

CLiC-it 2018

The Fifth Italian Conference on Computational Linguistics

Proceedings of the Conference

December 10-12, 2018

Copyright ©2018 for the individual papers by the papers' authors.
Copyright ©2018 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-2253>.
The papers are mirrored in the ACL Anthology.

ISBN 978-88-31978-41-5

Table of Contents

<i>Preface</i>	
Elena Cabrio, Alessandro Mazzei and Fabio Tamburini	1
<i>Computational Semantics in Neural Times</i>	
Johan Bos	6
<i>Disentangling the Thoughts: Latest News in Computational Argumentation</i>	
Iryna Gurevych	7
<i>A Distributional Study of Negated Adjectives and Antonyms</i>	
Laura Aina, Raffaella Bernardi and Raquel Fernández	8
<i>PRET: Prerequisite-Enriched Terminology. A Case Study on Educational Texts.</i>	
Chiara Alzetta, Frosina Koceva, Samuele Passalacqua, Ilaria Torre and Giovanni Adorni	15
<i>Distributional Analysis of Verbal Neologisms: Task Definition and Dataset Construction</i>	
Matteo Amore, Stephen McGregor and Elisabetta Jezek	22
<i>Parsing Italian Texts Together is Better Than Parsing Them Alone!</i>	
Oronzo Antonelli and Fabio Tamburini	28
<i>Buon appetito!"- Analyzing Happiness in Italian Tweets</i>	
Pierpaolo Basile and Nicole Novielli	35
<i>Long-term Social Media Data Collection at the University of Turin</i>	
Valerio Basile, Mirko Lai and Manuela Sanguinetti	41
<i>Neural Surface Realization for Italian</i>	
Valerio Basile and Alessandro Mazzei	47
<i>Hurtlex: A Multilingual Lexicon of Words to Hurt</i>	
Elisa Bassignana, Valerio Basile and Viviana Patti	52
<i>CoreNLP-it: A UD Pipeline for Italian based on Stanford CoreNLP</i>	
Alessandro Bondielli, Lucia C. Passaro and Alessandro Lenci	58
<i>DARC-IT: a DATaset for Reading Comprehension in Italian</i>	
Dominique Brunato, Martina Valeriani and Felice Dell’Orletta	63
<i>Modelling Italian Construction Flexibility with Distributional Semantics: Are Constructions Enough?</i>	
Lucia Busso, Ludovica Pannitto and Alessandro Lenci	69
<i>The SEEMPAD Dataset for Emphatic and Persuasive Argumentation</i>	
Elena Cabrio and Serena Villata	76
<i>Italian Event Detection Goes Deep Learning</i>	
Tommaso Caselli	82
<i>Enhancing the Latin Morphological Analyser LEMLAT with a Medieval Latin Glossary</i>	
Flavio Massimiliano Cecchini, Marco Passarotti, Paolo Ruffolo, Marinella Testori, Lia Draetta, Martina Fieromonte, Annarita Liano, Costanza Marini and Giovanni Piantanida	88
<i>Is Big Five Better than MBTI? A Personality Computing Challenge Using Twitter Data</i>	
Fabio Celli and Bruno Lepri	94

<i>Automatically Predicting User Ratings for Conversational Systems</i> Alessandra Cervone, Enrico Gambi, Giuliano Tortoreto, Evgeny Stepanov and Giuseppe Riccardi	100
<i>An Efficient Trie for Binding (and Movement)</i> Cristiano Chesi	106
<i>Generalizing Representations of Lexical Semantic Relations</i> Anupama Chingacham and Denis Paperno	112
<i>A NLP-based Analysis of Reflective Writings by Italian Teachers</i> Giulia Chiriatti, Valentina Della Gala, Felice Dell’Orletta, Simonetta Montemagni, Maria Chiara Pettenati, Maria Teresa Sagri and Giulia Venturi	119
<i>Sentences and Documents in Native Language Identification</i> Andrea Cimino, Felice Dell’Orletta, Dominique Brunato and Giulia Venturi	126
<i>Gender and Genre Linguistic Profiling: A Case Study on Female and Male Journalistic and Diary Prose</i> Eleonora Cocciu, Dominique Brunato, Giulia Venturi and Felice Dell’Orletta	132
<i>Conceptual Abstractness: from Nouns to Verbs</i> Davide Colla, Enrico Mensa, Aureliano Porporato and Daniele P. Radicioni	138
<i>Effective Communication without Verbs? Sure! Identification of Nominal Utterances in Italian Social Media Texts</i> Gloria Comandini, Manuela Speranza and Bernardo Magnini	144
<i>On the Readability of Deep Learning Models: the role of Kernel-based Deep Architectures</i> Danilo Croce, Daniele Rossini and Roberto Basili	150
<i>The CHROME Manifesto: Integrating Multimodal Data into Cultural Heritage Resources</i> Francesco Cutugno, Felice Dell’Orletta, Isabella Poggi, Renata Savy and Antonio Sorgente	156
<i>Italian in the Trenches: Linguistic Annotation and Analysis of Texts of the Great War</i> Irene De Felice, Felice Dell’Orletta, Giulia Venturi, Alessandro Lenci and Simonetta Montemagni	161
<i>Lexicon and Syntax: Complexity across Genres and Language Varieties</i> Pietro Dell’Oglio, Dominique Brunato and Felice Dell’Orletta	166
<i>Grammatical Class Effects in Production of Italian Inflected Verbs</i> Maria De Martino, Azzurra Mancuso and Alessandro Laudanna	172
<i>Integrating Terminology Extraction and Word Embedding for Unsupervised Aspect Based Sentiment Analysis</i> Luca Dini, Paolo Curtoni and Elena Melnikova	177
<i>A Linguistic Failure Analysis of Classification of Medical Publications: A Study on Stemming vs Lemmatization</i> Giorgio Maria Di Nunzio and Federica Vezzani	183
<i>Lexical Opposition in Discourse Contrast</i> Anna Feltracco, Bernardo Magnini and Elisabetta Jezek	188
<i>A new Pitch Tracking Smoother based on Deep Neural Networks</i> Michele Ferro and Fabio Tamburini	194

