

Verna Dankers

verna.dankers@mila.quebec

vernadankers.com

Education

- 2020-2025 **PhD at the UKRI Centre for Doctoral Training in Natural Language Processing**,
University of Edinburgh, Scotland
- *Supervisors*: Prof Dr Ivan Titov (*primary supervisor*), Dr Christopher Lucas
 - *Thesis title*: memorisation meets compositionality in natural language processing (successfully defended July 1st, 2025)
 - The CDT in NLP offers a PhD programme with integrated training. The programme was jointly run by the School of Informatics and the School of Philosophy, Psychology and Language Sciences and included a taught component equivalent to 90 ECTS. GPA 81/100
- 2017-2019 **MSc in Artificial Intelligence**, University of Amsterdam, Netherlands
- Graduated *cum laude*, GPA 9.1/10
 - *Master's Thesis Supervisors*: Dr Ekaterina Shutova, Dr Martha Lewis
 - *Research focus*: multitask learning for metaphor detection, published at EMNLP 2019.
- 2014-2017 **BSc in Artificial Intelligence**, University of Amsterdam, Netherlands
- Graduated *cum laude* and *with honours*, GPA 9.3/10
 - *Bachelor's Thesis Supervisors*: Dr Raquel Fernández, Dr Aysenur Bilgin

Employment

- 2025-now **Postdoc at Mila (Quebec AI institute) and McGill University**, Montreal, Canada
- *Research focus*: memorization and generalization's interdependency for LLMs
 - *Supervisor*: Prof Siva Reddy
- 2024 **Research intern at Microsoft Research, in the Machine Translation team**, Seattle, USA
- *Research focus*: Memorisation inheritance in knowledge distillation for machine translation, published at ACL 2025.
 - *Manager*: Vikas Raunak
 - Worked on-site from June to September, was employed remotely from October to December via [Appen](#).
- 2022-2023 **Contingent worker at FAIR labs at Meta, employed via Magnit**, remote work from Edinburgh, Scotland
- 2022 **Research scientist intern at FAIR labs at Meta**, Paris, France
- *Research focus*: Per-datum memorisation in machine translation, published at EMNLP 2023.
 - *Managers*: Dr Dieuwke Hupkes, Dr Brenden Lake

Conference Publications

- 2025 **Memorization inheritance in sequence-level knowledge distillation for neural machine translation**
Verna Dankers and Vikas Raunak. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 760-774 [[paper](#), [code](#)]
- 2024 **Generalisation first, memorisation second? Memorisation localisation for natural language classification tasks**
Verna Dankers and Ivan Titov. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 14348-14366 [[paper](#), [code](#)]

- 2023 **Memorisation cartography: mapping out the memorisation-generalisation continuum in neural machine translation**
Verna Dankers, Ivan Titov and Dieuwke Hupkes. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 8323-8343 [[paper](#), [code](#)]
- 2023 **Non-compositionality in sentiment: new data and analyses**
Verna Dankers and Christopher G. Lucas. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 5150-5162 [[paper](#), [code](#)]
- 2023 **Paper bullets: modeling propaganda with the help of metaphor**
Daniel Rodríguez, Verna Dankers, Preslav Nakov and Ekaterina Shutova. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 472-489 [[paper](#)]
- 2022 **Recursive neural networks with bottlenecks diagnose (non-)compositionality**
Verna Dankers and Ivan Titov. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 4361-4378 [[paper](#), [code](#)]
- 2022 **Can Transformer be too compositional? Analysing idiom processing in neural machine translation**
Verna Dankers, Christopher G. Lucas and Ivan Titov. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3608-3626 [[paper](#), [code](#)]
- 2022 **The paradox of the compositionality of natural language: a neural machine translation case study**
Verna Dankers, Elia Bruni and Dieuwke Hupkes. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4154-4175 [[paper](#), [code](#)]
- 2022 **Meta-learning for fast cross-lingual adaptation in dependency parsing**
Anna Langedijk, Verna Dankers, Phillip Lippe, Sander Bos, Bryan Cardenas Guevara, Helen Yannakoudakis and Ekaterina Shutova. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8503-8520 [[paper](#)]
- 2021 **Episodic memory demands modulate novel metaphor use during event narration**
Vesna Djokic, Verna Dankers and Ekaterina Shutova. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43 [[paper](#)]
- 2021 **Generalising to German plural noun classes, from the perspective of a recurrent neural network**
Verna Dankers*, Anna Langedijk*, Kate McCurdy, Adina Williams and Dieuwke Hupkes. In *Proceedings of the 25th Conference on Computational Natural Language Learning*, pages 94-108. *Equal contribution. [[paper](#), [code](#)] **Best paper award.**
- 2020 **The pragmatics behind politics: modelling metaphor, framing and emotion in political discourse**
Pere-Lluis Huguet Cabot, Verna Dankers, David Abadi, Agneta Fischer and Ekaterina Shutova. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 4479-4488 [[paper](#)]
- 2019 **Modelling the interplay of metaphor and emotion through multitask learning**
Verna Dankers, Marek Rei, Martha Lewis and Ekaterina Shutova. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 2218-2229 [[paper](#)]

