

CLiC-it 2019

The Sixth Italian Conference on Computational Linguistics

Proceedings of the Conference

November 13-15, 2019

Copyright ©2019 for the individual papers by the papers' authors.
Copyright ©2019 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-2481>.
The papers are mirrored in the ACL Anthology.

ISBN 979-1-280-13600-8

Table of Contents

<i>Preface</i>	
Raffaella Bernardi, Roberto Navigli and Giovanni Semeraro	1
<i>Visually-Grounded Dialogue Models: Past, Present, and Future</i>	
Raquel Fernández	11
<i>Impossible Languages and the Architecture of Human Grammars</i>	
Andrea Moro	12
<i>Recognizing and Reducing Bias in NLP Applications</i>	
Dirk Hovy	13
<i>Prerequisite or Not Prerequisite? That's the Problem! An NLP-based Approach for Concept Prerequisite Learning</i>	
Chiara Alzetta, Alessio Miaschi, Giovanni Adorni, Felice Dell'Orletta, Frosina Koceva, Samuele Passalacqua and Ilaria Torre	14
<i>An Empirical Analysis of Linguistic, Typographic, and Structural Features in Simplified German Texts</i>	
Alessia Battisti, Sarah Ebling and Martin Volk	22
<i>Almawave-SLU: A New Dataset for SLU in Italian</i>	
Valentina Bellomaria, Giuseppe Castellucci, Andrea Favalli and Raniero Romagnoli	29
<i>Nove Anni di jTEI: What's New?(Nine Years of jTEI: What's New?)</i>	
Federico Boschetti, Gabriella Pardelli and Giulia Venturi	35
<i>BullyFrame: Cyberbullying Meets FrameNet</i>	
Silvia Brambilla, Alessio Palmero Aprosio and Stefano Menini	41
<i>Lost in Text. A Cross-Genre Analysis of Linguistic Phenomena within Text</i>	
Chiara Buongiovanni, Francesco Gracci, Dominique Brunato and Felice Dell'Orletta	49
<i>Annotating Shakespeare's Sonnets with Appraisal Theory to Detect Irony</i>	
Nicolò Busetto and Rodolfo Delmonte	55
<i>Embeddings Shifts as Proxies for Different Word Use in Italian Newspapers</i>	
Michele Cafagna, Lorenzo De Mattei and Malvina Nissim	63
<i>Suitable Doesn't Mean Attractive. Human-Based Evaluation of Automatically Generated Headlines</i>	
Michele Cafagna, Lorenzo De Mattei, Davide Bacciu and Malvina Nissim	70
<i>There and Back Again: Cross-Lingual Transfer Learning for Event Detection</i>	
Tommaso Caselli and Ahmet Üstün	77
<i>PESInet: Automatic Recognition of Italian Statements, Questions, and Exclamations With Neural Networks</i>	
Sonia Cenceschi, Roberto Tedesco, Licia Sbattella, Davide Losio and Mauro Luchetti	85
<i>What Makes a Review helpful? Predicting the Helpfulness of Italian TripAdvisor Reviews</i>	
Giulia Chiriatti, Dominique Brunato, Felice Dell'Orletta and Giulia Venturi	92
<i>Is This an Effective Way to Annotate Irony Activators?</i>	
Alessandra Teresa Cignarella, Manuela Sanguinetti, Cristina Bosco and Paolo Rosso	98

<i>Robospierre, an Artificial Intelligence to Solve “La Ghigliottina”</i> Nicola Cirillo, Chiara Pericolo and Pasquale Tufano	106
<i>From Sartre to Frege in Three Steps: A* Search for Enriching Semantic Text Similarity Measures</i> Davide Colla, Marco Leontino, Enrico Mensa and Daniele P. Radicioni	113
<i>Is “manovra” Really “del popolo”? Linguistic Insights into Twitter Reactions to the Annual Italian Budget Law</i> Claudia Roberta Combei	121
<i>Cross-Platform Evaluation for Italian Hate Speech Detection</i> Michele Corazza, Stefano Menini, Elena Cabrio, Sara Tonelli and Serena Villata	129
<i>An Open Science System for Text Mining</i> Gianpaolo Coro, Giancarlo Panichi and Pasquale Pagano	136
<i>Detecting Irony in Shakespeare’s Sonnets with SPARSAR</i> Rodolfo Delmonte and Nicolò Busetto	143
<i>Towards an Italian Learner Treebank in Universal Dependencies</i> Elisa Di Nuovo, Cristina Bosco, Alessandro Mazzei and Manuela Sanguinetti	151
<i>Building an Italian Written-Spoken Parallel Corpus: a Pilot Study</i> Elisa Dominutti, Lucia Pifferi, Felice Dell’Orletta, Simonetta Montemagni and Valeria Quochi	159
<i>Italian and English Sentence Simplification: How Many Differences?</i> Martina Fieromonte, Dominique Brunato, Felice Dell’Orletta and Giulia Venturi	166
<i>Error Analysis in a Hate Speech Detection Task: The Case of HaSpeeDe-TW at EVALITA 2018</i> Chiara Francesconi, Cristina Bosco, Fabio Poletto and Manuela Sanguinetti	172
<i>Nunc Est Aestimandum: Towards an Evaluation of the Latin WordNet</i> Greta Franzini, Andrea Peverelli, Paolo Ruffolo, Marco Passarotti, Helena Sanna, Edoardo Signoroni, Viviana Ventura and Federica Zampedri	179
<i>Iride: an Industrial Perspective on Production Grade End to End Dialog System</i> Cristina Giannone, Valentina Bellomaria, Andrea Favalli and Raniero Romagnoli	187
<i>When Lexicon-Grammar Meets Open Information Extraction: a Computational Experiment for Italian Sentences</i> Raffaele Guarasci, Emanuele Damiano, Aniello Minutolo and Massimo Esposito	194
<i>Are Subtitling Corpora really Subtitle-like?</i> Alina Karakanta, Matteo Negri and Marco Turchi	201
<i>Asymmetries in Extraction From Nominal Copular Sentences: a Challenging Case Study for NLP Tools</i> Paolo Lorusso, Matteo Greco, Cristiano Chesi and Andrea Moro	207
<i>Objective Frequency Values of Canonical and Syntactically Modified Idioms: Preliminary Normative Data</i> Azzurra Mancuso and Alessandro Laudanna	215
<i>Gender Detection and Stylistic Differences and Similarities between Males and Females in a Dream Tales Blog</i> Raffaele Manna, Antonio Pascucci and Johanna Monti	221

