof Bouma, Markus Forsberg, Mans Hulden</i>	
<b>TWO-STEP MACHINE TRANSLATION WITH LATTICES</b> .....	3040
<i>Bushra Jawaid, Ondrej Bojar</i>	
<b>THE MUNICH BIOVOICE CORPUS: EFFECTS OF PHYSICAL EXERCISING, HEART RATE, AND SKIN CONDUCTANCE ON HUMAN SPEECH PRODUCTION</b> .....	3045
<i>Björn Schuller, Felix Friedmann, Florian Eyben</i>	
<b>DYSLIST: AN ANNOTATED RESOURCE OF DYSLEXIC ERRORS</b> .....	3050
<i>Luz Rello, Ricardo Baeza-Yates, Joaquim Llisterri</i>	
<b>ESTIMATION OF SPEAKING STYLE IN SPEECH CORPORA FOCUSING ON SPEECH TRANSCRIPTIONS</b> .....	3058
<i>Raymond Shen, Hideaki Kikuchi</i>	
<b>EVALUATION OF DIFFERENT STRATEGIES FOR DOMAIN ADAPTATION IN OPINION MINING</b> .....	3064
<i>Anne Garcia-Fernandez, Olivier Ferret, Marco Dinarelli</i>	

<b>CLINICAL DATA-DRIVEN PROBABILISTIC GRAPH PROCESSING</b> .....	3068
<i>Travis Goodwin, Sanda Harabagiu</i>	
<b>COMPARING SIMILARITY MEASURES FOR DISTRIBUTIONAL THESAURI</b> .....	3076
<i>Muntsa Padró, Marco Idiart, Aline Villavicencio, Carlos Ramisch</i>	
<b>PRUNING THE SEARCH SPACE OF THE WOLOF LFG GRAMMAR USING A PROBABILISTIC AND A CONSTRAINT GRAMMAR PARSER</b> .....	3084
<i>Cheikh M. Bamba Dione</i>	
<b>ARANLP: A JAVA-BASED LIBRARY FOR THE PROCESSING OF ARABIC TEXT</b> .....	3092
<i>Maha Althobaiti, Udo Kruschwitz, Massimo Poesio</i>	
<b>CRITERIA FOR IDENTIFYING AND ANNOTATING CAUSED MOTION CONSTRUCTIONS IN CORPUS DATA</b> .....	3097
<i>Jena D. Hwang, Annie Zaenen, Martha Palmer</i>	
<b>WEB-IMAGEABILITY OF THE BEHAVIORAL FEATURES OF BASIC-LEVEL CONCEPTS</b> .....	3105
<i>Yoshihiko Hayashi</i>	
<b>THE ALVEO VIRTUAL LABORATORY: A WEB BASED REPOSITORY API</b> .....	3111
<i>Steve Cassidy, Dominique Estival, Timothy Jones, Denis Burnham, Jared Burghold</i>	
<b>A COMPACT INTERACTIVE VISUALIZATION OF DEPENDENCY TREEBANK QUERY RESULTS</b> .....	3118
<i>Chris Culy, Marco Passarotti, Ulla König-Cardanobile</i>	
<b>BUILDING A CRISIS MANAGEMENT TERM RESOURCE FOR SOCIAL MEDIA: THE CASE OF FLOODS AND PROTESTS</b> .....	3126
<i>Irina Temnikova, Andrea Varga, Dogan Biyikli</i>	
<b>ILLINOISCLOUDNLP: TEXT ANALYTICS SERVICES IN THE CLOUD</b> .....	3134
<i>Hao Wu, Zhiye Fei, Aaron Dai, Mark Sammons, Dan Roth, Stephen Mayhew</i>	
<b>TEXT READABILITY AND WORD DISTRIBUTION IN JAPANESE</b> .....	3142
<i>Satoshi Sato</i>	
<b>THE USE OF A FILEMAKER PRO DATABASE IN EVALUATING SIGN LANGUAGE NOTATION SYSTEMS</b> .....	3147
<i>Julie Hochgesang</i>	
<b>MAPPING CPA PATTERNS ONTO ONTONOTES SENSES</b> .....	3154
<i>Octavian Popescu, Martha Palmer, Patrick Hanks</i>	
<b>LANGUAGE COLLAGE: GRAMMATICAL DESCRIPTION WITH THE LINGO GRAMMAR MATRIX</b> .....	3162
<i>Emily M. Bender</i>	
<b>APPLYING ACCESSIBILITY-ORIENTED CONTROLLED LANGUAGE (CL) RULES TO IMPROVE APPROPRIATENESS OF TEXT ALTERNATIVES FOR IMAGES: AN EXPLORATORY STUDY</b> .....	3167
<i>Silvia Rodríguez Vázquez, Pierrette Bouillon, Anton Bolfin</i>	
<b>A MULTI-DIALECT, MULTI-GENRE CORPUS OF INFORMAL WRITTEN ARABIC</b> .....	3175
<i>Ryan Cotterell, Chris Callison-Burch</i>	
<b>RECONSTRUCTING THE SEMANTIC LANDSCAPE OF NATURAL LANGUAGE PROCESSING</b> .....	3180
<i>Elisa Omodei, Jean-Philippe Cointet, Thierry Poibeau</i>	
<b>DISCOVERING AND VISUALISING STORIES IN NEWS</b> .....	3187
<i>Marieke Van Erp, Gleb Satyukov, Piek Vossen, Marit Nijssen</i>	
<b>SUPERVISED WITHIN-DOCUMENT EVENT COREFERENCE USING INFORMATION PROPAGATION</b> .....	3193
<i>Zhengzhong Liu, Jun Araki, Eduard Hovy, Teruko Mitamura</i>	
<b>VOCE CORPUS: ECOLOGICALLY COLLECTED SPEECH ANNOTATED WITH PHYSIOLOGICAL AND PSYCHOLOGICAL STRESS ASSESSMENTS</b> .....	3199
<i>Ana Aguiar, Mariana Kaiseler, Hugo Meinedo, Pedro Almeida, Mariana Cunha, Jorge Silva</i>	
<b>LANGUAGE RESOURCE ADDITION: DICTIONARY OR CORPUS?</b> .....	3206
<i>Shinsuke Mori, Graham Neubig</i>	
<b>THE DIRHA SIMULATED CORPUS</b> .....	3212
<i>Luca Cristoforetti, Mirco Ravanelli, Maurizio Omologo, Alessandro Sosi, Alberto Abad, Martin Hagnmueller, Petros Maragos</i>	
<b>THE SLOVAK CATEGORIZED NEWS CORPUS</b> .....	3218
<i>Daniel Hladek, Jan Stas, Jozef Juhar</i>	
<b>HIGH QUALITY WORD LISTS AS A RESOURCE FOR MULTIPLE PURPOSES</b> .....	3222
<i>Uwe Quasthoff, Dirk Goldhahn, Thomas Eckart, Erla Hallsteinsdóttir, Sabine Fiedler</i>	



<b>A QUALITY-BASED ACTIVE SAMPLE SELECTION STRATEGY FOR STATISTICAL MACHINE TRANSLATION</b> .....	3226
<i>Varvara Logacheva, Lucia Specia</i>	
<b>THE MULTILINGUAL PARAPHRASE DATABASE</b> .....	3232
<i>Juri Ganitkevitch, Chris Callison-Burch</i>	
<b>THE DEVELOPMENT OF DUTCH AND AFRIKAANS LANGUAGE RESOURCES FOR COMPOUND BOUNDARY ANALYSIS.</b> .....	3240
<i>Menno Van Zaanen, Gerhard Van Huyssteen, Suzanne Aussems, Chris Emmery, Roald Eiselen</i>	
<b>LANGUAGE PROCESSING INFRASTRUCTURE IN THE XLIKE PROJECT</b> .....	3247
<i>Lluís Padró, Zeljko Agic, Xavier Carreras, Blaz Fortuna, Esteban García-Cuesta, Zhixing Li, Tadej Stajner, Marko Tadic</i>	
<b>CONCEPTUAL TRANSFER: USING LOCAL CLASSIFIERS FOR TRANSFER SELECTION</b> .....	3253
<i>Gregor Thurmair</i>	
<b>METADATA AS LINKED OPEN DATA: MAPPING DISPARATE XML METADATA REGISTRIES INTO ONE RDF/OWL REGISTRY</b> .....	3260
