

Gianluca Lebani

Postdoctoral Fellow

Dipartimento di Filologia,
Letteratura e Linguistica
via Santa Maria, 36, I-56126 PISA
☎ +39 050 2215674

✉ gianluca.lebani@for.unipi.it
🏠 colinglab.humnet.unipi.it/people/gianluca

Research Interests

NLP Distributional Semantics, Corpus Linguistics, Computational Lexicography
Linguistics Verbal Semantics, Lexical Semantics
Psycholinguistics Language Acquisition, Language Processing

Current Position

May. 2013–current Postdoctoral Fellow at the Department of Philology, Literature, and Linguistics of the University of Pisa.
Mar. 2014–current Subject Expert and Teaching Assistant ("Cultore della Materia") in Glottology and Linguistics (L-LIN/01), University of Pisa.

Education

2008–2012 **Ph.D. in Cognitive and Brain Sciences**, *Center for Mind/Brain Sciences (University of Trento) and Fondazione Bruno Kessler, Trento, Italy.*
Programme: Language, Interaction and Computation
Thesis Title: STaRS.sys – designing and building a commonsense-knowledge enriched wordnet for therapeutic purposes
Advisor: E. Pianta
2004–2008 **M.Sc. degree in Linguistics**, *University of Siena, Siena, Italy, 110/110 cum laude.*
Thesis Title: Lexical Information and Word Sense Disambiguation
Relator: Prof. Dr. C. Chesi, Corelator: Prof. Dr. L. Rizzi
2000–2004 **B.A. degree in Communication Science**, *University of Siena, Siena, Italy, 110/110.*
Thesis Title: Autism and Relevance Theory
Relator: Prof. Dr. G. Usberti
Oct. 2002– **Erasmus exchange student**, *University of Genève, Genève, Switzerland.*
Mar. 2003 Department of Linguistics

Research and Academic Experience

2015–2016 **Member of the TRIPLE research group**, *Roma Tre University, Rome.*
Details: TRIPLE (Research Desk on Word and Lexicon) is a research group coordinated by Raffaele Simone, whose main goal is the study of the lexicon in Italian and other languages.
2015 **Scientific Consultant**, *Lexicon Branding Inc., Sausalito, CA, USA.*
Details: supply of corpus-based lexical information on demand
July–Aug. 2008 **Tagging for the Siena University Treebank**, *Interdepartmental Center of Cognitive Studies on Language. University of Siena, Siena, Italy.*
Details: the Siena University Treebank is a manually aligned bilingual treebank Italian - Italian Sign Language provided with both constituency and dependency annotation.

Mar.–Dec. 2007 **Experience in text processing and analysis of linguistic corpora**, *Interdipartimental Center of Cognitive Studies on Language. University of Siena, Siena, Italy.*
Internship

July–Dec. 2003 **Experience in autism diagnosis and therapy**, *Institute of General and Clinical Psychology. University of Siena, Siena, Italy.*
Internship carried out at the Dip. of Child Neuropsychiatry, Hospital 'Le Scotte', Siena

Languages

native **Italian**

fluent **English**

First Certificate of English (Cambridge ESOL)

basic **French**

Attended to two language courses at the University of Genève

Technical skills

programming Python [●●●●●], Javascript [●●●●○], PHP [●●○○○]
statistical analysis R [●●●●●], JASP [●●●●○]
markup languages L^AT_EX [●●●●○], XML [●●●●○], HTML [●●●●●], CSS [●●●●●]
DB administration MySQL [●●●○○], SQLite [●●●○○]
other Experiment Design: E-Prime [●●●○○]
OSs Windows [●●●●●], Linux [●●●●●]

Teaching

- Feb, 2016 *Introduzione all'Analisi Statistica dei Dati Linguistici (Introduction to the Statistical Analysis of Linguistic Data).*
Ph.D. School in Linguistics, Sapienza University of Rome and Roma Tre University
- Feb.–May, 2015 *Laboratorio di R (Introductory Statistics with R Laboratory, part of the course of Applied Linguistics held by Alessandro Lenci).*
Master Program in Linguistics, University of Pisa
- Apr., 2014 *Lectures on WordNet and LexIt (lectures within the course of Applied Linguistics held by Alessandro Lenci).*
Master Program in Linguistics, University of Pisa
- Feb. 19, 2014 *Esperimenti di analisi distribuzionale delle espansioni semantiche (Distributional Semantics Laboratory held with Alessandro Lenci).*
TRIPLE [SIT] Winter School, Rome

