

# Indice

<b>1. Creating a standard for evaluating distant supervision for relation extraction</b> Azad Abad, Alessandro Moschitti	1
<b>2. Towards Compositional Tree Kernels</b> Paolo Annesi, Danilo Croce, Roberto Basili	7
<b>3. Initial explorations in Kazakh to English statistical machine translation</b> Zhenisbek Assylbekov, Assulan Nurkas	12
<b>4. Adapting linguistic tools for the analysis of Italian medical records</b> Giuseppe Attardi, Vittoria Cozza, Daniele Sartiano	17
<b>5. Tecnologie del linguaggio e monitoraggio dell'evoluzione delle abilità di scrittura nella scuola secondaria di primo grado</b> Alessia Barbagli, Pietro Lucisano, Felice Dell'Orletta, Simonetta Montemagni, Giulia Venturi	23
<b>6. Italian irony detection in Twitter: A first approach</b> Francesco Barbieri, Francesco Ronzano, Horacio Saggion	28
<b>7. A retrieval model for automatic resolution of crossword puzzles in Italian language</b> Gianni Barlacchi, Massimo Nicosia, Alessandro Moschitti	33
<b>8. Analysing word meaning over time by exploiting temporal Random Indexing</b> Pierpaolo Basile, Annalina Caputo, Giovanni Semeraro	38
<b>9. Combining distributional semantic models and sense distribution for effective Italian word sense disambiguation</b> Pierpaolo Basile, Annalina Caputo, Giovanni Semeraro	43
<b>10. A Lesk-inspired unsupervised algorithm for lexical choice from WordNet synsets</b> Valerio Basile	48
<b>11. The Talmud system: A collaborative web application for the translation of the Babylonian Talmud into Italian</b> Andrea Bellandi, Davide Albanesi, Alessia Bellusci, Andrea Bozzi, Emiliano Giovannetti	53

<b>12. Towards a decision support system for text interpretation</b> Alessia Bellusci, Andrea Bellandi, Giulia Benotto, Amedeo Cappelli, Emiliano Giovannetti, Simone Marchi	58
<b>13. An Italian dataset of textual entailment graphs for text exploration of customer interactions</b> Luisa Bentivogli, Bernardo Magnini	63
<b>14. L'integrazione di informazioni contestuali e linguistiche nel riconoscimento automatico dell'ironia</b> Lorenzo Bernardini, Irina Prodanof	67
<b>15. A generic tool for the automatic syllabification of Italian</b> Brigitte Bigi, Caterina Petrone	73
<b>16. Errori di OCR e riconoscimento di entità nell'Archivio Storico de La Stampa</b> Andrea Bolioli, Eleonora Marchioni, Raffaella Ventaglio	78
<b>17. Computer assisted annotation of themes and motifs in ancient Greek epigrams: First steps</b> Federico Boschetti, Riccardo Del Gratta, Marion Lamé	83
<b>18. Defining an annotation scheme with a view to automatic text simplification</b> Dominique Brunato, Felice Dell'Orletta, Giulia Venturi, Simonetta Montemagni	87
<b>19. <i>Senso Comune</i> as a knowledge base of Italian language: the resource and its development</b> Tommaso Caselli, Isabella Chiari, Aldo Gangemi, Elisabetta Jezek, Alessandro Oltramari, Guido Vetere, Laure Vieu, Fabio Massimo Zanzotto	93
<b>20. CorEA: Italian news corpus with emotions and agreement</b> Fabio Celli, Giuseppe Riccardi, Arindam Ghosh	98
<b>21. Detecting attribution relations in speech: A corpus study</b> Alessandra Cervone, Silvia Pareti, Peter Bell, Irina Prodanof, Tommaso Caselli	103
<b>22. Adattamento al progetto dei modelli di traduzione automatica nella traduzione assistita</b> Mauro Cettolo, Nicola Bertoldi, Marcello Federico	108
<b>23. The new basic vocabulary of Italian as a linguistic resource</b> Isabella Chiari, Tullio De Mauro	113
<b>24. Sintassi e semantica dell'hashtag: studio preliminare di una forma di Scritture Brevi</b> Francesca Chiusaroli	117

