

---

# MARIO GIULIANELLI

[m.giulianelli@uva.nl](mailto:m.giulianelli@uva.nl)



[@glnmario](https://twitter.com/glnmario)



[glnmario](https://github.com/glnmario)



[glnmario.github.io](https://glnmario.github.io)

## Profile

---

PhD candidate at the [Institute for Logic, Language and Computation](#) of the [University of Amsterdam](#), in the [Dialogue Modelling Group](#) led by [Raquel Fernández](#). My research interests lie in the following areas:

- Computational semantics and pragmatics
- Neural language and dialogue models
- Natural language generation
- Language variation and change in communities of speakers
- Explainability, interpretability, and fair evaluation of NLP models

What I am most excited about is studying human and machine **strategies of communication** using computational models of **language understanding** and **generation**. This is the main topic of my PhD, which I address with machine learning, information theory, and statistical modelling.

## Education

---

### PhD Natural Language Processing

#### **University of Amsterdam**

Institute for Logic, Language and Computation

Amsterdam, Netherlands. September 2019 - September 2023.

#### **University of Gothenburg**

Invited research visit, Language Technology Research Lab.

Gothenburg, Sweden. November 2022.

#### **University of Aberdeen**

Research visit, School of Natural and Computing Sciences.

Aberdeen, UK. October 2022.

#### **University of Oslo**

Invited research visit, Language Technology Group,

Oslo, Norway. March 2022.

### MSc Artificial Intelligence

**University of Amsterdam** Amsterdam, Netherlands. 2017-2019.

Honours degree, cum laude. Final grade: 8.88 - ECTS A - First Class

MSc Thesis: *Lexical Semantic Change Analysis with Contextualised Word Representations*. In affiliation with the [Institute for Advanced Study](#).

Grade: 9.5 (top 5% of all Master's theses grades awarded at the University of Amsterdam). ADS Thesis Award 2020.

---

---

## **BA Computational Linguistics**

**Eberhard Karls Universität Tübingen** Tübingen, Germany. 2014-2017.

Final grade: 1.2 - ECTS A - First Class.

BA Thesis. *Semi-supervised emotion lexicon expansion with label propagation and specialized word embeddings*. Grade 1.0.

**French Generate Certificate of Education** (18/20), Rome, Italy. 2012-2014

**Italian Generate Certificate of Education** (100/100), Rome, Italy. 2009-2014.

## **Experience**

---

### **Work Experience**

#### **PhD Candidate, University of Amsterdam — 2019-2023**

I am a final year PhD candidate at the Institute for Logic, Language, and Computation, where I am part of the DREAM (Distributed dynamic REpresentations for diAlogue Management) ERC project and member of the Dialogue Modelling Group led by Raquel Fernández.

#### **NLP Intern, IBM Research and Development, Germany — 2016-2017**

I developed the full sentiment analysis and information extraction Italian pipeline for the IBM Watson Analytics for Social Media system.

#### **Research Assistant, University of Tübingen — 2015-2016**

I worked in the Theoretical Computational Linguistics Group with Prof Dr Walt Detmar Meurers, developing an app for second language acquisition.

### **Project coordination and supervision**

#### **Project coordinator, GenBench, 2021-2022**

I coordinated a team of 20 researchers (PhD students, faculty, and research scientists) for the realisation of the GenBench taxonomy and survey. I made key contributions to the design and write up of the taxonomy, as well as to the managing of our evolving survey of generalisation studies in NLP.

#### **MSc thesis supervisor, University of Amsterdam, 2021-present**

I supervise 3 student projects, respectively on language model **interpretability**, **voice conversion**, and **semantic change** modelling. The first has resulted in a published conference paper; the second in commercial technology; and the third is ongoing. I help the students develop project proposals and monitor their successful execution.

#### **Member of the PhD Council, Faculty of Science — 2020-2022**

I represented the Institute for Logic, Language, and Computation in the PhD Council of the Faculty of Science of the University of Amsterdam. As a council member, I have organised and coordinated **InterAction 2022**, where I moderated a panel discussion with invited panelists from and outside

---

---

academia. This event was designed for PhD students to learn through successful examples of interdisciplinary research and to foster collaboration.

#### **Workshop Organiser, BlackboxNLP, at EMNLP 2021**

I worked as an organiser of the 4th edition of BlackboxNLP, a well-established workshop on **interpretability** and **explainability** of machine learning systems for NLP, which took place in a hybrid format at EMNLP 2021. I contributed by designing the hybrid programme, leading the online workshop sessions, and acting as a meta-reviewer.

