

# Aina Garí Soler

*Postdoctoral Researcher in Natural Language Processing*

45 avenue du Président Wilson  
91120 Palaiseau, France  
☎ +33 6 48 44 21 81  
✉ [gari912@gmail.com](mailto:gari912@gmail.com)  
📁 [ainagari.github.io](https://github.com/ainagari)

## Experience

- Oct 2024 – **Postdoctoral Researcher**, *INRIA*, Paris, France.  
Sep 2025 Automatic analysis of alignment between speakers in conversations (SINNet ANR project) with Chloé Clavel and Matthieu Labeau.
- Sep 2021 – **Postdoctoral Researcher**, *Télécom-Paris*, Palaiseau, France.  
Sep 2024 Word usage in interaction and word meaning negotiation, with Chloé Clavel and Matthieu Labeau.
- Mar 2017 – **Research Assistant**, *Pompeu Fabra University*, Barcelona, Spain.  
Nov 2017 Collaboration in the project "Language, Deixis and the Disordered Mind" (FFI2013-40526-P) performing statistical analyses.
- Dec 2015 – **Student Assistant**, *Saarland University*, Saarbrücken, Germany.  
Aug 2016 Collaboration in the project "Information Density and Linguistic Encoding" (C1) correcting the automatic segmentation of recordings into speech sounds with Praat.
- Oct 2014 – **Corpus Annotator**, *University of Barcelona*, Barcelona, Spain.  
Jun 2015 Manual syntactic annotation of a corpus of Mexican Spanish.
- Oct 2012 – **Student Assistant**, *University Institute of Applied Linguistics (Pompeu Fabra University)*, Barcelona, Spain.  
Jun 2013 Internship in the *Neologisms*, *NLP* and *Forensic Linguistics* departments.

## Education

- Nov 2017 – **PhD in Computer Science**, *Université Paris Saclay, LISN (former LIMSI), CNRS*, Orsay, France.  
Jun 2021 PhD student in Multilingual Semantic Processing in the MULTISEM ANR project (ANR-16-CE33-0013). **Thesis title:** [Word Meaning Representation in Neural Language Models: Lexical Polysemy and Semantic Relationships](#). **Supervisors:** Marianna Apidianaki and Alexandre Allauzen.
- 2015 – 2017 **Erasmus Mundus European Masters Program in Language and Communication Technologies**.
- **Master's thesis:** [Modelling Differences between Near-Synonyms and its Application to Paraphrasing with Formality](#). **Supervisors:** Lonneke van der Plas and Dietrich Klakow.
  - 2016 – 2017 M.Sc. in Human Language Science and Technology, *University of Malta*, Msida, Malta. Classification: Distinction.
  - 2015 – 2016 M.Sc. in Language Science and Technology, *Saarland University*, Saarbrücken, Germany. GPA: 1.3 (excellent).

- 2010 – 2015 **Bachelor's Degree in Linguistics**, *University of Barcelona*, Barcelona, Spain.
- Majors in Russian and German, 13 distinctions, GPA: 9.2 / 10.
  - Eranet-Mundus Mobility Scholarship: 6 month academic exchange in Saint Petersburg State University, Saint Petersburg (Russia)

## Teaching and supervision

- 2021 – 2024 **Group project supervision** for the IA717 NLP class at Télécom-Paris (M1 Mastère Spécialisé en IA). Design of a project of Semantic Textual Similarity and guided supervision of 2-3 student groups per year.
- 2021 – 2024 Provided more than 60 hours of practical sessions (TP) and 4.5 hours of theoretical courses in NLP at Télécom-Paris and École Polytechnique, mostly at the Master level.
- 2021 **Master Thesis co-supervision** (30%) with Marianna Apidianaki. “Representations of troponymy in BERT” by Dmitry Narkevich, for the M.A. in Language Technology of the University of Helsinki.