<b>25. Annotation of complex emotions in real-life dialogues: The case of empathy</b>	
Morena Danieli, Giuseppe Riccardi, Firoj Alam	122
<b>26. Evaluating ImagAct-WordNet mapping for English and Italian through videos</b>	
Irene De Felice, Roberto Bartolini, Irene Russo, Valeria Quochi, Monica Monachini	128
<b>27. CLaSSES: a new digital resource for Latin epigraphy</b>	
Irene De Felice, Margherita Donati, Giovanna Marotta	132
<b>28. Online and multitask learning for Machine Translation quality estimation in real-world scenarios</b>	
José G. C. de Souza, Marco Turchi, Antonios Anastasopoulos, Matteo Negri	138
<b>29. A computational approach to poetic structure, rhythm and rhyme</b>	
Rodolfo Delmonte	144
<b>30. A reevaluation of dependency structure evaluation</b>	
Rodolfo Delmonte	151
<b>31. Analisi Linguistica e stilostatistica Uno studio predittivo sul campo</b>	
Rodolfo Delmonte	158
<b>32. An adaptable morphological parser for agglutinative languages</b>	
Marina Ermolaeva	164
<b>33. Distributed smoothed tree kernel</b>	
Lorenzo Ferrone, Fabio Massimo Zanzotto	169
<b>34. Polysemy alternations extraction using the PAROLE SIMPLE CLIPS Italian lexicon</b>	
Francesca Frontini, Valeria Quochi, Monica Monachini	175
<b>35. Rappresentazione dei concetti azionali attraverso prototipi e accordo nella categorizzazione dei verbi generali. Una validazione statistica</b>	
Gloria Gagliardi	180
<b>36. Correcting OCR errors for German in Fraktur font</b>	
Michel Génèreux, Egon W. Stemle, Verena Lyding, Lionel Nicolas	186
<b>37. Some issues on Italian to LIS automatic translation. The case of train announcements</b>	
Carlo Geraci, Alessandro Mazzei, Marco Angster	191

<b>38. ConParoleTue: crowdsourcing al servizio di un <i>Dizionario delle Collocazioni Italiane per Apprendenti (Dici-A)</i></b>	
Andrea Gobbi, Stefania Spina	197
<b>39. Making <i>Latent SVM<sup>struct</sup></i> practical for coreference resolution</b>	
Iryna Haponchyk, Alessandro Moschitti	203
<b>40. Nominal coercion in space: Mass/count nouns and distributional semantics</b>	
Manuela Hürlimann, Raffaella Bernardi, Denis Paperno	208
<b>41. Part-of-Speech tagging strategy for MIDIA: a diachronic corpus of the Italian language</b>	
Claudio Iacobini, Aurelio De Rosa, Giovanna Schirato	213
<b>42. Distributional analysis of copredication: Towards distinguishing systematic polysemy from coercion</b>	
Elisabetta Jezeq, Laure Vieu	219
<b>43. Publishing PAROLE SIMPLE CLIPS as linguistic linked open data</b>	
Fahad Khan, Francesca Frontini	224
<b>44. A preliminary comparison of state-of-the-art dependency parsers on the Italian Stanford Dependency Treebank</b>	
Alberto Lavelli	229
<b>45. SYMPATHy: Towards a comprehensive approach to the extraction of Italian Word Combinations</b>	
Alessandro Lenci, Gianluca E. Lebani, Sara Castagnoli, Francesca Masini, Malvina Nissim	234
<b>46. Più l'ascolto e più <i>Mi piace!</i> Social media e radio: uno studio preliminare del successo dei post</b>	
Eleonora Lisi, Emanuele Donati, Fabio Massimo Zanzotto	239
<b>47. Estimating lexical resources Impact in text-to-text inference tasks</b>	
Simone Magnolini, Bernardo Magnini	244
<b>48. Parting ways with the partitive view: a corpus-based account of the Italian particle “ne”</b>	
Alice Mariotti, Malvina Nissim	249
<b>49. On the lexical coverage of some resources on Italian cooking recipes</b>	
Alessandro Mazzei	254
<b>50. Event factuality in Italian: Annotation of news stories from the Ita-TimeBank</b>	
Anne-Lyse Minard, Alessandro Marchetti, Manuela Speranza	260