Supervision

- 2016 Alice Coli (Master Degree in Linguistics, *University of Pisa*)
Thesis Topic: *When the lawyer is a shark: a study on the categorization of metaphorical and literal word pairs*
Gaia Bonucelli (Master Degree in Linguistics, *University of Pisa*)
Thesis Topic: *Thematic roles: analysis of behavioral and distributional data*
- 2015 Marco S. G. Senaldi (Master Degree in Linguistics, *University of Pisa*)
Thesis Topic: *Casting Light on Idiom Flexibility: A Corpus-based Approach*
- 2013 Veronica Viola (Master Degree in Linguistics, *University of Pisa*)
Thesis Topic: *Towards an Italian VerbNet: a descriptive analysis of argument alternations*

Editorial Activity

area chair	CLiC-it 2016 Track: "Cognitive modeling of language processing and psycholinguistics"
local organizer	First Italian Conference on Computational Linguistics (CLiC-it 2014)
ad hoc reviewer	Behavior Research Methods
conference reviewer	COLING 2016, CogSci 2016, LREC 2016, ACL-IJCNLP 2015*, NAACL 2015*, CLiC-it 2014, EMNLP 2014, COLING 2014, LREC 2014, JSSP 2013* (subreviews are marked by *)
workshop reviewer	COLING 2016 Workshop on "Cognitive Aspects of the Lexicon", ACL 2016 Workshop on "Cognitive Aspects of Computational Language Learning", EMNLP 2015 Workshop on "Cognitive Aspects of Computational Language Learning", EACL 2014 Workshop on "Cognitive Aspects of Computational Language Learning", EMNLP 2013 Workshop on "Computational Models of Human Language Acquisition and Processing", ACL 2013 Student Research Workshop, ICWSM 2013 Workshop on "Computational Personality Recognition", LREC 2010 Workshop "Semantic Relations. Theory and Practice"

Publications

- 2016 G. E. Lebani and A. Lenci. "Beware the Jabberwock, dear reader! Testing the distributional reality of construction semantics". *Proceedings of the Workshop on Cognitive Aspects of the Lexicon (CogALex-V)*. Osaka, Japan, 2016, pp. 8–18
- M. S. G. Senaldi, G. E. Lebani, and A. Lenci. "Determining the Compositionality of Noun-Adjective Pairs with Lexical Variants and Distributional Semantics". *Proceedings of the Third Italian Conference on Computational Linguistics (CLiC-it 2016)*. Napoli, Italy, 2016. **Distinguished Young Paper Award**
- M. S. G. Senaldi, G. E. Lebani, and A. Lenci. "Lexical Variability and Compositionality: Investigating Idiomaticity with Distributional Semantic Models". *Proceedings of the 12th Workshop on Multiword Expressions*. Berlin, Germany, 2016, pp. 21–31
- G. Rambelli, G. E. Lebani, L. Prévot, and A. Lenci. "LexFr: adapting the LexIt framework to build a corpus-based French subcategorization lexicon". *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*. Portorož, Slovenia, 2016, pp. 930–937
- S. Castagnoli, G. E. Lebani, A. Lenci, F. Masini, M. Nissim, and L. C. Passaro. "POS-patterns or Syntax? Comparing methods for extracting Word Combinations". *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives (Full papers) - Fraseología computacional y basada en corpus: perspectivas monolingües y multilingües (Trabajos completos)*. Ed. by G. Corpas Pastor. 2016, pp. 101–114
- 2015 G. E. Lebani, A. Bondielli, and A. Lenci. "You are what you do. An empirical characterization of the semantic content of the thematic roles for a group of Italian verbs". *Journal of Cognitive Science* 16.4 (2015), pp. 401–430
- M. Nissim, S. Castagnoli, F. Masini, G. E. Lebani, L. C. Passaro, and A. Lenci. "Automatic extraction of Word Combinations from corpora: evaluating methods and benchmarks". *Proceedings of the Second Italian Conference on Computational Linguistics (CLiC-it 2015)*. Trento, Italy, 2015, pp. 204–209