## Related scientific experience

- **Senior Area Chair** of Lexical Semantics for COLING 2025.
- **Area Chair** of (Lexical) Semantics (and Pragmatics) for LREC-COLING 2024, NAACL 2024, ACL 2024, CoNLL 2025.
- **Reviewer** at ACL (2021, 2023), EMNLP (2021, 2022, 2023), EACL (2021), RANLP (2019, 2021, 2023), SemEval (2020), DeeLIO (2022), SRW@ACL (2023, 2024), SRW@EACL (2021,2023), RJC (2023), NLP4DH-IWCLUL (2024), SICon (2023, 2024), UTTER FSTP 2024, TAFFC, ARR.
- **Outstanding reviewer** at EACL 2021 and ACL 2021.
- **Invited talks** at the Informatics Department of the University of Turin (May 2024), Jozef Stefan Institute in Ljubjana, Slovenia (Jul 2022), the [FoTran group at the University of Helsinki \(Feb 2020\)](#) and the NLP group at the University of Pennsylvania (Mar 2018).
- **Co-organizer** of the [Social Influence in Conversations \(SICon 2024\) workshop](#), co-located with EMNLP.
- **Co-organizer** of the [2024 ELLIS Doctoral Symposium](#).
- Attendance to **Summer Schools**: ESSLLI 2016, ESSLLI 2017, LxMLS 2018.

## Awards

- Jul 2023 **Outstanding Paper Award** at the SICon 2023 Workshop for the paper “Measuring Lexico-Semantic Alignment in Debates with Contextualized Word Representations”.
- Jun 2016 Special Award of the Bachelor's Degree in Linguistics at the University of Barcelona (**best academic record**).

Jun 2015 Eva Monrós Award for the **best final project** in the Bachelor's Degree in Linguistics, “*Quantitative approaches to meaning*” (*Aproximacions quantitatives al significat*, in Catalan), Supervisors: M. Antònia Martí and Mariona Taulé.

## Skills

- Python (scikit-learn, PyTorch, Huggingface Transformers, NLTK), R, Unix Shell
- Windows, Linux
- MS Office, LaTeX

## Languages

- Catalan, Spanish: native languages
- English: C1 (8.0 / 9.0 IELTS 2015)
- Russian: B2 (TRKI-II 2014)
- French: B2, German: B1, Italian: B1

## Publications

- **Aina Garí Soler**, Jenny Myrendal, Chloé Clavel and Staffan Larsson (2025). **The NeWMe Corpus: A gold standard corpus for the study of Word Meaning Negotiation**. Submitted to Language Resources and Evaluation (LRE).
- **Aina Garí Soler**, Matthieu Labeau and Chloé Clavel (2024). **The Impact of Word Splitting on the Semantic Content of Contextualized Word Representations**. Transactions of the Association for Computational Linguistics (TACL), 12 299–320.
- **Aina Garí Soler** (2024). **The Use of the Grimacing Face Emoji and its Relation to Bruxism and Temporomandibular Joint Disorders**. OSF Preprints.
- **Aina Garí Soler**, Matthieu Labeau and Chloé Clavel (2023). **Measuring Lexico-Semantic Alignment in Debates with Contextualized Word Representations**. In Proceedings of the First Workshop on Social Influence in Conversations (SICon 2023), Toronto, Canada, July 14.
- **Aina Garí Soler**, Matthieu Labeau and Chloé Clavel (2023). **Un mot, deux facettes : traces des opinions dans les représentations contextualisées des mots**. In Actes de la 30e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 1 : conférence principale, Paris, France, June 5-8.
- **Aina Garí Soler**, Matthieu Labeau and Chloé Clavel (2022). **One Word, Two Sides: Traces of Stance in Contextualized Word Representations**. In Proceedings of the 29th International Conference on Computational Linguistics (COLING), Gyeongju, Korea, October 12-17.
- **Aina Garí Soler**, Matthieu Labeau and Chloé Clavel (2022). **Polysemy in Spoken Conversations and Written Texts**. In Proceedings of the 13th Language Resources and Evaluation Conference (LREC), Marseille, France, June 20-25.
- Marianna Apidianaki and **Aina Garí Soler** (2021). **ALL Dolphins Are Intelligent and SOME Are Friendly: Probing BERT for Nouns' Semantic Properties and their Prototypicality**. In Proceedings of the Fourth BlackBoxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP, Punta Cana, Dominican Republic and Online, November, 11.