<i>CROATPAS: A Resource of Corpus-derived Typed Predicate Argument Structures for Croatian</i> Costanza Marini and Elisabetta Jezek	228
<i>Enhancing a Text Summarization System with ELMo</i> Claudio Mastronardo and Fabio Tamburini	236
<i>KIParla Corpus: A New Resource for Spoken Italian</i> Caterina Mauri, Silvia Ballarè, Eugenio Gorla, Massimo Cerruti and Francesco Suriano	243
<i>Evaluating Speech Synthesis on Mathematical Sentences</i> Alessandro Mazzei, Michele Monticone and Cristian Bernareggi	250
<i>Automated Short Answer Grading: A Simple Solution for a Difficult Task</i> Stefano Menini, Sara Tonelli, Giovanni De Gasperis and Pierpaolo Vittorini	257
<i>Games for Learning Old and Special Alphabets – The Case Study of Gamifying Mrežnik</i> Josip Mihaljević	264
<i>Text Frame Detector: Slot Filling Based On Domain Knowledge Bases</i> Martina Miliani, Lucia C. Passaro and Alessandro Lenci	270
<i>Defining Action Types: on the roles of Thematic Structures and Argument Alternations</i> Massimo Moneglia, Alessandro Panunzi and Rossella Varvara	278
<i>HateChecker: a Tool to Automatically Detect Hater Users in Online Social Networks</i> Cataldo Musto, Angelo Sansonetti, Marco Polignano, Giovanni Semeraro and Marco Stranisci	285
<i>The Contribution of Embeddings to Sentiment Analysis on YouTube</i> Moniek Nieuwenhuis and Malvina Nissim	291
<i>A Novel Integrated Industrial Approach with Cobots in the Age of Industry 4.0 through Conversational Interaction and Computer Vision</i> Andrea Pazienza, Nicola Macchiarulo, Felice Vitulano, Antonio Fiorentini, Marco Cammisa, Leonardo Rigutini, Ernesto Di Iorio, Achille Globo and Antonio Trevisi	298
<i>Annotating Hate Speech: Three Schemes at Comparison</i> Fabio Poletto, Valerio Basile, Cristina Bosco, Viviana Patti and Marco Stranisci	304
<i>ALBERTo: Italian BERT Language Understanding Model for NLP Challenging Tasks Based on Tweets</i> Marco Polignano, Pierpaolo Basile, Marco de Gemmis, Giovanni Semeraro and Valerio Basile	312
<i>Evaluating the MuMe Dialogue System with the IDIAL protocol</i> Aureliano Porporato, Alessandro Mazzei, Daniele P. Radicioni and Rosa Meo	318
<i>To be Fair: a Case for Cognitively-Inspired Models of Meaning</i> Simon Preissner and Aurélie Herbelot	325
<i>The Impact of Self-Interaction Attention on the Extraction of Drug-Drug Interactions</i> Luca Putelli, Alfonso Emilio Gerevini, Alberto Lavelli and Ivan Serina	332
<i>Enriching Open Multilingual Wordnets with Morphological Features</i> Stefania Racioppa and Thierry Declerck	339
<i>A Comparison of Representation Models in a Non-Conventional Semantic Similarity Scenario</i> Andrea Amelio Ravelli, Oier Lopez de Lacalle and Eneko Agirre	345