Journal Publications

- 2023 **A taxonomy and review of generalization research in NLP**
Dieuwke Hupkes, Mario Giulianelli, Verna Dankers, Mikel Artetxe, Yanai Elazar, Tiago Pimentel, Christos Christodoulopoulos, Karim Lasri, Naomi Saphra, Arabella Sinclair, Dennis Ulmer, Florian Schottmann, Khuyagbaatar Batsuren, Kaiser Sun, Koustuv Sinha, Leila Khalatbari, Maria Ryskina, Rita Frieske, Ryan Cotterell and Zhijing Jin. In *Nature Machine Intelligence, volume 5*, pages 1161-1174 [[paper](#)]

2020 **Compositionality decomposed: How do neural networks generalise?**
Dieuwke Hupkes, **Verna Dankers**, Mathijs Mul and Elia Bruni. In *Journal of Artificial Intelligence*, 67, pages 757–795 [[paper](#), [extended abstract](#)]
Outstanding paper award, awarded after five years at IJCAI-2025.

Workshop Publications

2023 **Latent feature-based data splits to improve generalization evaluation: A hate speech detection case study**
Maike Züfle, **Verna Dankers** and Ivan Titov. In *GenBench: The first workshop on generalisation (benchmarking) in NLP*, p. 112 [[paper](#), [code](#)]

2020 **Being neighbourly: Neural metaphor identification in discourse**
Verna Dankers, Karan Malhotra, Gaurav Kudva, Volodymyr Medentsiy, and Ekaterina Shutova. In *Proceedings of the Second Workshop on Figurative Language Processing*, pages 227–234 [[paper](#)]

2019 **Transcoding compositionally: Using attention to find more generalisable solutions**
Kris Korrel, Dieuwke Hupkes, **Verna Dankers** and Elia Bruni. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 1–11 [[paper](#)]

2017 **Modelling word associations with word embeddings for a guesser agent in the Taboo city challenge competition**
Verna Dankers, Aysenur Bilgin, and Raquel Fernández. In *The 6th ESSENCE Workshop: the Taboo city challenge competition @ IJCAI-2017* [[paper](#)]

Preprints

2024 **Evaluating subword tokenization: Alien subword composition and OOV generalization challenge**
Khuyagbaatar Batsuren, Ekaterina Vylomova, **Verna Dankers**, Tsetsuukhei Delgerbaatar, Omri Uzan, Yuval Pinter, and Gábor Bella. *Preprint* [[paper](#)]

Service

Workshop organisation

2025 [L2M2](#): The first workshop on Large Language Model Memorization @ ACL 2025
2024 [GenBench](#): The second workshop on generalisation (benchmarking) in NLP @ EMNLP 2024
2023 [GenBench](#): The first workshop on generalisation (benchmarking) in NLP @ EMNLP 2023

Area chair

2026 ACL, CoNLL
2025 NAACL, EMNLP
2024 EACL, ACL

Reviewer

2025 NeurIPS Mechanistic Interpretability workshop, BlackboxNLP workshop, IEEE journal
2024 FigLang workshop, Computational Linguistics journal
2023 ACL (outstanding reviewer award), CoNLL, EMNLP, ARR, BlackboxNLP, Computational Linguistics journal
2022 ACL, EMNLP, ARR, BlackboxNLP workshop, FigLang workshop
2021 CoNLL (outstanding reviewer), ARR, BlackboxNLP workshop

Honours & Awards

- 2025 JAIR outstanding paper award, awarded at IJCAI 2025, Montreal, Canada
- 2023 Outstanding reviewer award at ACL 2023, Toronto, Canada
- 2022 Staff award for the role of teaching assistant in the FNLP course, University of Edinburgh, Scotland
- 2021 The Best Paper Award, awarded at the CoNLL conference, Punta Cana, Dominican Republic
- 2020-2024 PhD Funding, UKRI CDT in NLP, University of Edinburgh, Scotland
- Full scholarship covering tuition fees, programme costs and stipend (approx. £80,000 for 4 years).
- 2017-2019 ASML Technology Scholarship, Veldhoven, Netherlands
- Awarded a €10,000 scholarship and participated in a two-year professional development programme
- 2017 ESSENCE Workshop: Taboo city challenge competition @ IJCAI, 2nd price, Melbourne, Australia