<b>51. An English-Italian MWE dictionary</b> Johanna Monti	265
<b>52. ALCIDE: An online platform for the analysis of language and content in a digital environment</b> Giovanni Moretti, Sara Tonelli, Stefano Menini, Rachele Sprugnoli	270
<b>53. Inner speech, dialogue text and collaborativ learning in virtual learning communities</b> Stefanos Nikiforos, Katia Lida Kermanidis	275
<b>54. Gli errori di un sistema di riconoscimento automatico del parlato. Analisi linguistica e primi risultati di una ricerca interdisciplinare</b> Maria Palmerini, Renata Savy	281
<b>55. “Il Piave mormorava...”: Recognizing locations and other named entities in Italian texts on the Great War</b> Lucia Passaro, Alessandro Lenci	286
<b>56. The importance of being sum. Network analysis of a Latin dependency treebank</b> Marco Passarotti	291
<b>57. I-ChatbIT: An intelligent chatbot for the Italian Language</b> Arianna Pipitone, Vincenzo Cannella, Roberto Pirrone	296
<b>58. Two-dimensional wordlikeness effects in lexical organisation</b> Vito Pirrelli, Claudia Marzi, Marcello Ferro	301
<b>59. Toward disambiguating typed predicate-argument structures for Italian</b> Octavian Popescu, Ngoc Phuoc An Vo, Anna Feltracco, Elisabetta Jezek, Bernardo Magnini	306
<b>60. Il corpus Speaky</b> Fabio Poroli, Massimiliano Todisco, Michele Cornacchia, Cristina Delogo, Andrea Paoloni, Mauro Falcone	311
<b>61. Converting the parallel treebank ParTUT in Universal Stanford Dependencies</b> Manuela Sanguinetti, Cristina Bosco	316
<b>62. Developing corpora and tools for sentiment analysis: the experience of the University of Turin group</b> Manuela Sanguinetti, Emilio Sulis, Viviana Patti, Giancarlo Ruffo, Leonardo Allisio, Valeria Mussa and Cristina Bosco	322
<b>63. Unsupervised antonym-synonym discrimination in vector space</b> Enrico Santus, Qin Lu, Alessandro Lenci, Chu-Ren Huang	328

<b>64. Methods of textual archive preservation</b> Eva Sassolini, Sebastiana Cucurullo, Manuela Sassi	334
<b>65. Combining unsupervised syntactic and semantic models of thematic fit</b> Asad Sayeed, Vera Demberg	339
<b>66. Deep neural network adaptation for children's and adults' speech recognition</b> Romain Serizel, Diego Giuliani	344
<b>67. An Italian corpus for aspect based sentiment analysis of movie reviews</b> Antonio Sorgente, Giuseppe Vettigli, Francesco Mele	349
<b>68. Il <i>Perugia Corpus</i>: una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione</b> Stefania Spina	354
<b>69. Are quantum classifiers promising?</b> Fabio Tamburini	360
<b>70. Geometric and statistical analysis of emotions and topics in corpora</b> Francesco Tarasconi, Vittorio Di Tomaso	365
<b>71. Corpus ICoN: una raccolta di elaborati di italiano L2 prodotti in ambito universitario</b> Mirko Tavosanis	370
<b>72. Coreference resolution for Italian: Assessing the impact of linguistic components</b> Olga Uryupina, Alessandro Moschitti	374
<b>73. A context based model for Sentiment Analysis in Twitter for the Italian language</b> Andrea Vanzo, Giuseppe Castellucci, Danilo Croce, Roberto Basili	379
<b>74. Semantic role annotation of instrument subjects</b> Rossella Varvara, Elisabetta Jezek	384
<b>75. The Italian module for NooJ</b> Simonetta Vietri	389

# Indice

## WRITTEN LANGUAGE TASKS

### Dependency Parsing

#### The Evalita 2014 Dependency Parsing task

Cristina Bosco, Felice Dell’Orletta, Simonetta Montemagni,  
Manuela Sanguinetti, Maria Simi