<i>How Much Competence Is There in Performance? Assessing the Distributional Hypothesis in Word Bigrams</i>	
Johann Seltmann, Luca Ducceschi and Aurélie Herbelot	353
<i>Jointly Learning to See, Ask, Decide when to Stop, and then GuessWhat</i>	
Ravi Shekhar, Alberto Testoni, Raquel Fernández and Raffaella Bernardi	361
<i>Creating a Multilingual Terminological Resource using Linked Data: the case of Archaeological Domain in the Italian language</i>	
Giulia Speranza, Carola Carlino and Sina Ahmadi	367
<i>Vir is to Moderatus as Mulier is to Intemperans - Lemma Embeddings for Latin</i>	
Rachele Sprugnoli, Marco Passarotti and Giovanni Moretti	374
<i>Valutazione umana di Google Traduttore e DeepL per le traduzioni di testi giornalistici dall'inglese verso l'italiano (Human evaluation of Google Translator and DeepL for translations of journalistic texts from English into Italian)</i>	
Mirko Tivosanis	381
<i>Prendo la Parola in Questo Consesso Mondiale: A Multi-Genre 20th Century Corpus in the Political Domain</i>	
Sara Tonelli, Rachele Sprugnoli and Giovanni Moretti	388
<i>Reflexives, Impersonals and Their Kin: a Classification Problem</i>	
Kledia Topciu and Cristiano Chesi	396
<i>Annotation and Analysis of the PoliModal Corpus of Political Interviews</i>	
Daniela Trotta, Sara Tonelli, Alessio Palmero Aprosio and Annibale Elia	403
<i>Analyses of Literary Texts by Using Statistical Inference Methods</i>	
Mehmet Can Yavuz	410
<i>Neural Semantic Role Labeling using Verb Sense Disambiguation</i>	
Domenico Alfano, Roberto Abbruzzese and Donato Cappetta	417
<i>Kronos-it: a Dataset for the Italian Semantic Change Detection Task</i>	
Pierpaolo Basile, Giovanni Semeraro and Annalina Caputo	423
<i>How do Physiotherapists and Patients talk? Developing and annotating RiMotivAzione dialogue corpus.</i>	
Andrea Bolioli, Francesca Alloatti, Mariafrancesca Guadalupi, Roberta Iolanda Lanzi, Giorgia Pregolato and Andrea Turolla	429
<i>Standardizing Language with Word Embeddings and Language Modeling in Reports of Near Misses in Seveso Industries</i>	
Simone Bruno, Silvia Maria Ansaldi, Patrizia Agnello and Fabio Massimo Zanzotto	436
<i>Computational Linguistics Against Hate: Hate Speech Detection and Visualization on Social Media in the Contro L'OdioProject</i>	
Arthur T. E. Capozzi, Mirko Lai, Valerio Basile, Cataldo Musto, Marco Polignano, Fabio Poletto, Manuela Sanguinetti, Cristina Bosco, Viviana Patti, Giancarlo Ruffo, Giovanni Semeraro and Marco Stranisci	442
<i>Supporting Journalism by Combining Neural Language Generation and Knowledge Graphs</i>	
Marco Cremaschi, Federico Bianchi, Andrea Maurino and Andrea Primo Pierotti	448

<i>Deep Bidirectional Transformers for Italian Question Answering</i> Danilo Croce, Giorgio Brandi and Roberto Basili	454
<i>Applying Psychology of Persuasion to Conversational Agents through Reinforcement Learning: an Exploratory Study</i> Francesca Di Massimo, Valentina Carfora, Patrizia Catellani and Marco Piastra	460
<i>WebIsAGraph: A Very Large Hypernymy Graph from a Web Corpus</i> Stefano Faralli, Irene Finocchi, Simone Paolo Ponzetto and Paola Velardi	466
<i>CULTURE as a ‘Liquid’ Modern Word. Evidence from Synchronic and Diachronic Language Resources</i> Maristella Gatto	472
<i>A Dataset of Real Dialogues for Conversational Recommender Systems</i> Andrea Iovine, Fedelucio Narducci and Marco de Gemmis	478
<i>Quanti anni hai? Age Identification for Italian</i> Aleksandra Maslennikova, Paolo Labruna, Andrea Cimino and Felice Dell’Orletta	484
<i>Multi-task Learning Applied to Biomedical Named Entity Recognition Task</i> Tahir Mehmood, Alfonso Gerevini, Alberto Lavelli and Ivan Serina	490
<i>Mining Italian Short Argumentative Texts</i> Ivan Namor, Pietro Totis, Samuele Garda and Manfred Stede	496
<i>Fixing Comma Splices in Italian with BERT</i> Daniele Puccinelli, Silvia Demartini and Renée E. D’Aoust	502
<i>A Comparative Study of Models for Answer Sentence Selection</i> Federico Rossetto, Alessio Gravina, Silvia Severini and Giuseppe Attardi	508
<i>An Italian Question Answering System for Structured Data based on Controlled Natural Languages</i> Lucia Siciliani, Pierpaolo Basile, Giovanni Semeraro and Matteo Mennitti	514
<i>The Tenuousness of Lemmatization in Lexicon-based Sentiment Analysis</i> Marco Vassallo, Giuliano Gabrieli, Valerio Basile and Cristina Bosco	520