- G. E. Lebani, M. S. G. Senaldi, and A. Lenci. "Modeling Idiom Variability with Entropy and Distributional Semantics". *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics (QITL 6)*. Tübingen, Germany, 2015
- A. Lenci, G. E. Lebani, M. S. Senaldi, S. Castagnoli, F. Masini, and M. Nissim. "Mapping the Constructicon with SYMPATHy. Italian Word Combinations between fixedness and productivity". *Proceedings of the NetWordS Final Conference on Word Knowledge and Word Usage: Representations and Processes in the Mental Lexicon*. Pisa, Italy, 2015, pp. 144–149
- 2014 A. Lenci, G. E. Lebani, S. Castagnoli, F. Masini, and M. Nissim. "SYMPATHy: Towards a comprehensive approach to the extraction of Italian Word Combinations". *Proceedings of the First Italian Conference on Computational Linguistics (CLiC-it 2014)*. Pisa, Italy, 2014, pp. 234–238
- L. C. Passaro, G. E. Lebani, L. Pollacci, E. Chersoni, and A. Lenci. "The CoLing Lab system for Sentiment Polarity Classification of tweets". *Proceedings of the Fourth International Workshop EVALITA 2014*. Pisa, Italy, 2014, pp. 87–92
- G. E. Lebani, V. Viola, and A. Lenci. "Bootstrapping an Italian VerbNet: data-driven analysis of verb alternations". *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. Reykjavik, Iceland, 2014, pp. 1127–1134
- L. Romeo, G. E. Lebani, N. Bel, and A. Lenci. "Choosing which to use? A study of distributional models for nominal lexical semantic classification". *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. Reykjavik, Iceland, 2014, pp. 4366–4373
- F. Boschetti, A. Cimino, F. Dell'Orletta, G. E. Lebani, L. Passaro, P. Picchi, G. Venturi, S. Montemagni, and A. Lenci. "Computational Analysis of Historical Documents: An Application to Italian War Bulletins in World War I and II". *Proceedings of the Workshop on "Language Resources and Technologies for Processing and Linking Historical Documents and Archives" (LRT4HDA 2014)*. Reykjavik, Iceland, 2014
- 2012 G. E. Lebani and E. Pianta. "Encoding Commonsense Lexical Knowledge into WordNet". *Proceedings of the 6th Global WordNet Conference (GWC 2012)*. Matsue, Japan, 2012, pp. 159–166
- 2010 G. E. Lebani and E. Pianta. "Human Language Technologies supporting Therapeutic Practices for Language Disorders: the Project STaRS.sys". *Proceedings of the 7th annual meeting of the Italian Society of Cognitive Science (AISC-2010)*. Trento, Italy, 2010, pp. 52–56
- G. E. Lebani and E. Pianta. "Exploiting Lexical Resources for Therapeutic Purposes: the Case of WordNet and STaRS.sys". *Proceedings of the 2nd Workshop on Cognitive Aspects of the Lexicon (CogALex-2)*. Beijing, China, 2010, pp. 12–17
- G. E. Lebani and E. Pianta. "A Feature Type Classification for Therapeutic Purposes: a preliminary evaluation with non-expert speakers". *Proceedings of the Fourth Linguistic Annotation Workshop (LAW IV)*. Uppsala, Sweden, 2010, pp. 157–161
- 2008 C. Chesi, G. Lebani, and M. Pallottino. "A Bilingual Treebank (ITA-LIS) suitable for Machine Translation: what Cartography and Minimalism teach us". *STiL - Studies in Linguistics 2* (2008), pp. 166–185

Abstracts

- 2016 M. S. G. Senaldi, G. E. Lebani, and A. Lenci. "Predicting crowdsourced idiomatic flexibility judgments from corpus-based statistics". *The 22nd AMLaP conference, Architectures and Mechanisms for Language Processing*. Bilbao, Spain, 2016
- 2015 S. Castagnoli, F. Masini, M. Nissim, G. E. Lebani, M. S. Senaldi, and A. Lenci. "Extracting and Analysing Italian Word Combinations with SYntactically Marked PATterns". *PARSEME 4th general meeting*. Valletta, Malta, 2015
- S. Castagnoli, G. E. Lebani, A. Lenci, F. Masini, M. Nissim, and V. Piunno. "Towards a corpus-based online dictionary of Italian Word Combinations". *COSTNENeL (Working Group 3) Meeting, Automatic Knowledge Acquisition (AKA) for Lexicography*. Herstmonceux Castle, UK, 2015
- 2013 A. Lenci and G. E. Lebani. "Per una classificazione semantico-distribuzionale dei verbi italiani". *XLVII Congresso Internazionale della Società di Linguistica Italiana (SLI 2013)*. Salerno, Italy, 2013

last updated: January 28, 2017