#### **Project coordinator, University of Amsterdam, 2021**

I coordinated a team of 11 Bachelor, Master, and PhD students leading to 3 (out of 3) accepted task submissions (Fantasy Reasoning, Metaphor Understanding, and Implicit Relations) to the Beyond The Imitation Game Benchmark (Big-bench), a collaborative benchmark intended to probe large language models and extrapolate their future capabilities. My job was to help the students develop their ideas into actionable research projects and to manage their successful execution.

### **Teaching**

#### **Course Coordinator, MSc AI, University of Amsterdam — 2020-2021**

I coordinated teaching assistants and I was responsible for lab assignments during two editions of the MSc AI course *Natural Language Processing 1*.

#### **Teaching Assistant, MSc AI and Logic, University of Amsterdam — 2020**

I held seminar sessions and guided final projects for *Computational Dialogue Modelling*, an advanced course offered to MSc AI and MSc Logic students. I participated in this role for two consecutive editions of the course.

#### **Teaching Assistant, MSc AI, University of Amsterdam — 2018-2019**

I held lab sessions and corrected coding and theory assignments for the MSc AI course *Natural Language Processing 1*, obtaining an average student feedback of 8.0 in 2018 and 9.0 in 2019.

#### **Teaching Assistant, BA ComplLing, University of Tübingen — 2015-2016**

I held lab sessions and corrected assignments for two courses of the BA Computational Linguistics: *Data Structures and Algorithms I* and *II* (in Java).

### **Awards and scholarships**

---

#### **ADS Master Thesis Award — 2020**

The Amsterdam Data Science Thesis Awards promotes excellence in Data Science and AI from students in Amsterdam-based knowledge institutes.

#### **IAS Excellence Student — 2019**

The Institute for Advance Study of the University of Amsterdam offered me the chance to work on my M thesis alongside renowned researchers and talented early-career scholars.

---

---

### **Best Paper Award at BlackBoxNLP — 2018**

Best Paper Award awarded at the Workshop on *Analyzing and Interpreting Neural Networks for NLP* (BlackBoxNLP), held in Brussels at EMNLP 2018, one of the main international NLP venues.

### **Amsterdam Science Talent Scholarship — 2017-2019**

The Amsterdam Science Talent Scholarship selects students on the basis of their academic excellence and promise in the proposed field.

### **Deutschlandstipendium — 2016-2017**

The *Deutschlandstipendium* supports highly talented students in German universities, without consideration of income and nationality.

### **Certificat d'Inscription au Tableau d'Honneur — 2014**

Acknowledgement for advanced knowledge of French by the *Association des Membres de l'Ordre des Palmes Académiques*.

---

## **Languages and Technologies**

### **Natural languages**

Fluent in Italian and English, advanced in French and German, beginner in Dutch and Latin.

### **Programming languages**

Advanced in Python, proficient in Java, intermediate in Matlab, Prolog, Netlogo, R, JAGS, JavaScript.

### **Python libraries I use nearly every day**

NumPy, SciPy, scikit-learn, Pandas, PyTorch, HuggingFace, NLTK, Spacy, matplotlib, seaborn, plotly

### **Markup languages**

HTML, CSS, XML, XSLT, XPath

### **Phonetic analysis**

Praat

---

## **Invited talks**

**Language Technology Research Lab.** University of Gothenburg.

15 November 2022. Gothenburg, Sweden.

**Inria & Loria.** 27 October 2022. Nancy, France.

**School of Natural and Computing Sciences.** University of Aberdeen.

12 October 2022. Aberdeen, Scotland.

**Advanced Information Retrieval (MSc AI and Informatics).** University of Amsterdam. 21 September 2022. Amsterdam, Netherlands.

**Language Technology Seminar.** Language Technology Group. University of Oslo. 14 March 2022. Norway.

---

---

**NLPitch.** Institute for Logic, Language and Computation. University of Amsterdam. 26 October 2020. Amsterdam, Netherlands.

**AILC Lectures on Computational Linguistics.** Università Tor Vergata. 17 June 2021. Rome, Italy.

**Cognitive Machine Learning Lab.** Ecole Normale Supérieure. 29 June 2021. Paris, France.

**Symposium on Meaning Variation in Social Contexts.** Institute for Logic, Language and Computation. University of Amsterdam. 5 November 2020.

**Computational Linguistic Seminar.** Institute for Logic, Language and Computation. University of Amsterdam. 5 October 2021.

