Ludovica Pannitto

Experience

University of Salerno, Adjunct/Contract professor — April 2026 – April 2027, Salerno, Italy

- Responsibility for "Basic Computer Science" (*Informatica di Base*), 9 cfu, for Bachelor course in "Studi Diplomatici, Internazionali e sulla Sicurezza Globale"

University of Bologna, NLP Lab Manager — June 2023 – June 2026, Bologna, Italy

- Managing the "Experimental Laboratory" at the Department of Modern Languages, Literatures and Cultures.
- Overseeing computational linguistics projects and interdisciplinary research.
- Supervising student interns in annotation, tool development, and scientific communication.

Education

PhD **CIMeC - University of Trento**, Cognitive and Brain Sciences — Nov 2018 – May 2023, Rovereto (TN), Italy

- Dissertation: *CALaMo: a Constructionist perspective on the Analysis of linguistic behaviour of Language Models*
- Supervisor: Dr. Aurélie Herbelot

MA **University of Pisa**, Digital Humanities – Language Technologies — Oct 2016 – Oct 2018, Pisa, Italy

- Final Mark: 110 cum laude/110
- Dissertation: *Event Knowledge in Compositional Distributional Semantics*
- Supervisor: Prof. Alessandro Lenci
- **Coursework (selection):** Introduction to Artificial Intelligence, Information Retrieval, Natural Language Processing, Applied Linguistics, Discrete Mathematics and Linear Algebra, Algorithm Engineering, Advanced Computational Linguistics

BA **University of Pisa**, Digital Humanities — Oct 2013 – Jul 2016, Pisa, Italy

- Final Mark: 110 cum laude/110
- **Coursework (selection):** Theoretical foundations and programming, General Linguistics, Computational Linguistics, Text Encoding, Databases and Web Programming, Probability Calculation and Statistics

Advanced Training

2nd UniDive training school on Linguistic Diversity in NLP — Jan 2026, Erevan, Armenia

Julia High Performance — Feb 2025, CINECA, Bologna, Italy

Introduction to Fortran for Scientific Computing — Nov 2024, CINECA, Bologna, Italy

1st UniDive training school — Jul 2024, Chișinău, Moldova

ASPP 14th Advanced Scientific Programming in Python — Aug 2022, Bilbao, Spain

ESSLLI 30th European Summer School in Logic, Language and Information — Aug 2018, Sofia, Bulgaria

Supervision

Il parlato spontaneo nell'interlingua di un apprendente L1 bangla: un'analisi corpus-based per la didattica dell'italiano L2 2026

- Master thesis, co-supervision with Prof. Rosa Pugliese

'Layer su Layer': identifying and disambiguating the Italian NPN construction through BERT's contextual embeddings 2025

- Master thesis, co-supervision with Prof. Francesca Masini

The Fate of Interaction after ASR Decoding Optimization 2025

- Master thesis, co-supervision with Prof. Adriano Ferraresi, Prof. Caterina Mauri, Prof. Maja Miličević Petrović

Categorizzazione ad-hoc e gestualità: una indagine corpus-based sul corpus `gest-it` 2025

- Master thesis, co-supervision with Prof. Caterina Mauri

Teaching

Hands-on Speech Corpus Processing: Standardization, Automation, and GitOps for Reproducibility Spring 2026, Como, Italy

- Hands-on tutorial on spoken language corpora pre-processing and management, focusing on reproducibility, automation, and collaborative workflows.
- Focus on creation of UD-compliant resource creation, GitOps methodologies, FAIR data principles, and best practices for ensuring data integrity and interoperability.

Informatica di Base Spring 2026, Salerno, Italy

- University course (9 CFU), University of Salerno.

Data Plumbers' Corner 2024-2026, Bologna, Italy

- Ideator and Co-instructor (with Arianna Masciolini and Luca Rinaldi) for Data Plumbers' Corner — a recurring series on data processing and software tooling for research.
- Designed and delivered hands-on modules covering: Bash scripting, UNIX shell, regular expressions, Python for data processing, version control with Git, collaborative software development on GitHub, documentation practices and software maintenance, basic computer architecture and filesystem concepts.

Semantics with Transformers: a case study of the NPN construction Fall 2025, Bologna, Italy

- Course held in collaboration with Greta Gorzoni.
- Collective reading (2 hours) of research paper demonstrating the study case and identification of processing pipeline.
- Introduction (4 hours) to basic Python for reproducing a research pipeline.
- Basic tools (6 hours) for dataset creation, data annotation and inter-coder agreement. Demonstration of interaction with BERT and Scikit-Learn library.