<i>Marta Villegas, Maite Melero, Núria Bel</i>	
<b>SHARING RESOURCES BETWEEN FREE/OPEN-SOURCE RULE-BASED MACHINE TRANSLATION SYSTEMS: GRAMMATICAL FRAMEWORK AND APERTIUM</b> .....	3268
<i>Grégoire Détrez, Víctor M. Sánchez-Cartagena, Aarne Ranta</i>	
<b>ANNOTATING ARGUMENTS: THE NOMAD COLLABORATIVE ANNOTATION TOOL</b> .....	3275
<i>Georgios Petasis</i>	
<b>WHO CARES ABOUT SARCASTIC TWEETS? INVESTIGATING THE IMPACT OF SARCASM ON SENTIMENT ANALYSIS.</b> .....	3283
<i>Diana Maynard, Mark Greenwood</i>	
<b>A STREAM COMPUTING APPROACH TOWARDS SCALABLE NLP</b> .....	3289
<i>Xabier Artola, Zuhaitz Beloki, Aitor Soroa</i>	
<b>ISLEX ? A MULTILINGUAL WEB DICTIONARY</b> .....	3295
<i>Pórdís Úlfarsdóttir</i>	
<b>EXPLOITING CATENAE IN A PARALLEL TREEBANK ALIGNMENT</b> .....	3301
<i>Manuela Sanguinetti, Cristina Bosco, Loredana Cupi</i>	
<b>SUBLANGUAGE CORPUS ANALYSIS TOOLKIT: A TOOL FOR ASSESSING THE REPRESENTATIVENESS AND SUBLANGUAGE CHARACTERISTICS OF CORPORA</b> .....	3309
<i>Irina Temnikova, William A. Baumgartner, Negacy D. Hailu, Ivelina Nikolova, Tony McEnery, Adam Kilgarriff, Galia Angelova, K. Bretonnel Cohen</i>	
<b>A LARGE-SCALE EVALUATION OF PRE-EDITING STRATEGIES FOR IMPROVING USER-GENERATED CONTENT TRANSLATION</b> .....	3314
<i>Violeta Seretan, Pierrette Bouillon, Johanna Gerlach</i>	
<b>CORRECTING ERRORS IN A NEW GOLD STANDARD FOR TAGGING ICELANDIC TEXT</b> .....	3321
<i>Sigrún Helgadóttir, Hrafn Loftsson, Eiríkur Rögnvaldsson</i>	
<b>SEMI-COMPOSITIONAL METHOD FOR SYNONYM EXTRACTION OF MULTI-WORD TERMS</b> .....	3326
<i>Béatrice Daille, Amir Hazem</i>	
<b>TUKE-BNEWS-SK: SLOVAK BROADCAST NEWS CORPUS CONSTRUCTION AND EVALUATION</b> .....	3332
<i>Matus Pleva, Jozef Juhar</i>	
<b>THE HUNGARIAN GIGAWORD CORPUS</b> .....	3337
<i>Csaba Oravecz, Tamás Váradi, Bálint Sass</i>	
<b>HINDI TO ENGLISH MACHINE TRANSLATION: USING EFFECTIVE SELECTION IN MULTI-MODEL SMT</b> .....	3342
<i>Kunal Sachdeva, Rishabh Srivastava, Sambhav Jain, Dipti Sharma</i>	
<b>FROM NATURAL LANGUAGE TO ONTOLOGY POPULATION IN THE CULTURAL HERITAGE DOMAIN. A COMPUTATIONAL LINGUISTICS-BASED APPROACH.</b> .....	3347
<i>Maria Pia Di Buono, Mario Monteleone</i>	
<b>EXPERIENCES WITH PARALLELISATION OF AN EXISTING NLP PIPELINE: TAGGING HANSARD</b> .....	3353
<i>Stephen Wattam, Paul Rayson, Marc Alexander, Jean Anderson</i>	
<b>NAMED ENTITY CORPUS CONSTRUCTION USING WIKIPEDIA AND DBPEDIA ONTOLOGY</b> .....	3357
<i>Younggyun Hahn, Jungyeul Park, Kyungtae Lim, Youngsik Kim, Dosam Hwang, Key-Sun Choi</i>	
<b>A MODEL TO GENERATE ADAPTIVE MULTIMODAL JOB INTERVIEWS WITH A VIRTUAL RECRUITER</b> .....	3362
<i>Zoraida Callejas, Brian Ravenet, Magalie Ochs, Catherine Pelachaud</i>	

<b>THE SETIMES.HR LINGUISTICALLY ANNOTATED CORPUS OF CROATIAN</b> .....	3367
<i>Željko Agić, Nikola Ljubešić</i>	
<b>ACTIV-ES: A COMPARABLE, CROSS-DIALECT CORPUS OF ‘EVERYDAY’ SPANISH FROM ARGENTINA, MEXICO, AND SPAIN</b> .....	3371
<i>Jerid Francom, Mans Hulden, Adam Ussishkin</i>	
<b>MULTIPLE CHOICE QUESTION CORPUS ANALYSIS FOR DISTRACTOR CHARACTERIZATION</b> .....	3376
<i>Van-Minh Pho, Thibault André, Anne-Laure Ligozat, Brigitte Grau, Gabriel Illouz, Thomas Francois</i>	
<b>CROATIAN DEPENDENCY TREEBANK 2.0: NEW ANNOTATION GUIDELINES FOR IMPROVED PARSING</b> .....	3384
<i>Željko Agić, Daša Berovic, Danijela Merkle, Marko Tadić</i>	
<b>EURONEWS: A MULTILINGUAL SPEECH CORPUS FOR ASR</b> .....	3391
<i>Roberto Gretter</i>	
<b>CONSTITUENCY PARSING OF BULGARIAN: WORD- VS CLASS-BASED PARSING</b> .....	3395
<i>Masood Ghayoomi, Kiril Simov, Petya Osenova</i>	
<b>A MULTIMODAL CORPUS OF RAPID DIALOGUE GAMES</b> .....	3400
<i>Maike Paetzel, David Nicolas Racca, David Devault</i>	
<b>NEW SPANISH SPEECH CORPUS DATABASE FOR THE ANALYSIS OF PEOPLE SUFFERING FROM PARKINSON'S DISEASE</b> .....	3407
<i>Juan Rafael Orozco-Arroyave, Julián David Arias-Londoño, Jesús Francisco Vargas-Bonilla, María Claudia Gonzalez-Rátiva, Elmar Nöth</i>	
<b>THOMAS AQUINAS IN THE TUNDRA: INTEGRATING THE INDEX THOMISTICUS TREEBANK INTO CLARIN-D</b> .....	3413
<i>Scott Martens, Marco Passarotti</i>	
<b>IDENTIFICATION OF TECHNOLOGY TERMS IN PATENTS</b> .....	3421
<i>Peter Anick, Marc Verhagen, James Pustejovsky</i>	
<b>TOWARDS LINKED HYPERNYMS DATASET 2.0: COMPLEMENTING DBPEDIA WITH HYPERNYM DISCOVERY</b> .....	3428
<i>Tomas Kliegr, Ondrej Zamazal</i>	
<b>AUTOMATIC MAPPING LEXICAL RESOURCES: A LEXICAL UNIT AS THE KEYSTONE</b> .....	3435
<i>Eduard Bejcek, Kettnerová Václava, Marketa Lopatkova</i>	
<b>TERMWISE: A CAT-TOOL WITH CONTEXT-SENSITIVE TERMINOLOGICAL SUPPORT</b> .....	3442
<i>Kris Heylen, Stephen Bond, Dirk De Hertog De Hertog, Ivan Vulic, Hendrik Kockaert</i>	
<b>MEASURING THE IMPACT OF SPELLING ERRORS ON THE QUALITY OF MACHINE TRANSLATION</b> .....	3447
<i>Irina Galinskaya, Valentin Gusev, Elena Mescheryakova, Mariya Shmatova</i>	
<b>TOWARDS MULTILINGUAL CONVERSATIONS IN THE MEDICAL DOMAIN: DEVELOPMENT OF MULTILINGUAL MEDICAL DATA AND A NETWORK-BASED ASR SYSTEM</b> .....	3454
<i>Sakriani Sakti, Keigo Kubo, Sho Matsumiya, Graham Neubig, Tomoki Toda, Satoshi Nakamura, Fumihiko Adachi, Ryosuke Isotani</i>	
