

# Gianluca Lebani

*Computational Linguist*

Department of Linguistics and  
Comparative Cultural Studies  
Ca' Foscari University of Venice

✉ [gianluca.lebani@unive.it](mailto:gianluca.lebani@unive.it)  
📄 <https://gianlucalebani.github.io/>

## Current Position

July 2019 – current Research fellow at the Department of Linguistics and Comparative Cultural Studies of the Ca' Foscari University of Venice

## Research Interests

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

## Research and Academic Experience

- Sept. 2018–June 2019 **Lecturer**, *Department of Languages, Literature and Communication, Utrecht University, Utrecht.*
- May–Nov. 2018 **Visiting Researcher**, *Computational Linguistics Laboratory, Department of Philology, Literature, and Linguistics, University of Pisa, Pisa.*
- Sept.–Dec. 2017 **Lecturer**, *Amsterdam University College, Amsterdam.*
- May 2013–Jan. 2017 **Postdoctoral Fellow**, *Computational Linguistics Laboratory, Department of Philology, Literature, and Linguistics, University of Pisa, Pisa.*
- Mar. 2014–Jan. 2017 **Subject Expert and Teaching Assistant in Glottology and Linguistics**, *University of Pisa, Pisa.*
- 2015–2016 **Member of the TRIPLE research group**, *Roma Tre University, Rome.*
- 2015 **Scientific Consultant**, *Lexicon Branding Inc., Sausalito, CA, USA.*
- July–Aug. 2008 **Tagging for the Siena University Treebank**, *Interdisciplinary Center for Cognitive Studies of Language, University of Siena, Siena, Italy.*

## 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  
Advisor: Dr. C. Chesi, Co-advisor: 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  
Advisor: Prof. Dr. G. Usberti

Oct. 2002– **Erasmus exchange student**, *University of Genève*, Genève, Switzerland.  
Mar. 2003 Department of Linguistics

---

## Technical skills

programming Python [ ●●●●● ], Javascript [ ●●●●○ ], PHP [ ●●○○○ ]  
statistical analysis R [ ●●●●● ], JASP [ ●●●●● ]  
markup languages L<sup>A</sup>T<sub>E</sub>X [ ●●●●○ ], XML [ ●●●●○ ], HTML & CSS [ ●●●●● ]  
DB administration MySQL [ ●●●○○ ], SQLite [ ●●●○○ ]  
other Experiment Design: E-Prime [ ●●●○○ ]  
OSs Windows [ ●●●●● ], Linux [ ●●●●● ]

---

## Languages

native **Italian**  
fluent **English**  
basic **French**

---

## Teaching

AY 2020-21 *Digital Text Analysis*.  
onward MA Language Sciences, Ca' Foscari University of Venice

AY 2019-20 *Computational Linguistics*.  
onward MA Language Sciences, Ca' Foscari University of Venice

AY 2018-19 *Computational Linguistics (coordinator: F. Adriaans)*.  
BA Artificial Intelligence, Utrecht University

AY 2018-19 *Introduction to Linguistics for AI (coordinator: R.W.F. Nouwen)*.  
BA Artificial Intelligence, Utrecht University

AY 2018-19 *Language and Computation (coordinator: F. Adriaans)*.  
BA Linguistics, Utrecht University

AY 2018-19 *Words in the Mind*.  
BA Linguistics, Utrecht University

AY 2018-19 *Machine Learning (coordinator: T. Deoskar)*.  
BA Artificial Intelligence, Utrecht University

AY 2018-19 *Digital Methods in Linguistics*.  
RMA Linguistics, Utrecht University

AY 2017-18 *Text Mining & Collective Intelligence*.  
Sciences Curriculum, Amsterdam University College

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

AY 2014-15 *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

AY 2013-14 *Lectures on WordNet and Lexlt (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

- 2020 Carolina Gomiero (MD in Language Sciences, *Ca' Foscari University of Venice*)  
Thesis Topic: *The Distribution of Indefinite Determiners in the Venetan Dialects*  
Luca Molinari (MD in Language Sciences, *Ca' Foscari University of Venice*)  
Thesis Topic: *The expression of indefiniteness and optionality in the dialect of Piacenza*
- 2016 Alice Coli (MD 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 (MD in Linguistics, *University of Pisa*)  
Thesis Topic: *Thematic roles: analysis of behavioral and distributional data*
- 2015 Marco S. G. Senaldi (MD in Linguistics, *University of Pisa*)  
Thesis Topic: *Casting Light on Idiom Flexibility: A Corpus-based Approach*
- 2013 Veronica Viola (MD 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, Cognitive Processing  |
| conference reviewer | EACL 2021; AACL-IJCNLP 2020, ACL 2020, 2018; CogSci 2020, 2019, 2018, 2017, 2016; COLING 2020, 2018, 2016, 2014; EMNLP 2020; *SEM 2019, 2017; CLiC-it 2019, 2018, 2017, 2014; CONLL 2019; EMNLP-IJCNLP 2019; IWCS 2019; LREC 2018, 2016, 2014  |
|                     | workshop reviewer EMNLP 2020/NAACL 2019 Workshop on "Cognitive Modeling and Computational Linguistics"; ACL 2018/ACL 2016/EMNLP 2015/EACL 2014 Workshops on "Cognitive Aspects of Computational Language Learning", ESSLLI 2017 Student Session  |
| ; LIGHTS 2018       | COLING 2016 Workshop on "Cognitive Aspects of the Lexicon", 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

- 2018 A. Lenci, G. E. Lebani, and L. C. Passaro. "The Emotions of Abstract Words: A Distributional Semantic Analysis". *Topics in Cognitive Science* 10.3 (2018), pp. 550–572. **Among the journal's top cited papers**
- G. E. Lebani and A. Lenci. "A Distributional Model of Verb-Specific Semantic Roles Inferences". *Language, Cognition, and Computational Models*. Ed. by T. Poibeau and A. Villavicencio. Cambridge University Press, 2018. Chap. 6, pp. 118–158

- 2017 M. S. G. Senaldi, G. E. Lebani, and A. Lenci. "Determining the Compositionality of Noun-Adjective Pairs with Lexical Variants and Distributional Semantics". *Italian Journal of Computational Linguistics* 3.1 (2017), pp. 43–58
- G. E. Lebani and A. Lenci. "Modelling the meaning of constructions with distributional semantics". *The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding, Technical Report SS-17-02*. Stanford, CA, 2017, pp. 197–204
- A. Lenci, F. Masini, M. Nissim, S. Castagnoli, G. E. Lebani, L. C. Passaro, and M. S. G. Senaldi. "How to harvest Word Combinations from corpora: Methods, evaluation and perspectives". *Studi e Saggi Linguistici* 55.2 (2017), pp. 45–68
- 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 12<sup>th</sup> 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 10<sup>th</sup> 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 6<sup>th</sup> 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 9<sup>th</sup> 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 9<sup>th</sup> 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 6<sup>th</sup> 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 7<sup>th</sup> 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 2<sup>nd</sup> 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

last updated: August 14, 2020