Computational Tools for the Study of Meaning Fall 2024, Bologna, Italy

- Tutorial (2 hours) on text structure and corpus creation with SketchEngine.
- Introduction (6 hours) to regular expressions, bash scripting and Python for data processing and pipeline construction.
- Fundamentals (4 hours) of Distributional Semantics and evaluation of vector space models.

Distributional Semantics Fall 2023, Bologna, Italy

- Tutorial (12 hours) on the fundamentals of distributional approaches to meaning.

Introduction to Programming in Python for the Humanities Fall 2021, Bologna, Italy

- Introduction (8 hours) basic programming concepts and text processing with Python.

Publications

Reuse by Design: A Pivot-Based Architecture for the KIParla Corpus of Spoken Italian 2026
Ludovica Pannitto, Caterina Mauri
Journal of Open Humanities Data
10.5334/johd.527

'Layer su Layer': Identifying and Disambiguating the Italian NPN Construction in BERT's family 2026
Greta Gorzoni, Ludovica Pannitto, Francesca Masini
Proceedings of CMCL 2026 (15th Workshop on Cognitive Modeling and Computational Linguistics)
10.63317/532xoam2mj8s

Analyzing Environmental Discourse through Construction-Based Pattern Extraction 2026
Elisa Chierchiello, Eliana Di Palma, Ludovica Pannitto and Cristina Bosco
Proceedings of NLP4Ecology - Workshop at LREC 2026
10.63317/3d7q5k2cavae

Say Again? The Limits of Whisper with Conversation. A Case Study on the KIParla Corpus. 2026
Martina Simonotti, Ludovica Pannitto, Caterina Mauri, Adriano Ferraresi and Gabriele Carioli
Proceedings of SPEAKABLE (Speech Language Models in Low-Resource Settings - Workshop at LREC 2026)
10.63317/2so5y449gb4w

Coconstructions in spoken data: UD annotation guidelines and first results 2026
Ludovica Pannitto, Kaja Dobrovoljc, Sylvain Kahane, Elena Battaglia, Bruno Guillaume, Caterina Mauri and Eleonora Zucchini
Proceedings of the Ninth Workshop on Universal Dependencies (UDW 2026)
10.63317/3kcohrckgnkz

Temporal, aspectual, and modal evaluation in nominal reference: the case of Italian 2026
Caterina Mauri, Ludovica Pannitto
Folia Linguistica (Accepted)
10.1515/flin-2025-0142

Is Semi-Automatic Transcription Useful in Corpus Creation? Preliminary Considerations on the KIParla Corpus 2026
Martina Simonotti, Ludovica Pannitto, Eleonora Zucchini, Silvia Ballaré, Caterina Mauri
Proceedings of the Language Resources and Evaluation Conference (LREC 2026)
10.63317/3gw6i7cbs5rc

Survey of Tools for Manual Linguistic Annotation: Supporting Diversity through Interactive Exploration 2026
Ludovica Pannitto, Kaja Dobrovoljc, Bruno Guillaume
Proceedings of the Language Resources and Evaluation Conference (LREC 2026)
10.63317/4u4be6bbtj8e

Shaping reference through temporality: The case of Italian nominals 2026
Caterina Mauri, Ludovica Pannitto
Lingue e Linguaggio (Accepted)

Introducing KIParla Forest: seeds for a UD annotation of interactional syntax 2025
Ludovica Pannitto, Eleonora Zucchini, Silvia Ballaré, Cristina Bosco, Caterina Mauri, Manuela Sanguinetti
Proceedings of the Eighth International Conference on Dependency Linguistics (Depling, SyntaxFest 2025)
aclanthology.org/2025.depling-1.5/

Constraining constructions with WordNet: pros and cons for the semantic annotation of fillers in the Italian Constructicon 2025
Flavio Pisciotta, Ludovica Pannitto, Lucia Busso, Beatrice Bernasconi, Francesca Masini
Proceedings of the 13th Global WordNet Conference (GWC 2025)
aclanthology.org/2025.gwc-1.9.pdf/

Deriving semantic classes of Italian adjectives via word embeddings: a large-scale investigation 2025
Ivan Lacić, Ludovica Pannitto
Proceedings of the 13th Global WordNet Conference (GWC 2025)
aclanthology.org/2025.gwc-1.34.pdf/