<b>AUTOMATIC DETECTION OF OTHER-REPETITION OCCURRENCES: APPLICATION TO FRENCH CONVERSATIONAL SPEECH</b> .....	3459
<i>Brigitte Bigi, Roxane Bertrand, Mathilde Guardiola</i>	
<b>THE SLOVENE BNSI BROADCAST NEWS DATABASE AND REFERENCE SPEECH CORPUS GOS: TOWARDS THE UNIFORM GUIDELINES FOR FUTURE WORK</b> .....	3466
<i>Andrej Zgank, Ana Zwitter Vitez, Darinka Verdonik</i>	
<b>FROM SYNSETS TO VIDEOS: ENRICHING ITALWORDNET MULTIMODALLY</b> .....	3470
<i>Roberto Bartolini, Valeria Quochi, Irene De Felice, Irene Russo, Monica Monachini</i>	
<b>LANGUAGE EDITING DATASET OF ACADEMIC TEXTS</b> .....	3478
<i>Vidas Daudaravicius</i>	
<b>A TOOLKIT FOR EFFICIENT LEARNING OF LEXICAL UNITS FOR SPEECH RECOGNITION</b> .....	3483
<i>Matti Varjokallio, Mikko Kurimo</i>	
<b>TOWARDS AUTOMATIC TRANSFORMATION BETWEEN DIFFERENT TRANSCRIPTION CONVENTIONS: PREDICTION OF INTONATION MARKERS FROM LINGUISTIC AND ACOUSTIC FEATURES</b> .....	3487
<i>Yuichi Ishimoto, Tomoyuki Tsuchiya, Hanae Koiso, Yasuharu Den</i>	
<b>JAPANESE CONVERSATION CORPUS FOR TRAINING AND EVALUATION OF BACKCHANNEL PREDICTION MODEL</b> .....	3492
<i>Hiroaki Noguchi, Yasuhiro Katagiri, Yasuharu Den</i>	

<b>AIX MAP TASK CORPUS: THE FRENCH MULTIMODAL CORPUS OF TASK-ORIENTED DIALOGUE</b> .....	3497
<i>Jan Gorisch, Corine Astésano, Ellen Gurman Bard, Brigitte Bigi, Laurent Prévot</i>	
<b>ADAPTING A PART-OF-SPEECH TAGSET TO NON-STANDARD TEXT: THE CASE OF STTS</b> .....	3502
<i>Heike Zinsmeister, Ulrich Heid, Kathrin Beck</i>	
<b>ALERT!... CALM DOWN, THERE IS NOTHING TO WORRY ABOUT. WARNING AND SOOTHING SPEECH SYNTHESIS.</b> .....	3510
<i>Milan Rusko, Sakhia Darjaa, Marian Trnka, Marian Ritomsky, Robert Sabo</i>	
<b>MULTIWORD EXPRESSIONS IN MACHINE TRANSLATION</b> .....	3516
<i>Valia Kordoni, Iliana Simova</i>	
<b>CROMER: A TOOL FOR CROSS-DOCUMENT EVENT AND ENTITY COREFERENCE</b> .....	3520
<i>Christian Girardi, Manuela Speranza, Rachele Sprugnoli, Sara Tonelli</i>	
<b>RSS-TOBI - A PROSODICALLY ENHANCED ROMANIAN SPEECH CORPUS</b> .....	3525
<i>Tiberiu Boros, Adriana Stan, Oliver Watts, Stefan Daniel Dumitrescu</i>	
<b>COREFERENCE RESOLUTION FOR LATVIAN</b> .....	3530
<i>Arturs Znotins, Peteris Paikens</i>	
<b>HOW COULD VEINS SPEED UP THE PROCESS OF DISCOURSE PARSING</b> .....	3535
<i>Elena Mitocariu, Daniel Anechitei, Dan Cristea</i>	
<b>HOW TO CONSTRUCT A MULTI-LINGUAL DOMAIN ONTOLOGY</b> .....	3543
<i>Nitsan Chrizman, Alon Itai</i>	
<b>AUTOMATIC LANGUAGE IDENTITY TAGGING ON WORD AND SENTENCE-LEVEL IN MULTILINGUAL TEXT SOURCES: A CASE-STUDY ON LUXEMBOURGISH</b> .....	3549
<i>Thomas Lavergne, Gilles Adda, Martine Adda-Decker, Lori Lamel</i>	
<b>TOWARDS SHARED DATASETS FOR NORMALIZATION RESEARCH</b> .....	3554
<i>Orphee De Clercq, Sarah Schulz, Bart Desmet, Veronique Hoste</i>	
<b>RULE-BASED REORDERING SPACE IN STATISTICAL MACHINE TRANSLATION</b> .....	3560
<i>Nicolas Pécheux, Alexander Allauzen, François Yvon</i>	
<b>KALAKA-3: A DATABASE FOR THE RECOGNITION OF SPOKEN EUROPEAN LANGUAGES ON YOUTUBE AUDIOS</b> .....	3567
<i>Luis Javier Rodriguez-Fuentes, Mikel Penagarikano, Amparo Varona, Mireia Diez, German Bordel</i>	
<b>MINING ONLINE DISCUSSION FORUMS FOR METAPHORS</b> .....	3574
<i>Andrew Gargett, John Barnden</i>	
<b>TAGNTEXT: A PARALLEL CORPUS FOR THE INDUCTION OF RESOURCE-SPECIFIC NON-TAXONOMICAL RELATIONS FROM TAGGED IMAGES</b> .....	3580
<i>Theodosia Togia, Ann Copestake</i>	
<b>CORILGA: A GALICIAN MULTILEVEL ANNOTATED SPEECH CORPUS FOR LINGUISTIC ANALYSIS</b> .....	3588
<i>Carmen Garcia-Mateo, Antonio Cardenal, Xose Luis Regueira, Elisa Fernández Rei, Marta Martínez, Roberto Seara, Rocío Varela, Noemi Basanta</i>	
<b>OVERVIEW OF TODAI ROBOT PROJECT AND EVALUATION FRAMEWORK OF ITS NLP-BASED PROBLEM SOLVING</b> .....	3593
<i>Akira Fujita, Akihiro Kameda, Ai Kawazoe, Yusuke Miyao</i>	
<b>A DATABASE OF FULL BODY VIRTUAL INTERACTIONS ANNOTATED WITH EXPRESSIVITY SCORES</b> .....	3601
<i>Demulier Virginie, Elisabetta Bevacqua, Florian Focone, Tom Giraud, Pamela Carreno, Brice Isableu, Sylvie Gibet, Pierre De Loor, Jean-Claude Martin</i>	
<b>ACCESS CONTROL BY QUERY REWRITING: THE CASE OF KORAP</b> .....	3607
<i>Piotr Banski, Nils Diewald, Michael Hanl, Marc Kupietz, Andreas Witt</i>	
<b>BASQUE SPEECON-LIKE AND BASQUE SPEECHDAT MDB-600: SPEECH DATABASES FOR THE DEVELOPMENT OF ASR TECHNOLOGY FOR BASQUE</b> .....	3613
<i>Igor Odriozola, Inma Hernaez, María Inés Torres, Luis Javier Rodriguez-Fuentes, Mikel Penagarikano, Eva Navas</i>	
<b>DIVE-ARABIC: GULF ARABIC DIALOGUE IN A VIRTUAL ENVIRONMENT</b> .....	3621
<i>Andrew Gargett, Sam Hellmuth, Ghazi Algethami</i>	
<b>A MULTIMODAL INTERPRETER FOR 3D VISUALIZATION AND ANIMATION OF VERBAL CONCEPTS</b> .....	3627
<i>Coline Claude-Lachenaud, Eric Charton, Benoît Ozell, Michel Gagnon</i>	
<b>ERLANGEN-CLP: A LARGE ANNOTATED CORPUS OF SPEECH FROM CHILDREN WITH CLEFT LIP AND PALATE</b> .....	3635
<i>Tobias Bocklet, Andreas Maier, Korbinian Riedhammer, Ulrich Eysholdt, Elmar Nöth</i>	
<b>CLASSIFYING INCONSISTENCIES IN DBPEDIA LANGUAGE SPECIFIC CHAPTERS</b> .....	3639
<i>Elena Cabrio, Serena Villata, Fabien Gandon</i>	