- **Aina Garí Soler** and Marianna Apidianaki (2021). **Let's Play Mono-Poly: BERT Can Reveal Words' Polysemy Level and Partitionability into Senses**. Transactions of the Association for Computational Linguistics (TACL), 12:299–320.
- **Aina Garí Soler** and Marianna Apidianaki (2021). **Scalar Adjective Identification and Multilingual Ranking**. In Proceedings of the 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2021), Mexico City, Mexico, June 6-11.
- **Aina Garí Soler** and Marianna Apidianaki (2020). **BERT Knows Punta Cana is not just beautiful, it's gorgeous: Ranking Scalar Adjectives with Contextualised Representations**. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Nov 16-20.
- **Aina Garí Soler** and Marianna Apidianaki (2020). **MULTISEM at SemEval-2020 Task 3: Fine-tuning BERT for Lexical Meaning**. In Proceedings of the 14th International Workshop on Semantic Evaluation, Dec 12-13, Barcelona, Spain.
- **Aina Garí Soler**, Marianna Apidianaki and Alexandre Allauzen (2019). **LIMSI-MultiSem at the IJCAI SemDeep-5 WiC Challenge: Context Representations for Word Usage Similarity Estimation**. In 5th Workshop on Semantic Deep Learning (SemDeep-5)
- **Aina Garí Soler**, Marianna Apidianaki and Alexandre Allauzen (2019). **Word Usage Similarity Estimation with Sentence Representations and Automatic Substitutes**. Proceedings of the 8th Joint Conference on Lexical and Computational Semantics (STARSEM 2019), 6-7 Jun, Minneapolis, USA.
- Syrielle Montariol, **Aina Garí Soler** and Alexandre Allauzen (2019). **Exploring sentence informativeness**. Proceedings of TALN, Jul 1-5 , Toulouse, France.
- **Aina Garí Soler**, Anne Cocos, Marianna Apidianaki and Chris Callison-Burch (2019). **A comparison of context-sensitive models for lexical substitution**. Proceedings of the 13th International Conference on Computational Semantics (IWCS 2019), 23-27 May, Gothenburg, Sweden.
- Antonia Tovar, **Aina Garí Soler**, Jesús Ruiz-Idiago, Celia Mareca Viladrich, Edith Pomarol-Clotet, Joana Rosselló, Wolfram Hinzen (2020). **Language disintegration in spontaneous speech in Huntington's disease: a more fine-grained analysis**. In Journal of Communication Disorders 83 (2020): 105970
- Antonia Tovar Torres, Wolfgang Sebastian Schmeisser Nieto, **Aina Garí Soler**, Catalina Morey Matamalas, Wolfram Hinzen (2019). **Language disintegration under conditions of severe formal thought disorder**. Glossa: A Journal of General Linguistics, 4(1), 134.
- **Aina Garí**, Marianna Apidianaki and Alexandre Allauzen (2018). **A comparative study of word embeddings and other features for lexical complexity detection in French**. In Actes de la Conférence TALN. Volume 1 - Articles longs, articles courts de TALN, May 14-18, Rennes, France.
- Taulé, Mariona, M. Antònia Martí, Ann Bies, **Aina Garí**, Montserrat Nofre, Zhiyi Song, Stephanie Strassel and Joe Ellis (2015). **Spanish Treebank Annotation of Informal Non-Standard Web Text**. Daniel, F. and Díaz O. (Eds.) Proceedings of ICWE-2015, Lecture Notes in Computer Science. Springer Verlag. ISSN 0302-9743
- Muntsa Padró, Núria Bel and **Aina Garí** (2013). **Verb SCF extraction for Spanish with dependency parsing**. Procesamiento del Lenguaje Natural 51, 93-100. September 2013