**NLP 1 (MSc AI course).** University of Amsterdam. 18 November 2020. Online.

**Cognitive Science & Artificial Intelligence PhD Group.** University of Tilburg. 1 May 2020. Tilburg, Netherlands.

**Cool Logic Seminar.** Institute for Logic, Language and Computation. University of Amsterdam. 1 February 2019.

## Publications

---

### Book chapters

Nina Tahmasebi, Andrey Kutuzov, Mario Giulianelli. (2023). **Computational models of semantic change.** In A. Ledgeway, E. Aldridge, A. Breitbarth, K.E. Kiss., J. Salmons and A. Simonenko (Eds.). *The Wiley Blackwell Companion to Diachronic Linguistics*. John Wiley & Sons.

### Journal articles

Mario Giulianelli and Daniel de Kok. (2018). **Semi-supervised emotion lexicon expansion with label propagation.** *Computational Linguistics in the Netherlands Journal* 8.

### Peer-reviewed conference papers

Mario Giulianelli. (2022). **Towards Pragmatic Production Strategies for Natural Language Generation Tasks.** To appear in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)*.

Mario Giulianelli, Arabella Sinclair, Raquel Fernández. (2022). **Construction Repetition Reduces Information Rate in Dialogue.** To appear in *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (AACL-IJCNLP 2022)*.

Samuel Ryb, Mario Giulianelli, Arabella Sinclair, Raquel Fernández. (2022). **AnaLog: Testing Analytical and Deductive Logic Learnability in Language Models.** In *Proceedings of \*SEM 2022: The 11th Joint Conference on Lexical and Computational Semantics*.

---

---

Mario Giulianelli, Arabella Sinclair, Raquel Fernández. (2021). **Is Information Density Uniform in Task-Oriented Dialogues?** In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)*.

Mario Giulianelli and Raquel Fernández. (2021). **Analysing Human Strategies of Information Transmission as a Function of Discourse Context.** In *Proceedings of the 25th Conference on Computational Natural Language Learning (CoNLL 2021)*.

Mario Giulianelli, Andrey Kutuzov, Lidia Pivovarov. (2021). **Grammatical Profiling for Semantic Change Detection.** In *Proceedings of the 25th Conference on Computational Natural Language Learning (CoNLL 2021)*.

Mario Giulianelli, Marco Del Tredici, and Raquel Fernández. (2020). **Analysing Lexical Semantic Change with Contextualised Word Representations.** In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*

Ece Takmaz, Mario Giulianelli, Sandro Pezzelle, Arabella Sinclair and Raquel Fernández. (2020). **Refer, Reuse, Reduce: Generating Subsequent References in Visual and Conversational Contexts.** In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)*

### Peer-reviewed workshop papers

Mario Giulianelli, Andrey Kutuzov, Lidia Pivovarov. (2022). **Do Not Fire the Linguist: Grammatical Profiles Help Language Models Detect Semantic Change.** In *Proceedings of the 3rd Workshop on Computational Approaches to Historical Language Change (LChange'22)*.

Andrey Kutuzov and Mario Giulianelli. (2020). **UiO-UvA at SemEval-2020 Task 1: Contextualised Embeddings for Lexical Semantic Change Detection.** In *Proceedings of the 14th International Workshop on Semantic Evaluation*.

Mario Giulianelli, Jacqueline Harding, Florian Mohnert, Dieuwke Hupkes, and Willem Zuidema. (2018). **Under the Hood: Using Diagnostic Classifiers to Investigate and Improve how Language Models Track Agreement Information.** In *Proceedings of the 1st Workshop on Analyzing and Interpreting Neural Networks for NLP (BlackBoxNLP)*. [Best paper award](#).

### Preprints

Dieuwke Hupkes, Mario Giulianelli, Verna Dankers, Mikel Artexte, Tiago Pimentel, et al. (2022). **State-of-the-art generalisation research in NLP: A taxonomy and review.** *Preprint. Presubmitted to [Nature Machine Intelligence](#)*.

Aarohi Srivastava, ..., Mario Giulianelli, and others. (2022). **Beyond the Imitation Game: Quantifying and Extrapolating the Capabilities of Language Models.** *Preprint. Under (single-blind) submission in [Transaction of Machine Learning Research](#)*.

---

---

## Service and professional activities

---

**Organising:** BlackBoxNLP 2021 (EMNLP Workshop), REST-CL (a scientific retreat for computational linguistics PhD students in Europe), InterAction 2022 (an event to foster interdisciplinary collaborations for science PhD students of the University of Amsterdam).

**Chairing:** Workshop on Computational Approaches to Historical Language Change (L'Change) at ACL 2022.

**Reviewing:** ACL, EMNLP, ACL ARR, Transactions on Audio, Speech and Language Processing, BlackBoxNLP, L'Change, CLIN Journal, EACL Student Research Workshop.

---