<b>TVD: A REPRODUCIBLE AND MULTIPLY ALIGNED TV SERIES DATASET</b> .....	3647
<i>Anindya Roy, Camille Guinaudeau, Herve Bredin, Claude Barras</i>	
<b>TOWARDS ELECTRONIC SMS DICTIONARY CONSTRUCTION: AN ALIGNMENT-BASED APPROACH</b> .....	3655
<i>Cédric Lopez, Reda Bestandji, Mathieu Roche, Rachel Panckhurst</i>	
<b>COMPOUNDS AND DISTRIBUTIONAL THESAURI</b> .....	3661
<i>Olivier Ferret</i>	
<b>TALC-SEF A MANUALLY-REVISED POS-TAGGED LITERARY CORPUS IN SERBIAN, ENGLISH AND FRENCH</b> .....	3667
<i>Antonio Balvet, Dejan Stosic, Aleksandra Miletic</i>	
<b>CROWDSOURCING FOR EVALUATING MACHINE TRANSLATION QUALITY</b> .....	3673
<i>Shinsuke Goto, Donghui Lin, Toru Ishida</i>	
<b>THE HALLIDAY CENTRE TAGGER: AN ONLINE PLATFORM FOR SEMI-AUTOMATIC TEXT ANNOTATION AND ANALYSIS</b> .....	3681
<i>Billy T. M. Wong, Ian C. Chow, Jonathan J. Webster, Hengbin Yan</i>	
<b>FLOW GRAPH CORPUS FROM RECIPE TEXTS</b> .....	3685
<i>Shinsuke Mori, Hirokuni Maeta, Yoko Yamakata, Tetsuro Sasada</i>	
<b>FREEPAL: A LARGE COLLECTION OF DEEP LEXICO-SYNTACTIC PATTERNS FOR RELATION EXTRACTION</b> .....	3693
<i>Johannes Kirschnick, Alan Akbik, Holmer Hensen</i>	
<b>CORRECTING AND VALIDATING SYNTACTIC DEPENDENCY IN THE SPOKEN FRENCH TREEBANK RHAPSODIE</b> .....	3698
<i>Rachel Bawden, Marie-Amélie Botalla, Kim Gerdes, Sylvain Kahane</i>	

Volume 5

<b>MORFEUSZ RELOADED</b> .....	3704
<i>Marcin Wolinski</i>	
<b>MODELING, MANAGING, EXPOSING, AND LINKING ONTOLOGIES WITH A WIKI-BASED TOOL</b> .....	3710
<i>Mauro Dragoni, Alessio Bosca, Matteo Casu, Andi Rexha</i>	
<b>A MODEL FOR PROCESSING ILLOCUTIONARY STRUCTURES AND ARGUMENTATION IN DEBATES</b> .....	3718
<i>Kasia Budzynska, Mathilde Janier, Chris Reed, Patrick Saint-Dizier, Manfred Stede, Olena Yakorska</i>	
<b>STUDENT ACHIEVEMENT AND FRENCH SENTENCE REPETITION TEST SCORES</b> .....	3726
<i>Deryle Lonsdale, Benjamin Millard</i>	
<b>HUMAN ANNOTATION OF ASR ERROR REGIONS: IS "GRAVITY" A SHARABLE CONCEPT FOR HUMAN ANNOTATORS?</b> .....	3733
<i>Daniel Luzzati, Cyril Grouin, Ioana Vasilescu, Martine Adda-Decker, Eric Bilinski, Nathalie Camelin, Juliette Kahn, Carole Lailler, Lori Lamel, Sophie Rosset</i>	
<b>SWISSADMIN: A MULTILINGUAL TAGGED PARALLEL CORPUS OF PRESS RELEASES</b> .....	3740
<i>Yves Scherrer, Luka Nerima, Lorenza Russo, Maria Ivanova, Eric Wehrli</i>	
<b>TO PAY OR TO GET PAID: ENRICHING A VALENCY LEXICON WITH DIATHESES</b> .....	3745
<i>Anna Vernerová, Václava Kettnerová, Marketa Lopatkova</i>	
<b>UM-CORPUS: A LARGE ENGLISH-CHINESE PARALLEL CORPUS FOR STATISTICAL MACHINE TRANSLATION</b> .....	3753
<i>Liang Tian, Derek F. Wong, Lidia S. Chao, Paulo Quaresma, Francisco Oliveira, Lu Yi</i>	
<b>IXA PIPELINE: EFFICIENT AND READY TO USE MULTILINGUAL NLP TOOLS</b> .....	3759
<i>Rodrigo Agerri, Josu Bermudez, German Rigau</i>	
<b>ANNOTATING THE FOCUS OF NEGATION IN JAPANESE TEXT</b> .....	3765
<i>Suguru Matsuyoshi, Ryo Otsuki, Fumiyo Fukumoto</i>	
<b>NIF4OGGD - NLP INTERCHANGE FORMAT FOR OPEN GERMAN GOVERNMENTAL DATA</b> .....	3773
<i>Mohamed Sherif, Sandro Coelho, Ricardo Usbeck, Sebastian Hellmann, Jens Lehmann, Martin Brümmer, Andreas Both</i>	
<b>A GOLD STANDARD FOR CLIR EVALUATION IN THE ORGANIC AGRICULTURE DOMAIN</b> .....	3778
<i>Alessio Bosca, Matteo Casu, Matteo Dragoni, Nikolaos Marianos</i>	
<b>HEURISTIC HYPER-MINIMIZATION OF FINITE STATE LEXICONS</b> .....	3782
<i>Senka Drobac, Krister Lindén, Tommi Pirinen, Miikka Silfverberg</i>	
<b>META-SHARE: ONE YEAR AFTER</b> .....	3788
<i>Stelios Piperidis, Harris Papageorgiou, Christian Spurk, Georg Rehm, Khalid Choukri, Olivier Hamon, Nicoletta Calzolari, Riccardo Del Gratta, Bernardo Magnini, Christian Girardi</i>	

<b>USING A SERIOUS GAME TO COLLECT A CHILD LEARNER SPEECH CORPUS</b> .....	3795
<i>Claudia Baur, Manny Rayner, Nikos Tsourakis</i>	
<b>THE LRE MAP DISCLOSED</b> .....	3802
<i>Riccardo Del Gratta, Gabriella Pardelli, Sara Goggi</i>	
<b>THE DEVELOPMENT OF THE MULTILINGUAL LUNA CORPUS FOR SPOKEN LANGUAGE SYSTEM PORTING</b> .....	3810
<i>Evgeny Stepanov, Giuseppe Riccardi, Ali Orkan Bayer</i>	
<b>VERBS OF SAYING WITH A TEXTUAL CONNECTING FUNCTION IN THE PRAGUE DISCOURSE TREEBANK</b> .....	3814
<i>Magdalena Rysova</i>	
<b>RANKING JOB OFFERS FOR CANDIDATES: LEARNING HIDDEN KNOWLEDGE FROM BIG DATA</b> .....	3820
<i>Marc Poch, Nria Bel, Sergio Espeja, Felipe Navio</i>	
<b>AN OPEN-SOURCE HEAVILY MULTILINGUAL TRANSLATION GRAPH EXTRACTED FROM WIKTIONARIES AND PARALLEL CORPORA</b> .....	3827
<i>Valrie Hanoka, Benot Sagot</i>	
<b>CROWD-SOURCING EVALUATION OF AUTOMATICALLY ACQUIRED, MORPHOLOGICALLY RELATED WORD GROUPINGS</b> .....	3835
<i>Claudia Borg, Albert Gatt</i>	
<b>AN INNOVATIVE WORLD LANGUAGE CENTRE : CHALLENGES FOR THE USE OF LANGUAGE TECHNOLOGY</b> .....	3843
<i>Auur Hauksdttir</i>	
<b>A LANGUAGE-INDEPENDENT AND FULLY UNSUPERVISED APPROACH TO LEXICON INDUCTION AND PART-OF-SPEECH TAGGING FOR CLOSELY RELATED LANGUAGES</b> .....	3848
<i>Yves Scherrer, Benot Sagot</i>	
<b>NEW BILINGUAL SPEECH DATABASES FOR AUDIO DIARIZATION</b> .....	3855
<i>David Tavarez, Eva Navas, Daniel Erro, Ibon Saratxaga, Inma Hernaez</i>	
<b>A LDA-BASED TOPIC CLASSIFICATION APPROACH FROM HIGHLY IMPERFECT AUTOMATIC TRANSCRIPTIONS</b> .....	3860
