

CLiC-it 2017

The Fourth Italian Conference on Computational Linguistics

Proceedings of the Conference

December 11-13, 2017

Copyright ©2017 for the individual papers by the papers' authors.
Copyright ©2017 for the volume as a collection by its editors.
This volume and its papers are published under the Creative Commons License Attribution 4.0 International (CC BY 4.0).

These proceedings have been published in the CEUR Workshop Proceedings series.
The original papers are available at: <https://ceur-ws.org/Vol-2006>.
The papers are mirrored in the ACL Anthology.

ISBN 978-88-99982-94-2

Table of Contents

<i>Preface</i>	
Roberto Basili, Malvina Nissim and Giorgio Satta	1
<i>Searching for General Models that Learn Compositionally</i>	
Marco Baroni	6
<i>Doing Stuff with Long Short Term Memory networks</i>	
Yoav Goldberg	7
<i>Computational Sociolinguistics – An Emerging Partnership</i>	
Rada Mihalcea	8
<i>I Verbi Neologici nell’Italiano del Web: Comportamento Sintattico e Selezione dell’Ausiliare</i>	
Matteo Amore	9
<i>Predicting Land Use of Italian Cities using Structural Semantic Models</i>	
Gianni Barlacchi, Bruno Lepri and Alessandro Moschitti	15
<i>Predicting Controversial News Using Facebook Reactions</i>	
Angelo Basile, Tommaso Caselli and Malvina Nissim	20
<i>Bi-directional LSTM-CNNs-CRF for Italian Sequence Labeling</i>	
Pierpaolo Basile, Giovanni Semeraro and Pierluigi Cassotti	26
<i>Monitoring Adolescents’ Distress using Social Web data as a Source: the InsideOut Project</i>	
Roberto Basili, Valentina Bellomaria, Niels Jonas Bugge, Danilo Croce, Francesco De Michele, Federico Fiori Nastro, Paolo Fiori Nastro, Chantal Michel, Stefanie Schmidt and Frauke Schultze-Lutter	32
<i>Gimme the Usual"- How Handling of Pragmatics Improves Chatbots</i>	
Alessia Bianchini, Francesco Tarasconi, Raffaella Ventaglio and Mariafrancesca Guadalupi	38
<i>Deep-learning the Ropes: Modeling Idiomaticity with Neural Networks</i>	
Yuri Bizzoni, Marco S.G. Senaldi and Alessandro Lenci	44
<i>Neural Sentiment Analysis for a Real-World Application</i>	
Daniele Bonadiman, Giuseppe Castellucci, Andrea Favalli, Raniero Romagnoli and Alessandro Moschitti	50
<i>Emo2Val: Inferring Valence Scores from fine-grained Emotion Values</i>	
Alessandro Bondielli, Lucia Passaro and Alessandro Lenci	56
<i>Toward a Bilingual Lexical Database on Connectives: Exploiting a German/Italian Parallel Corpus</i>	
Peter Bourgonje, Yulia Grishina and Manfred Stede	61
<i>Developing a Large Scale FrameNet for Italian: the IFrameNet Experience</i>	
Silvia Brambilla, Danilo Croce, Fabio Tamburini and Roberto Basili	67
<i>-io Nouns through the Ages. Analysing Latin Morphological Productivity with Lemlat</i>	
Marco Budassi, Eleonora Litta and Marco Passarotti	73
<i>Gender Stereotypes in Film Language: a Corpus-assisted Analysis</i>	
Lucia Busso and Gianmarco Vignozzi	79

<i>How Deepis Learning Word Inflection?</i>	
Franco Alberto Cardillo, Marcello Ferro, Claudia Marzi and Vito Pirrelli	85
<i>Tree LSTMs for Learning Sentence Representations</i>	
Héctor Cerezo Costas, Manuela I. Martín-Vicente and Francisco J. Gonzalez-Castano	91
<i>Irony Detection: from the Twittersphere to the News Space</i>	
Alessandra Cervone, Evgeny Stepanov, Fabio Celli and Giuseppe Riccardi	97
<i>Phase-based Minimalist Parsing and Complexity in Non-local Dependencies</i>	
Cristiano Chesi	103
<i>TWITTIRO': a Social Media Corpus with a Multi-layered Annotation for Irony</i>	
Alessandra Teresa Cignarella, Cristina Bosco and Viviana Patti	109
<i>Identifying Predictive Features for Textual Genre Classification: the Key Role of Syntax</i>	
Andrea Cimino, Martijn Wieling, Felice Dell'Orletta, Simonetta Montemagni and Giulia Venturi	
115	
<i>CorAIt – A Non-native Speech Database for Italian</i>	
Claudia Roberta Combei	121
<i>Dealing with Italian Adjectives in Noun Phrase: a Study Oriented to Natural Language Generation</i>	
Giorgia Conte, Cristina Bosco and Alessandro Mazzei	127
<i>Variabili Rilevanti nella Rappresentazione delle Parole nel Lessico Mentale: Dati Psicolinguistici da una Banca-dati di Nomi e Verbi Italiani</i>	
Maria De Martino, Azzurra Mancuso and Alessandro Laudanna	133
<i>Tagging Semantic Types for Verb Argument Positions</i>	
Francesca Della Moretta, Anna Feltracco, Elisabetta Jezek and Bernardo Magnini	139
<i>Linguistic Features and Newsworthiness: an Analysis of News style</i>	
Maria Pia Di Buono and Jan Snajder	144
<i>Can Monolingual Embeddings Improve Neural Machine Translation?</i>	
Mattia Antonino Di Gangi and Marcello Federico	149
<i>Distributed Processes for Spoken Questions and Commands Understanding</i>	
Dario Di Mauro, Antonio Origlia and Francesco Cutugno	155
<i>A Reproducible Approach with R Markdown to Automatic Classification of Medical Certificates in French</i>	
Giorgio Maria Di Nunzio, Federica Beghini, Federica Vezzani and Geneviève Henrot	161
<i>Contrast-Ita Bank: A corpus for Italian Annotated with Discourse Contrast Relations</i>	
Anna Feltracco, Bernardo Magnini and Elisabetta Jezek	167
<i>Ironic Gestures and Tones in Twitter</i>	
Simona Frenda	173
<i>Emerging Sentiment Language Model for Emotion Detection</i>	
Anastasia Giachanou, Francisco Rangel, Fabio Crestani and Paolo Rosso	179
<i>Stylometry in Computer-Assisted Translation: Experiments on the Babylonian Talmud</i>	
Emiliano Giovannetti, Davide Albanesi, Andrea Bellandi, David Dattilo and Felice Dell'Orletta	
185	