<i>Using and Evaluating TRACER for an Index Fontium Computatus of the Summa contra Gentiles of Thomas Aquinas</i>	
Greta Franzini, Marco Passarotti, Maria Moritz and Marco Böhler	200
<i>Inter-Annotator Agreement in linguistica: una rassegna critica</i>	
Gloria Gagliardi	207
<i>Auxiliary Selection in Italian Intransitive Verbs: A Computational Investigation based on Annotated Corpora</i>	
Ilaria Ghezzi, Cristina Bosco and Alessandro Mazzei	214
<i>Constructing an Annotated Resource for Part-Of-Speech Tagging of Mishnaic Hebrew</i>	
Emiliano Giovannetti, Davide Albanesi, Andrea Bellandi, Simone Marchi and Alessandra Pechioli	220
<i>Concept Tagging for Natural Language Understanding: Two Decadelong Algorithm Development</i>	
Jacopo Gobbi, Evgeny Stepanov and Giuseppe Riccardi	225
<i>The Language-invariant Aspect of Compounding: Predicting Compound Meanings Across Languages</i>	
Fritz Günther and Marco Marelli	231
<i>From General to Specific : Leveraging Named Entity Recognition for Slot Filling in Conversational Language Understanding</i>	
Samuel Louvan and Bernardo Magnini	236
<i>What's in a Food Name: Knowledge Induction from Gazetteers of Food Main Ingredient</i>	
Bernardo Magnini, Vevake Balaraman, Simone Magnolini and Marco Guerini	242
<i>La sentiment analysis come strumento di studio del parlato emozionale?</i>	
Paolo Mairano, Enrico Zovato and Vito Quinci	248
<i>The iDAI.publication: Extracting and Linking Information in the Publications of the German Archaeological Institute (DAI)</i>	
Francesco Mambrini	254
<i>Source-driven Representations for Hate Speech Detection</i>	
Flavio Merenda, Claudia Zaghi, Tommaso Caselli and Malvina Nissim	259
<i>Progettare Chatbot: considerazioni e linee guida</i>	
Eleonora Mollo, Amon Rapp, Dario Mana and Rossana Simeoni	265
<i>Advances in Multiword Expression Identification for the Italian language: The PARSEME Shared Task Edition 1.1</i>	
Johanna Monti, Silvio Ricardo Cordeiro, Carlos Ramisch, Federico Sangati, Agata Savary and Veronika Vincze	272
<i>PARSEME-IT - Issues in verbal Multiword Expressions Identification and Classification</i>	
Johanna Monti, Valeria Caruso and Maria Pia Di Buono	279
<i>Local Associations and Semantic Ties in Overt and Masked Semantic Priming</i>	
Andrea Nadalini, Marco Marelli, Roberto Bottini and Davide Crepaldi	284
<i>Online Neural Automatic Post-editing for Neural Machine Translation</i>	
Matteo Negri, Marco Turchi, Nicola Bertoldi and Marcello Federico	289
<i>EnetCollect in Italy</i>	
Lionel Nicolas, Verena Lyding, Luisa Bentivogli, Federico Sangati, Johanna Monti, Irene Russo, Roberto Gretter and Daniele Falavigna	295

<i>Towards SMT-Assisted Error Annotation of Learner Corpora</i> Nadezda Okinina and Lionel Nicolas	300
<i>Towards Personalised Simplification based on L2 Learners' Native Language</i> Alessio Palmero Apro시오, Stefano Menini, Sara Tonelli, Luca Ducceschi and Leonardo Herzog	306
<i>Tint 2.0: an All-inclusive Suite for NLP in Italian</i> Alessio Palmero Apro시오 and Giovanni Moretti	312
<i>MEDEA: Merging Event knowledge and Distributional vEctor Addition</i> Ludovica Pannitto and Alessandro Lenci	319
<i>LatInfLexi: an Inflected Lexicon of Latin Verbs</i> Matteo Pellegrini and Marco Passarotti	325
<i>Word Embeddings in Sentiment Analysis</i> Ruggero Petrolito and Felice Dell'Orletta	331
<i>Identifying Citation Contexts: a Review of Strategies and Goals.</i> Agata Rotondi, Angelo Di Iorio and Freddy Limpens	336
<i>DialettiBot: a Telegram Bot for Crowdsourcing Recordings of Italian Dialects</i> Federico Sangati, Ekaterina Abramova and Johanna Monti	343
<i>Bootstrapping Enhanced Universal Dependencies for Italian</i> Maria Simi and Simonetta Montemagni	349
<i>Analysing the Evolution of Students' Writing Skills and the Impact of Neo-standard Italian with the help of Computational Linguistics</i> Rachele Sprugnoli, Sara Tonelli, Alessio Palmero Apro시오 and Giovanni Moretti	355
<i>Arretium or Arezzo? A Neural Approach to the Identification of Place Names in Historical Texts</i> Rachele Sprugnoli	361
<i>Multi-source Transformer for Automatic Post-Editing</i> Amirhossein Tebbifakhr, Ruchit Agrawal, Matteo Negri and Marco Turchi	367
<i>Classifying Italian newspaper text: news or editorial?</i> Pietro Totis and Manfred Stede	373
<i>Multi-Word Expressions in spoken language: PoliSdict</i> Daniela Trotta, Michele Stingo, Teresa Albanese, Raffaele Guarasci and Annibale Elia	378