<i>Mohamed Morchid, Richard Dufour, Georges Linares</i>	
<b>AN OPEN SOURCE PART-OF-SPEECH TAGGER FOR NORWEGIAN: BUILDING ON EXISTING LANGUAGE RESOURCES</b> .....	3866
<i>Cristina Snchez Marco</i>	
<b>BILINGUAL DICTIONARIES FOR ALL EU LANGUAGES</b> .....	3873
<i>Ahmet Aker, Monica Paramita, Marcis Pinnis, Robert Gaizauskas</i>	
<b>SYNERGY OF NEDERLAB AND @PHILOSTEI: DIACHRONIC AND MULTILINGUAL TEXT-INDUCED CORPUS CLEAN-UP</b> .....	3880
<i>Martin Reynaert</i>	
<b>QUALITY ESTIMATION FOR SYNTHETIC PARALLEL DATA GENERATION</b> .....	3887
<i>Raphael Rubino, Antonio Toral, Nikola Ljubeic, Gema Ramrez-Snchez</i>	
<b>ADDING A THIRD LANGUAGE TO A LEXICAL RESOURCE DESCRIBING LEGAL TERMINOLOGY: THE ASSIGNMENT OF EQUIVALENTS</b> .....	N/A
<i>: Janine Pimentel</i>	
<b>AN OUT-OF-DOMAIN TEST SUITE FOR DEPENDENCY PARSING OF GERMAN</b> .....	3894
<i>Wolfgang Seeker, Jonas Kuhn</i>	
<b>REPRESENTING MULTILINGUAL DATA AS LINKED DATA: THE CASE OF BABELNET 2.0</b> .....	3902
<i>Maud Ehrmann, Francesco Cecconi, Daniele Vannella, John Philip McCrae, Philipp Cimiano, Roberto Navigli</i>	
<b>NOMAD: LINGUISTIC RESOURCES AND TOOLS AIMED AT POLICY FORMULATION AND VALIDATION</b> .....	3910
<i>George Kiomourtzis, George Giannakopoulos, Georgios Petasis, Pythagoras Karampiperis, Vangelis Karkaletsis</i>	
<b>ENCOMPASSING A SPECTRUM OF LT USERS IN THE CLARIN-DK INFRASTRUCTURE</b> .....	3917
<i>Lina Henriksen, Dorte Haltrup Hansen, Bente Maegaard, Bolette Sandford Pedersen, Claus Povlsen</i>	
<b>AUTOMATICALLY ENRICHING SPOKEN CORPORA WITH SYNTACTIC INFORMATION FOR LINGUISTIC STUDIES</b> .....	3924
<i>Alexis Nasr, Frederic Bechet, Benoit Favre, Thierry Bazillon, Jose Deulofeu, Andre Valli</i>	
<b>PROPA-L: A SEMANTIC FILTERING SERVICE FROM A LEXICAL NETWORK CREATED USING GAMES WITH A PURPOSE</b> .....	3929
<i>Mathieu Lafourcade, Karn Fort</i>	
<b>LESS IS MORE? TOWARDS A REDUCED INVENTORY OF CATEGORIES FOR TRAINING A PARSER FOR THE ITALIAN STANFORD DEPENDENCIES</b> .....	3935
<i>Maria Simi, Cristina Bosco, Simonetta Montemagni</i>	
<b>META-CLASSIFIERS EASILY IMPROVE COMMERCIAL SENTIMENT DETECTION TOOLS</b> .....	3943
<i>Mark Cieliebak, Oliver Drr, Fatih Uzdilli</i>	

<b>UNIXMAN CORPUS: A RESOURCE FOR LANGUAGE LEARNING IN THE UNIX DOMAIN</b> .....	3948
<i>Kyle Richardson, Jonas Kuhn</i>	
<b>GRAPAT: A TOOL FOR GRAPH ANNOTATIONS</b> .....	3953
<i>Jonathan Sonntag, Manfred Stede</i>	
<b>STANDARDISATION AND INTEROPERATION OF MORPHOSYNTACTIC AND SYNTACTIC ANNOTATION TOOLS FOR SPANISH AND THEIR ANNOTATIONS</b> .....	3958
<i>Antonio Pareja-Lora, Guillermo Cárcamo-Escorza, Alicia Ballesteros-Calvo</i>	
<b>A FRAMEWORK FOR COMPILING HIGH QUALITY KNOWLEDGE RESOURCES FROM RAW CORPORA</b> .....	3965
<i>Gongye Jin, Daisuke Kawahara, Sadao Kurohashi</i>	
<b>FUZZY V-MEASURE - AN EVALUATION METHOD FOR CLUSTER ANALYSES OF AMBIGUOUS DATA</b> .....	3971
<i>Jason Utt, Sylvia Springorum, Maximilian Köper, Sabine Schulte Im Walde</i>	
<b>CLUSTERING TWEETS USING WIKIPEDIA CONCEPTS</b> .....	3978
<i>Guoyu Tang, Yunqing Xia, Weizhi Wang, Raymond Lau, Fang Zheng</i>	
<b>THE TUTORBOT CORPUS ? A CORPUS FOR STUDYING TUTORING BEHAVIOUR IN MULTIPARTY FACE-TO-FACE SPOKEN DIALOGUE</b> .....	3984
<i>Maria Koutsombogera, Samer Al Moubayed, Bajibabu Bollepalli, Ahmed Hussen Abdelaziz, Martin Johansson, José David Aguas Lopes, Jekaterina Novikova, Catharine Oertel, Kalin Stefanov, Gül Varol</i>	
<b>TWEETCAT: A TOOL FOR BUILDING TWITTER CORPORA OF SMALLER LANGUAGES</b> .....	3990
<i>Nikola Ljubešić, Darja Fiser, Tomaz Erjavec</i>	
<b>HINDENCORP - HINDI-ENGLISH AND HINDI-ONLY CORPUS FOR MACHINE TRANSLATION</b> .....	3995
<i>Ondrej Bojar, Vojtech Diatka, Pavel Rychly, Pavel Stranak, Vit Suchomel, Aleš Tamchyna, Daniel Zeman</i>	
<b>COMBINING DEPENDENCY INFORMATION AND GENERALIZATION IN A PATTERN-BASED APPROACH TO THE CLASSIFICATION OF LEXICAL-SEMANTIC RELATION INSTANCES</b> .....	4001
<i>Silvia Neculescu, Sara Mendes, Núria Bel</i>	
<b>USING AUDIO BOOKS FOR TRAINING A TEXT-TO-SPEECH SYSTEM</b> .....	4009
<i>Aimilios Chalamandaris, Pirros Tsiakoulis, Sotiris Karabetos, Spyros Raptis</i>	
<b>USING A SLEDGEHAMMER TO CRACK A NUT? LEXICAL DIVERSITY AND EVENT COREFERENCE RESOLUTION</b> .....	4014
<i>Agata Cybulska, Piek Vossen</i>	
<b>CAWAC -- A WEB CORPUS OF CATALAN AND ITS APPLICATION TO LANGUAGE MODELING AND MACHINE TRANSLATION</b> .....	4022
<i>Nikola Ljubešić, Antonio Toral</i>	
<b>RECENT DEVELOPMENTS IN DEREKO</b> .....	4027
<i>Marc Kupietz, Harald Lüngen</i>	
<b>WHY CHINESE WEB-AS-CORPUS IS WACKY? OR: HOW BIG DATA IS KILLING CHINESE CORPUS LINGUISTICS</b> .....	4035
<i>Shu-Kai Hsieh</i>	
<b>AUTOMATIC ACQUISITION OF URDU NOUNS (ALONG WITH GENDER AND IRREGULAR PLURALS)</b> .....	4039
<i>Tafseer Ahmed Khan</i>	
<b>RE-USING AN ARGUMENT CORPUS TO AID IN THE CURATION OF SOCIAL MEDIA COLLECTIONS</b> .....	4044
<i>Clare Llewellyn, Claire Grover, Jon Oberlander, Ewan Klein</i>	
<b>BILLIONS OF PARALLEL WORDS FOR FREE: BUILDING AND USING THE EU BOOKSHOP CORPUS</b> .....	4051
<i>Raivis Skadinš, Jörg Tiedemann, Roberts Rozis, Daiga Deksnė</i>	