1

#### Dependency Parsing Techniques for Information Extraction

Giuseppe Attardi, Maria Simi

9

#### Comparing State-of-the-art Dependency

#### Parsers for the EVALITA 2014 Dependency Parsing Task

Alberto Lavelli

15

#### Testing parsing improvements with combination and translation in Evalita 2014

Alessandro Mazzei

21

## Evaluation of Events and Temporal Information

### EVENTI. Evaluation of Events and Temporal Information at Evalita 2014

Tommaso Caselli, Rachele Sprugnoli, Manuela Speranza  
and Monica Monachini

27

#### Experiments in Identification of Italian Temporal Expressions

Giuseppe Attardi and Luca Baronti

35

#### HeidelTime at EVENTI: Tuning Italian Resources and Addressing TimeML’s Empty Tags

Giulio Manfredi, Jannik Strötgen, Julian Zell and Michael Gertz

39

#### FBK-HLT-time: a complete Italian Temporal Processing system for EVENTI-Evalita 2014

Paramita Mirza and Anne-Lyse Minard

44

## **Sentiment Polarity Classification**

### **Overview of the Evalita 2014 SENTiment POLarity Classification Task**

Valerio Basile, Andrea Bolioli, Viviana Patti, Paolo Rosso  
and Malvina Nissim 50

### **UNIBA at EVALITA 2014-SENTIPOLC Task: Predicting tweet sentiment polarity combining micro-blogging, lexicon and semantic features**

Pierpaolo Basile and Nicole Novielli 58

### **ITGETARUNS A Linguistic Rule-Based System for Pragmatic Text Processing**

Rodolfo Delmonte 64

### **Subjectivity, Polarity And Irony Detection: A Multi-Layer Approach**

Elisabetta Fersini, Enza Messina, Federico Alberto Pozzi 70

### **IRADABE: Adapting English Lexicons to the Italian Sentiment Polarity Classification task**

Irazú Hernández Farias, Davide Buscaldi  
and Belém Priego Sánchez 75

### **Linguistically-motivated and Lexicon Features for Sentiment Analysis of Italian Tweets**

Andrea Cimino, Stefano Cresci, Felice Dell'Orletta  
and Maurizio Tesconi 81

### **The CoLing Lab system for Sentiment Polarity Classification of tweets**

Lucia C. Passaro, Gianluca E. Lebani, Laura Pollacci,  
Emmanuele Chersoni and Alessandro Lenci 87

### **The FICLIT+CS@UniBO System at the VALITA 2014 Sentiment Polarity Classification Task**

Pierluigi Di Gennaro and Arianna Rossi 93

### **A Multiple Kernel Approach for Twitter Sentiment Analysis in Italian**

Giuseppe Castellucci, Danilo Croce, Diego De Cao  
and Roberto Basili 98

### **Relying on intrinsic word features to characterise subjectivity, polarity and irony of Tweets**

Francesco Barbieri, Francesco Ronzano and Horacio Saggion 104

### **Self-Evaluating Workflow for Language-Independent Sentiment Analysis**

Arseni Anisimovich 108



## **SPEECH TASKS**

### **Emotion Recogniton Task (ERT)**

#### **EVALITA 2014: Emotion Recognition Task (ERT)**

Antonio Origlia, Vincenzo Galatà 112

#### **A Preliminary Application of Echo State Networks to Emotion Recognition**

Claudio Gallicchio, Alessio Micheli 116

#### **Emotion Recognition with a Kernel Quantum Classifier**

Fabio Tamburini 120

### **Forced Alignment on Children Speech (FACS)**

#### **Forced Alignment on Children Speech**

Piero Cosi, Francesco Cutugno, Vincenzo Galatà and Antonio Origlia 124

#### **The SPPAS participation to Evalita 2014**

Brigitte Bigi 127

### **Human and Machine Dialect Identification from Natural Speech and Artificial Stimuli (HDMI)**

#### **Human and Machine Language / Dialect Identification from Natural Speech and Artificial Stimuli:**

##### **a Pilot Study with Italian Listeners**

Antonio Romano and Claudio Russo 131

### **Speech Activity Detection and Speaker Localization in Domestic Environments (SASLODOM)**

#### **SASLODOM: Speech Activity detection and Speaker Localization in DOMestic environments**

Alessio Brutti, Mirco Ravanelli, Maurizio Omologo 139

#### **The L2F system for the EVALITA-2014 speech activity detection challenge in domestic environments**

Alberto Abad, Miguel Matos, Hugo Meinedo,  
Ramon F. Astudillo, Isabel Trancoso 147

#### **Neural Networks Based Methods for Voice Activity Detection in a Multi-room Domestic Environment**

Giacomo Ferroni, Roberto Bonfigli, Emanuele Principi,  
Stefano Squartini, and Francesco Piazza 153