Did Somebody Say 'Gest-IT'? A Pilot Exploration of Multimodal Data Management 2024
Ludovica Pannitto, Lorenzo Albanesi, Laura Marion, Federica Maria Martines, Carmelo Caruso, Claudia S. Bianchini, Francesca Masini and Caterina Mauri
Proceedings of CLiC-It 2024
aclanthology.org/2024.clicit-1.80

Verbal and visual metaphors of an unfolding crisis: Emerging styles of representation in British and Italian front page news at the onset of the COVID-19 pandemic 2024
Antonella Luporini, Claudia R. Combei, Ludovica Pannitto
DIVE-IN: An International Journal on Diversity and Inclusion
10.6092/issn.2785-3233/21623

CALaMo: A Constructionist Assessment of Language Models 2023
Ludovica Pannitto, Aurélie Herbelot
Proceedings of the First Workshop on Construction Grammars and NLP (CxGs+NLP)
aclanthology.org/2023.cxgsnlp-1.3

Can Recurrent Neural Networks Validate Usage-Based Theories of Grammar Acquisition? 2022
Ludovica Pannitto, Aurélie Herbelot
Frontiers in Psychology
10.3389/fpsyg.2022.741321

Che cos'è la linguistica computazionale 2022
Malvina Nissim, Ludovica Pannitto
https://www.carocci.it/prodotto/che-cose-la-linguistica-computazionale
Carocci Editore

Teaching NLP with Bracelets and Restaurant Menus: An Interactive Workshop for Italian Students 2021
Ludovica Pannitto et al.
Proceedings of the Fifth Workshop on Teaching NLP (ACL 2021)
10.18653/v1/2021.teachingnlp-1.26

A Dissemination Workshop for Introducing Young Italian Students to NLP 2021
Ludovica Pannitto et al.
Proceedings of the Fifth Workshop on Teaching NLP (ACL 2021)
10.18653/v1/2021.teachingnlp-1.7

Recurrent Babbling: Evaluating the Acquisition of Grammar from Limited Input Data 2020
Ludovica Pannitto, Aurélie Herbelot
Proceedings of CoNLL 2020
10.18653/v1/2020.conll-1.13

Are Word Embeddings Really a Bad Fit for the Estimation of Thematic Fit? 2020
Emmanuele Chersoni, Ludovica Pannitto et al.
Proceedings of LREC 2020
aclanthology.org/2020.lrec-1.700

A Structured Distributional Model of Sentence Meaning and Processing 2019
Emmanuele Chersoni, Enrico Santus, Ludovica Pannitto, Alessandro Lenci, Philippe Blache, Chu-Ren Huang
Natural Language Engineering
10.1017/S1351324919000214

Event Knowledge in Compositional Distributional Semantics 2019
Ludovica Pannitto, Alessandro Lenci
Italian Journal of Computational Linguistics (IJCoL)
10.4000/ijcol.463

Refining the Distributional Inclusion Hypothesis for Unsupervised Hypernym Identification 2018
Ludovica Pannitto, Lavinia Salicchi, Alessandro Lenci
Italian Journal of Computational Linguistics (IJCoL)
10.4000/ijcol.506

Modelling Italian Construction Flexibility with Distributional Semantics 2018
Lucia Busso, Ludovica Pannitto, Alessandro Lenci
Proceedings of CLiC-It 2018
10.4000/books.aaccademia.3106

MEDEA: Merging Event Knowledge and Distributional vEctor Addition 2018
Ludovica Pannitto, Alessandro Lenci
Proceedings of CLiC-It 2018
10.4000/books.aaccademia.3577

FA3L at SemEval-2017: A ThRee Embeddings RNN for Question Answering 2017
Giuseppe Attardi, Ludovica Pannitto et al.
Proceedings of SemEval 2017
10.18653/v1/S17-2048

Talks

Data annotation for AI-driven data analysis, Nov 2025

- Talk in collaboration with Giorgia Troiani, PhD, at "III Data Analysis Appetiser", PhD in English Linguistics - Universidad Complutense de Madrid - Spain

Treebanking the KIParla corpus of Spoken Italian., Oct 2025

- Department of Computer Science, University of Turin - Italy

Kickstarting the Italian Constructicon: from idea to implementation through data structuring., May 2025

- Workshop "Pedagogical Application of Constructicography Workshop 2025" - University of Gothenburg - Sweden