<b>GENERATING POLARITY LEXICONS WITH WORDNET PROPAGATION IN 5 LANGUAGES</b> .....	4057
<i>Isa Maks, Ruben Izquierdo, Francesca Frontini, Rodrigo Aggeri, Piek Vossen, Andoni Azpeitia</i>	
<b>ONLINE OPTIMISATION OF LOG-LINEAR WEIGHTS IN INTERACTIVE MACHINE TRANSLATION</b> .....	4064
<i>Mara China-Rios, Germán Sanchis Trilles, Daniel Daniel Ortiz-Martínez, Francisco Casacuberta</i>	
<b>EXTENDING HEIDELTIME FOR TEMPORAL EXPRESSIONS REFERRING TO HISTORIC DATES</b> .....	4068
<i>Jannik Strotgen, Thomas Bogel, Julian Zell, Ayser Armiti, Tran Van Canh, Michael Gertz</i>	
<b>THE USAGE REVIEW CORPUS FOR FINE GRAINED MULTI LINGUAL OPINION ANALYSIS</b> .....	4076
<i>Roman Klinger, Philipp Cimiano</i>	

<b>AN EXERCISE IN REUSE OF RESOURCES: ADAPTING GENERAL DISCOURSE COREFERENCE RESOLUTION FOR DETECTING LEXICAL CHAINS IN PATENT DOCUMENTATION</b> .....	4084
<i>Nadjet Bouayad-Agha, Alicia Burga, Gerard Casamayor, Joan Codina, Rogelio Nazar, Leo Wanner</i>	
<b>VOAR: A VISUAL AND INTEGRATED ONTOLOGY ALIGNMENT ENVIRONMENT</b> .....	4092
<i>Bernardo Severo, Cassia Trojahn, Renata Vieira</i>	
<b>A 500 MILLION WORD POS-TAGGED ICELANDIC CORPUS</b> .....	4099
<i>Thomas Eckart, Erla Hallsteinsdóttir, Sigrún Helgadóttir, Uwe Quasthoff, Dirk Goldhahn</i>	
<b>A COLLECTION OF SCHOLARLY BOOK REVIEWS FROM THE PLATFORMS OF ELECTRONIC SOURCES IN HUMANITIES AND SOCIAL SCIENCES OPENEDITION.ORG</b> .....	4104
<i>Chahinez Benkoussas, Hussam Hamdan, Patrice Bellot, Frédéric Béchet, Elodie Faath</i>	
<b>RAPID DEPLOYMENT OF PHRASE STRUCTURE PARSING FOR RELATED LANGUAGES: A CASE STUDY OF INSULAR SCANDINAVIAN</b> .....	4110
<i>Anton Karl Ingason, Hrafn Loftsson, Eiríkur Rognvaldsson, Einar Freyr Sigurðsson, Joel C. Wallenberg</i>	
<b>N<sup>3</sup> - A COLLECTION OF DATASETS FOR NAMED ENTITY RECOGNITION AND DISAMBIGUATION IN THE NLP INTERCHANGE FORMAT</b> .....	4115
<i>Michael Röder, Ricardo Usbeck, Sebastian Hellmann, Daniel Gerber, Andreas Both</i>	
<b>ASSESSMENT OF NON-NATIVE PROSODY FOR SPANISH AS L2 USING QUANTITATIVE SCORES AND PERCEPTUAL EVALUATION</b> .....	4120
<i>Valentín Cardeñoso-Payo, César González-Ferreras, David Escudero</i>	
<b>EXPLOITING THE LARGE-SCALE GERMAN BROADCAST CORPUS TO BOOST THE FRAUNHOFER IAIS SPEECH RECOGNITION SYSTEM</b> .....	4126
<i>Michael Stadtschnitzer, Jochen Schwenninger, Daniel Stein, Joachim Koehler</i>	
<b>SUMMARIZING NEWS CLUSTERS ON THE BASIS OF THEMATIC CHAINS</b> .....	4130
<i>Natalia Loukachevitch, Aleksey Alekseev</i>	
<b>BECAUSE SIZE DOES MATTER: THE HAMBURG DEPENDENCY TREEBANK</b> .....	4138
<i>Kilian A. Foth, Arne Köhn, Niels Beuck, Wolfgang Menzel</i>	
<b>DISCOVERING THE ITALIAN LITERATURE: INTERACTIVE ACCESS TO AUDIO INDEXED TEXT RESOURCES</b> .....	4146
<i>Vincenzo Galatà, Alberto Benin, Piero Così, Giuseppe Riccardo Leone, Giulio Paci, Giacomo Sommovilla, Fabio Tesser</i>	
<b>ENABLING LANGUAGE RESOURCES TO EXPOSE TRANSLATIONS AS LINKED DATA ON THE WEB</b> .....	4151
<i>Jorge Gracia, Elena Montiel-Ponsoda, Daniel Vila-Suero, Guadalupe Aguado-De-Cea</i>	
<b>PIVOT-BASED MULTILINGUAL DICTIONARY BUILDING USING WIKTIONARY</b> .....	4156
<i>Judit Acs</i>	
<b>EXPLORING THE UTILITY OF COREFERENCE CHAINS FOR IMPROVED IDENTIFICATION OF PERSONAL NAMES</b> .....	4161
<i>Andrea Glaser, Jonas Kuhn</i>	
<b>MULTILINGUAL EXTENDED WORDNET KNOWLEDGE BASE: SEMANTIC PARSING AND TRANSLATION OF GLOSSES</b> .....	4169
<i>Tatiana Erekhinskaya, Meghana Satpute, Dan Moldovan</i>	
<b>RELATION INFERENCE IN LEXICAL NETWORKS ... WITH REFINEMENTS</b> .....	4174
<i>Manel Zarrouk, Mathieu Lafourcade</i>	
<b>A MULTIMODAL DATASET FOR DECEPTION DETECTION</b> .....	4180
<i>Veronica Perez-Rosas, Rada Mihalcea, Alexis Narvaez, Mihai Burzo</i>	
<b>C-PHONOGENRE: A 7-HOURS CORPUS OF 7 SPEAKING STYLES IN FRENCH: RELATIONS BETWEEN SITUATIONAL FEATURES AND PROSODIC PROPERTIES</b> .....	4185
<i>Jean-Philippe Goldman, Tea Prsir, Antoine Auchlin</i>	
<b>THE AMARA CORPUS: BUILDING PARALLEL LANGUAGE RESOURCES FOR THE EDUCATIONAL DOMAIN</b> .....	4189
<i>Ahmed Abdelali, Francisco Guzman, Hassan Sajjad, Stephan Vogel</i>	
<b>DEPENDENCY PARSING REPRESENTATION EFFECTS ON THE ACCURACY OF SEMANTIC APPLICATIONS - AN EXAMPLE OF AN INFLECTIVE LANGUAGE</b> .....	4196
<i>Lauma Pretkalnina, Arturs Znotins, Laura Rituma, Didzis Goško</i>	
<b>CO-CLUSTERING OF BILINGUAL DATASETS AS A MEAN FOR ASSISTING THE CONSTRUCTION OF THEMATIC BILINGUAL COMPARABLE CORPORA</b> .....	4204
<i>Guiyao Ke, Pierre-Francois Marteau</i>	
<b>RULED-BASED, INTERLINGUAL MOTIVATED MAPPING OF PLWORDNET ONTO SUMO ONTOLOGY</b> .....	4212
<i>Pawel Kedzia, Maciej Piasecki</i>	

<b>EXTENDING THE COVERAGE OF A MWE DATABASE FOR PERSIAN CPS EXPLOITING VALENCY ALTERNATIONS</b> .....	4220
<i>Pollet Samvelian, Pegah Faghiri, Sarra El Ayari</i>	
<b>FINDING A TRADEOFF BETWEEN ACCURACY AND RATER'S WORKLOAD IN GRADING CLUSTERED SHORT ANSWERS</b> .....	4224
<i>Andrea Horbach, Alexis Palmer, Magdalena Wolska</i>	
<b>MACROSYNTACTIC SEGMENTERS OF A FRENCH SPOKEN CORPUS</b> .....	4232
<i>Ilaine Wang, Sylvain Kahane, Isabelle Tellier</i>	
<b>VULNERABILITY IN ACQUISITION, LANGUAGE IMPAIRMENTS IN DUTCH: CREATING A VALID DATA ARCHIVE</b> .....	4238