<i>Towards an Italian Lexicon for Polarity Classification (polarITA): a Comparative Analysis of Lexical Resources for Sentiment Analysis</i>	
Delia Irazú Hernández Farías, Irene Laganà, Viviana Patti and Cristina Bosco	191
<i>Multilingual Neural Machine Translation for Low Resource Languages</i>	
Surafel Melaku Lakew, Mattia Antonino Di Gangi and Marcello Federico	197
<i>Stem and Fragment Priming on Verbal Forms of Italian</i>	
Alessandro Laudanna and Giulia Bracco	203
<i>Metadata Annotation for Dramatic Texts</i>	
Vincenzo Lombardo, Rossana Damiano and Antonio Pizzo	209
<i>Deep Learning for Automatic Image Captioning in Poor Training Conditions</i>	
Caterina Masotti, Danilo Croce and Roberto Basili	215
<i>E Pluribus Unum. Representing Compounding in a Derivational Lexicon of Latin</i>	
Silvia Micheli and Eleonora Litta	220
<i>Sanremo's Winner Is... Category-driven Selection Strategies for Active Learning</i>	
Anne-Lyse Minard, Manuela Speranza, Mohammed R. H. Qwaider and Bernardo Magnini . .	225
<i>Dalla Word Sense Disambiguation alla Sintassi: il Problema dell'Articolo Partitivo in Italiano</i>	
Ignazio Mauro Mirto and Emanuele Cipolla	231
<i>PARSEME-It Corpus An annotated Corpus of Verbal Multiword Expressions in Italian</i>	
Johanna Monti, Maria Pia Di Buono and Federico Sangati	236
<i>MODAL: A Multilingual Corpus Annotated for Modality</i>	
Malvina Nissim and Paola Pietrandrea	242
<i>AHyDA: Automatic Hypernym Detection with Feature Augmentation</i>	
Ludovica Pannitto, Lavinia Salicchi and Alessandro Lenci	248
<i>INFORMed PA: A NER for the Italian Public Administration Domain</i>	
Lucia Passaro, Alessandro Lenci and Anna Gabbolini	254
<i>Mining Offensive Language on Social Media</i>	
Serena Pelosi, Alessandro Maisto, Pierluigi Vitale and Simonetta Vietri	260
<i>Automatic Evaluation of Employee Satisfaction</i>	
Marco Piersanti, Giulia Brandetti and Pierluigi Failla	265
<i>Hate Speech Annotation: Analysis of an Italian Twitter Corpus</i>	
Fabio Poletto, Marco Stranisci, Manuela Sanguinetti, Viviana Patti and Cristina Bosco	271
<i>PoS Taggers in the Wild: A Case Study with Swiss Italian Student Essays</i>	
Daniele Puccinelli, Silvia Demartini, Aris Piatti, Sara Giulivi, Luca Cignetti and Simone Fornara	277
<i>Find Problems before They Find You with AnnotatorPro's Monitoring Functionalities</i>	
Mohammed R. H. Qwaider, Anne-Lyse Minard, Manuela Speranza and Bernardo Magnini . .	283
<i>Evaluating a Rule Based Strategy to Map IMAGACT and T-PAS</i>	
Andrea Amelio Ravelli, Lorenzo Gregori and Anna Feltracco	289
<i>Domain-specific Named Entity Disambiguation in Historical Memoirs</i>	
Marco Rovera, Federico Nanni, Simone Paolo Ponzetto and Anna Goy	295

<i>Toward a Treebank Collecting German Aesthetic Writings of the Late 18th Century</i> Alessio Salomoni	300
<i>Assessing the Use of Terminology in Phrase-Based Statistical Machine Translation for Academic Course Catalogues Translation</i> Randy Scansani, Marcello Federico and Luisa Bentivogli	306
<i>A little bit of bella pianura: Detecting Code-Mixing in Historical English Travel Writing</i> Rachele Sprugnoli, Sara Tonelli, Giovanni Moretti and Stefano Menini	312
<i>PARAD-it: Eliciting Italian Paradigmatic Relations with Crowdsourcing</i> Irene Sucameli and Alessandro Lenci	318
<i>The Impact of Phrases on Italian Lexical Simplification</i> Sara Tonelli, Alessio Palmero Aprosio and Marco Mazzon	324
<i>Analysis of Italian Word Embeddings</i> Rocco Tripodi and Stefano Li Pira	329
<i>Applicazione di un Metodo Attribuzionistico Quantitativo alla Monodia Liturgica Medievale</i> Francesco Unguendoli, Giampaolo Cristadoro and Marco Beghelli	335
<i>Commercial Applications through Community Question Answering Technology</i> Antonio Uva, Valerio Storch, Casimiro Carrino, Ugo Di Iorio and Alessandro Moschitti	341