Treebanking the KIParla corpus of Spoken Italian: new formats and new perspectives for interoperability., May 2025

- Higher seminar series - Humanistiska fakulteten - University of Gothenburg - Sweden

Academic service

Conference Organization

International conference "Talking Data": Methodological and theoretical challenges raised by spoken interaction data — 2025, Bologna, Italy
Organizing Committee

CLiC-it: Eleventh Italian Conference on Computational Linguistics — 2025, Cagliari, Italy
Senior Program Committee

SpLAn-UD: UniDive Workshop on Spoken Language Annotation for Universal Dependencies — 2025, Bologna, Italy
Organizing Committee

Doctoral Student Day — 2021, online, Italy
Organizing Committee

Peer reviewing

LREC	2026
PACLIC39, RANLP, CMCL, GWC2025	2025
LREC-COLING	2024
COLING	2022
ACL-IJCNLP, IWCS, CMCL, EACL-SRW	2021
COLING, LiNCR,	2020
CLiC-it, IWCS	2019

Dissemination and Outreach

Public Engagement and Citizen Science

- Co-organized and delivered dissemination workshops on NLP for high school students in Italy, including interactive activities such as *"Teaching NLP with Bracelets and Restaurant Menus"* (ACL 2021) and *"A Dissemination Workshop for Introducing Young Italian Students to NLP"* (ACL 2021).
- Participation in multiple science festivals and outreach events, including *Festival della Scienza* (2020, 2021), *Bergamo Scienza* (2019), and the *European Researchers' Night* (2024), with activities on computational linguistics for the general public.
- Contributor to the public blog *Linguisticamente* with the article *"Ma un computer mi capisce? Cos'è e a cosa serve la Linguistica Computazionale"* (2020), aimed at explaining computational linguistics to a broad audience.
- Involved in citizen science initiatives such as *"KIParla: Exploring Diversity in Spoken Italian"* presented at *LingComm25* (2025) and other public communication events.
- Co-authored the book "Che cos'è la linguistica computazionale" (Carocci Editore, 2022), introducing computational linguistics to a general Italian audience.

Community Building and Science Communication

- Member of the discussion panel *"LingComm with Young People"* during the *International Conference on Linguistics Communication* (2021).
- Participant in the project *"Penne Amiche della Scienza"* for schools (2020–2021), promoting scientific thinking and language awareness.
- Active in the development of public engagement initiatives for the Italian Association for Computational Linguistics (AILC).

Awards

"Emanuele Pianta" Award for the Best Master Thesis, issued by Italian Association of Computational Linguistics for the best Master Thesis (Laurea Magistrale) in Computational Linguistics submitted at an Italian University — 2019

Academic Collaborations

- **UniDive COST Action CA21167:** Active member contributing to Working Group 1 (WG1) and Working Group 2 (WG2). Co-lead of:
  - *Task 1.5: Annotation of Spoken Data* (with Sylvain Kahane and Kaja Dobrovoljc)
  - *Task 2.5: Extensions of UD-PARSEME annotation for constructions* (with Francesca Masini)
- **KIParla Project:** Data Manager for the KIParla corpus (`https://kiparla.it/`), overseeing computational developments and annotation strategies for the first UD treebank of spoken Italian.
- **Italian Association for Computational Linguistics (AILC):** Active collaborator in dissemination, public engagement, and development of activities aimed at increasing the visibility of computational linguistics in Italy.

Skills

Languages:

- **Italian:** Native speaker, **English:** Fluent (C2 level), **Italian Sign Language (LIS):** Fluent, **French:** Basic proficiency (B1 level), **German:** Basic proficiency (B1 level)

Programming and Technical Skills:

- **Programming Languages:** Python (advanced), Bash (advanced), Regular Expressions (advanced), Julia (basic), Fortran (basic)
- **Data Management:** Git (advanced), Linux distributions (advanced), corpus creation and annotation, reproducible workflows (GitOps)
- **Natural Language Processing:** Distributional Semantics, corpus processing, syntax and semantic annotation, computational modeling of language
- **Tools and Technologies:** NLTK, spaCy, Hugging Face Transformers, SketchEngine, Universal Dependency tools, TEI encoding tools

Soft Skills:

- Project management and leadership in interdisciplinary research settings; Mentoring and supervision of students in linguistics, computational linguistics, and NLP; Public speaking, science communication, and outreach; Collaborative work in international, cross-disciplinary teams.

Bologna, 16/07/2026