<i>Jetske Klatter, Roeland Van Hout, Henk Van Den Heuvel, Paula Fikkert, Anne Baker, Jan De Jong, Frank Wijnen, Eric Sanders, Paul Trilsbeek</i>	
<b>THE EXTENDED DIRNDL CORPUS AS A RESOURCE FOR COREFERENCE AND BRIDGING RESOLUTION</b> .....	4246
<i>Anders Björkelund, Kerstin Eckart, Arndt Riestler, Nadja Schauffler, Katrin Schweitzer</i>	
<b>A FLEXIBLE LANGUAGE LEARNING PLATFORM BASED ON LANGUAGE RESOURCES AND WEB SERVICES</b> .....	4253
<i>Elena Volodina, Ildikó Pilán, Lars Borin, Therese Lindström Tiedemann</i>	
<b>TOWARDS INTEROPERABLE DISCOURSE ANNOTATION. DISCOURSE FEATURES IN THE ONTOLOGIES OF LINGUISTIC ANNOTATION</b> .....	4259
<i>Christian Chiarcos</i>	
<b>MORPHOLOGICAL PARSING OF SWAHILI USING CROWDSOURCED LEXICAL RESOURCES</b> .....	4268
<i>Patrick Littell, Kaitlyn Price, Lori Levin</i>	
<b>IMPROVING ENTITY LINKING USING SURFACE FORM REFINEMENT</b> .....	4275
<i>Eric Charton, Marie-Jean Meurs, Ludovic Jean-Louis, Michel Gagnon</i>	
<b>ELRA'S CONSOLIDATED SERVICES FOR THE HLT COMMUNITY</b> .....	4282
<i>Victoria Arranz, Khalid Choukri, Valérie Mapelli, Hélène Mazo</i>	
<b>SINGLE CLASSIFIER APPROACH FOR VERB SENSE DISAMBIGUATION BASED ON GENERALIZED FEATURES</b> .....	4288
<i>Daisuke Kawahara, Martha Palmer</i>	
<b>EXTRACTING SEMANTIC RELATIONS FROM PORTUGUESE CORPORA USING LEXICAL-SYNTACTIC PATTERNS</b> .....	4292
<i>Raquel Amaro</i>	
<b>LINKEDHEALTHANSWERS: TOWARDS LINKED DATA-DRIVEN QUESTION ANSWERING FOR THE HEALTH CARE DOMAIN</b> .....	4297
<i>Artem Ostankov, Florian Rohrbach, Ulli Waltinger</i>	
<b>AN ANALYSIS OF AMBIGUITY IN WORD SENSE ANNOTATIONS</b> .....	4305
<i>David Jurgens</i>	
<b>VOLIP: A CORPUS OF SPOKEN ITALIAN AND A VIRTUOUS EXAMPLE OF REUSE OF LINGUISTIC RESOURCES</b> .....	4312
<i>Iolanda Alfano, Francesco Cutugno, Aurelio De Rosa, Claudio Iacobini, Renata Savy, Miriam Voghera</i>	
<b>CHASING THE PERFECT SPLITTER: A COMPARISON OF DIFFERENT COMPOUND SPLITTING TOOLS</b> .....	4317
<i>Carla Parra Escartín</i>	
<b>A COMPARISON OF MT ERRORS AND ESL ERRORS</b> .....	4325
<i>Homa B. Hashemi, Rebecca Hwa</i>	
<b>NEW FUNCTIONS FOR A MULTIPURPOSE MULTIMODAL TOOL FOR PHONETIC AND LINGUISTIC ANALYSIS OF VERY LARGE SPEECH CORPORA</b> .....	4330
<i>Philippe Martin</i>	
<b>HOT TOPICS AND SCHISMS IN NLP: COMMUNITY AND TREND ANALYSIS WITH SAFFRON ON ACL AND LREC PROCEEDINGS</b> .....	4334
<i>Paul Buitelaar, Georgeta Bordea, Barry Coughlan</i>	
<b>THE AMERICAN LOCAL NEWS CORPUS</b> .....	4340
<i>Ann Irvine, Joshua Langfus, Chris Callison-Burch</i>	
<b>HAMLEDT 2.0: THIRTY DEPENDENCY TREEBANKS STANFORDIZED</b> .....	4344
<i>Rudolf Rosa, Jan Masek, David Marecek, Martin Popel, Daniel Zeman, Zdenek Zabokrtský</i>	
<b>BUILDING THE SENSE-TAGGED MULTILINGUAL PARALLEL CORPUS</b> .....	4352
<i>Shan Wang, Francis Bond</i>	
<b>MULTILINGUAL CORPORA WITH COREFERENTIAL ANNOTATION OF PERSON ENTITIES</b> .....	4359
<i>Marcos Garcia, Pablo Gamallo</i>	



<b>SANA: A LARGE SCALE MULTI-GENRE, MULTI-DIALECT LEXICON FOR ARABIC SUBJECTIVITY AND SENTIMENT ANALYSIS</b> .....	4364
<i>Muhammad Abdul-Mageed, Mona Diab</i>	
<b>EVALUATION OF TECHNOLOGY TERM RECOGNITION WITH RANDOM INDEXING</b> .....	4372
<i>Behrang Zadeh, Siegfried Handschuh</i>	
<b>OPTIMIZING A DISTRIBUTIONAL SEMANTIC MODEL FOR THE PREDICTION OF GERMAN PARTICLE VERB COMPOSITIONALITY</b> .....	4378
<i>Stefan Bott, Sabine Schulte Im Walde</i>	
<b>A HINDI-ENGLISH CODE-SWITCHING CORPUS</b> .....	4386
<i>Anik Dey, Pascale Fung</i>	
<b>THE LANGUAGE APPLICATION GRID</b> .....	4390
<i>Nancy Ide, James Pustejovsky, Christopher Cieri, Eric Nyberg, Di Wang, Keith Suderman, Marc Verhagen, Jonathan Wright</i>	
<b>DISMO: A MORPHOSYNTACTIC, DISFLUENCY AND MULTI-WORD UNIT ANNOTATOR. AN EVALUATION ON A CORPUS OF FRENCH SPONTANEOUS AND READ SPEECH</b> .....	4399
<i>George Christodoulides, Mathieu Avanzi, Jean-Philippe Goldman</i>	
<b>INTEGRATION OF WORKFLOW AND PIPELINE FOR LANGUAGE SERVICE COMPOSITION</b> .....	4405
<i>Trang Mai Xuan, Yohei Murakami, Donghui Lin, Toru Ishida</i>	
<b>SEGMENTATION EVALUATION METRICS, A COMPARISON GROUNDED ON PROSODIC AND DISCOURSE UNITS</b> .....	4413
<i>Klim Peshkov, Laurent Prévot</i>	
<b>KOKO: AN L1 LEARNER CORPUS FOR GERMAN</b> .....	4418
<i>Andrea Abel, Aivars Glaznieks, Lionel Nicolas, Egon Stemle</i>	
<b>IMPROVING EVALUATION OF ENGLISH-CZECH MT THROUGH PARAPHRASING</b> .....	4426
<i>Petra Barancikova, Rudolf Rosa, Ales Tamchyna</i>	
<b>DISCLOSE MODELS, HIDE THE DATA - HOW TO MAKE USE OF CONFIDENTIAL CORPORA WITHOUT SEEING SENSITIVE RAW DATA</b> .....	4432
<i>Erik Faessler, Johannes Hellrich, Udo Hahn</i>	
<b>WHEN transliteration met crowdsourcing : AN EMPIRICAL STUDY OF transliteration VIA crowdsourcing USING EFFICIENT, NON-REDUNDANT AND FAIR QUALITY CONTROL</b> .....	4440
<i>Mitesh M. Khapra, Ananthkrishnan Ramanathan, Anoop Kunchukuttan, Karthik Visweswariah, Pushpak Bhattacharyya</i>	
<b>OPEN PHILOLOGY AT THE UNIVERSITY OF LEIPZIG</b> .....	4447
<i>Frederik Baumgardt, Giuseppe Celano, Gregory R. Crane, Stella Dee, Maryam Foradi, Emily Franzini, Greta Franzini, Monica Lent, Maria Moritz, Simona Stoyanova</i>	
<b>A MODULAR SYSTEM FOR RULE-BASED TEXT CATEGORISATION</b> .....	N/A
<i>: Marco Del Tredici, Malvina Nissim</i>	
<b>DCEP -DIGITAL CORPUS OF THE EUROPEAN PARLIAMENT</b> .....	4451
<i>Najeh Hajlaoui, David Kolovratnik, Jaakko Väyrynen, Ralf Steinberger, Daniel Varga</i>	
<b>FACING THE IDENTIFICATION PROBLEM IN LANGUAGE-RELATED SCIENTIFIC DATA ANALYSIS</b> .....	4459
<i>Joseph Mariani, Christopher Cieri, Gil Francopoulo, Patrick Paroubek, Marine Delaborde</i>	
<b>SMILE AND LAUGHTER IN HUMAN-MACHINE INTERACTION: A STUDY OF ENGAGEMENT</b> .....	4466
<i>Mariette Soury, Laurence Devillers</i>	
<b>EXPLOITING NETWORKS IN LAW</b> .....	4471
<i>Livio Robaldo, Guido Boella, Luigi Di Caro, Andrea Violato</i>	
<b>UTILIZING CONSTITUENT STRUCTURE FOR COMPOUND ANALYSIS</b> .....	4476
<i>Kristín Bjarnadóttir, Jón Daðason</i>	
<b>LARGE SCALE ARABIC ERROR ANNOTATION: GUIDELINES AND FRAMEWORK</b> .....	4481
<i>Wajdi Zaghouni, Behrang Mohit, Nizar Habash, Ossama Obeid, Nadi Tomeh, Alla Rozovskaya, Noura Farra, Sarah Alkuhlani, Kemal Oflazer</i>	
<b>EL-WOZ: A CLIENT-SERVER WIZARD-OF-OZ INTERFACE</b> .....	4489
<i>Thomas Pellegrini, Vahid Hedayati, Angela Costa</i>	
<b>PARSING CHINESE SYNTHETIC WORDS WITH A CHARACTER-BASED DEPENDENCY MODEL</b> .....	4493
<i>Fei Cheng, Kevin Duh, Yuji Matsumoto</i>	
<b>ETER : A NEW METRIC FOR THE EVALUATION OF HIERARCHICAL NAMED ENTITY RECOGNITION</b> .....	4499
<i>Mohamed Ben Jannet, Martine Adda-Decker, Olivier Galibert, Juliette Kahn, Sophie Rosset</i>	
<b>DETECTING SUBEVENT STRUCTURE FOR EVENT COREFERENCE RESOLUTION</b> .....	4507
<i>Jun Araki, Zhengzhong Liu, Eduard Hovy, Teruko Mitamura</i>	

<b>AN EFFICIENT AND USER-FRIENDLY TOOL FOR MACHINE TRANSLATION QUALITY ESTIMATION .....</b>	<b>4513</b>
<i>Kashif Shah, Marco Turchi, Lucia Specia</i>	
<b>RESOURCE CREATION AND EVALUATION FOR MULTILINGUAL SENTIMENT ANALYSIS IN SOCIAL MEDIA TEXTS .....</b>	<b>4518</b>
<i>Alexandra Balahur, Marco Turchi, Ralf Steinberger, Jose Manuel Perea-Ortega, Guillaume Jacquet, Dilek Kucuk, Vanni Zavarella, Adil El Ghali</i>	
<b>NAMED ENTITY TAGGING A VERY LARGE UNBALANCED CORPUS: TRAINING AND EVALUATING NE CLASSIFIERS .....</b>	<b>4523</b>
<i>Joachim Bingel, Thomas Haider</i>	
<b>VALIDATION ISSUES INDUCED BY AN AUTOMATIC PRE-ANNOTATION MECHANISM IN THE BUILDING OF NON-PROJECTIVE DEPENDENCY TREEBANKS .....</b>	<b>4529</b>
<i>Ophélie Lacroix, Denis Béchet</i>	
<b>MAT: A TOOL FOR L2 PRONUNCIATION ERRORS ANNOTATION .....</b>	<b>4534</b>
<i>Renlong Ai, Marcela Charfuelan</i>	
<b>WORD SEMANTIC SIMILARITY FOR MORPHOLOGICALLY RICH LANGUAGES .....</b>	<b>4538</b>
<i>Kalliopi Zervanou, Elias Iosif, Alexandros Potamianos</i>	
<b>LEXTERM MANAGER: DESIGN FOR AN INTEGRATED LEXICOGRAPHY AND TERMINOLOGY SYSTEM .....</b>	<b>4545</b>
<i>Joshua Elliot, Logan Kearsley, Jason Housley, Alan Melby</i>	
<b>FOCUSING ANNOTATION FOR SEMANTIC ROLE LABELING .....</b>	<b>4551</b>
<i>Daniel Peterson, Martha Palmer, Shumin Wu</i>	
<b>OFF-ROAD LAF: ENCODING AND PROCESSING ANNOTATIONS IN NLP WORKFLOWS .....</b>	<b>4556</b>
<i>Emanuele Lapponi, Erik Velldal, Stephan Oepen, Rune Lain Knudsen</i>	
<b>DEVELOPING A FRAMEWORK FOR DESCRIBING RELATIONS AMONG LANGUAGE RESOURCES .....</b>	<b>4563</b>
<i>Penny Labropoulou, Christopher Cieri, Maria Gavrilidou</i>	
<b>EVALUATING WEB-AS-CORPUS TOPICAL DOCUMENT RETRIEVAL WITH AN INDEX OF THE OPENDIRECTORY .....</b>	<b>4571</b>
<i>Clément De Groc, Xavier Tannier</i>	
<b>WORD ALIGNMENT-BASED REORDERING OF SOURCE CHUNKS IN PB-SMT .....</b>	<b>4575</b>
<i>Santanu Pal, Sudip Kumar Naskar, Sivaji Bandyopadhyay</i>	
<b>TAALPORTAAL: AN ONLINE GRAMMAR OF DUTCH AND FRISIAN .....</b>	<b>4582</b>
<i>Frank Landsbergen, Carole Tiberius, Roderik Derrison</i>	
<b>A FRAMEWORK FOR PUBLIC HEALTH SURVEILLANCE .....</b>	<b>4587</b>
<i>Andrew Yates, Jon Parker, Nazli Goharian, Ophir Frieder</i>	
<b>MULTILINGUAL TEST SETS FOR MACHINE TRANSLATION OF SEARCH QUERIES FOR CROSS-LINGUAL INFORMATION RETRIEVAL IN THE MEDICAL DOMAIN .....</b>	<b>4595</b>
<i>Zdenka Uresova, Jan Hajic, Pavel Pecina, Ondrej Dusek</i>	
<b>A TOOL SUITE FOR CREATING QUESTION ANSWERING BENCHMARKS .....</b>	<b>4599</b>
<i>Axel-Cyrille Ngonga Ngomo, Norman Heino, René Speck, Prodromos Malakasiotis</i>	
<b>THEMATIC COHESION: MEASURING TERMS DISCRIMINATORY POWER TOWARD THEMES .....</b>	<b>4607</b>
<i>Clément De Groc, Xavier Tannier, Claude De Loupy</i>	
<b>TERMINOLOGY RESOURCES AND TERMINOLOGY WORK BENEFIT FROM CLOUD SERVICES .....</b>	<b>4612</b>
<i>Tatiana Gornostay, Andrejs Vasiljevs</i>	
<b>BIDIRECTIONAL CONVERTER BETWEEN SYNTACTIC ANNOTATIONS : FROM FRENCH TREEBANK DEPENDENCIES TO PASSAGE ANNOTATIONS, AND BACK .....</b>	<b>4618</b>
<i>Munshi Asadullah, Patrick Paroubek, Anne Vilnat</i>	
<b>VARCLASS: AN OPEN-SOURCE LANGUAGE IDENTIFICATION TOOL FOR LANGUAGE VARIETIES .....</b>	<b>4624</b>
<i>Marcos Zampieri, Binyam Gebre</i>	
<b>RECSA: RESOURCE FOR EVALUATING CROSS-LINGUAL SEMANTIC ANNOTATION .....</b>	<b>4628</b>
<i>Achim Rettinger, Lei Zhang, Daša Berovic, Danijela Merkle, Matea Srebacic, Marko Tadic</i>	
<b>Author Index</b>	