Teaching Experience

- 2025-now **Guest lecturer McGill University**
- Guest lecture on interpretability for *Natural Language Understanding with Deep Learning / Computational Semantics Siva Reddy* (COMP 545; LING 545)
 - Hosting seminars on memorization for *Large Language Models* (COMP 767; LING 782)
- 2021-2024 **Teaching Support at the School of Informatics**, University of Edinburgh, Scotland
- *Marker* for the Researching Responsible and Trustworthy Natural Language Processing course for first-year PhD students.
 - *Marker* for the Doing Research in NLP course for first-year PhD students.
 - *Teaching Assistant* for the undergraduate course Foundations of NLP.
 - *Tutor* for the Accelerated Natural Language Processing postgraduate course.
- 2019-2020 **Teaching Assistant MSc programme Artificial Intelligence**, University of Amsterdam, Netherlands
- *Courses taught*: Statistical Methods of Natural Language Semantics, Advanced Techniques of Computational Linguistics, Natural Language Processing 2.
 - *Responsibilities*: independently supervised research projects during courses, co-supervised MSc dissertations, gave guest lectures on ‘multitask learning’.
- 2015-2018 **Teaching Assistant BSc programme Artificial Intelligence**, University of Amsterdam, Netherlands
- *Courses taught*: Introduction to Logic, Introduction to Artificial Intelligence, Linguistics and Language Processing, Linear Algebra, Fundamentals of Fuzzy Logic.
 - *Responsibilities*: organised weekly tutorials for 20 students, graded homework and exams.

Supervision

- 2026 Matteo Boglioni. I co-supervise Matteo’s internship at Mila, on benchmarking unlearning methods in LLMs.
- 2025 Márk Bodrácška. I co-supervise Márk’s MSc in AI thesis project about memorisation localisation in LLMs.
- 2023 Maike Züfle. I co-supervised Maike’s MSc in AI thesis project (published at the GenBench 2023 workshop).
- 2021 Daniel R. Baleato. I co-supervised Daniel’s MSc in Data Science thesis project (published at EACL 2023).
- 2019 Karan Malhotra, Gaurav Kudva, Volodymyr Medentsiy. I supervised their MSc in AI research project (published at the Figlang 2020 workshop).
- 2019 Anna Langedijk, Sander Bos, Bryan Cardenas Guevara. I supervised their MSc in AI research project (published at ACL 2022).

Talks & Panels

- 2026 *LLM memory: from monoliths to modularity*, Mila Conference 2026
- 2025 *Memorization: myth or mystery?*, IVADO Workshop on Deploying Autonomous Agents: Lessons, Risks, and Real-World Impact
- Memorization: myth or mystery?*, BlackboxNLP keynote at EMNLP 2025
- How to write NLP reviews and rebuttals*, an ACL mentorship panel at ACL 2025

Analysing memorisation in classification and translation through localisation and cartography,
Invited Talk, IT University of Copenhagen

Analysing memorisation in classification and translation through localisation and cartography,
Invited Talk, Mila/McGill University

2024

Analysing memorisation in classification and translation through localisation and cartography,
Computational Linguistics Seminar, University of Amsterdam

Memorisation in translation and classification: the what and the where, InDeep Seminar, Tilburg University

2023

Memorisation meets non-compositionality, Institute for Natural Language Processing, University of Stuttgart

Idiom processing in Transformer, a translation case study, CardiffNLP seminar, Cardiff University

Memorisation maps for neural machine translation, Sheffield CDT conference, Sheffield University

Idiom processing in Transformer, a translation case study, NLP Research Group, Sheffield University

Competences

- **Technical Skills:** Python (*advanced*), Latex (*advanced*), Matlab (*intermediate*), Prolog (*beginner*), Java (*beginner*), C++ (*beginner*), R (*beginner*), Bash (*beginner*)
- **Languages:** Dutch (*native*), English (*fluent*, TOEFL 119/120), German (*intermediate*), French (*pre-intermediate*)

Community Building & Outreach

I think the role of a researcher encompasses more than contributing to teaching and research, and also involves providing **peer support**, **mentoring** and contributing to **public engagement**.

2025-now

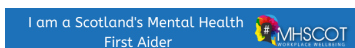
Throughout my time in Montreal, I volunteered as a mentor for the following programmes:

- The *Gender Equity Mentoring in AI (GEMinAI)* Program;
- The *Women in Machine Learning (WiML)* mentorship programme.

2020-2024

During my time in Edinburgh, I contributed to the following outreach activities:

- For two years in a row, I co-organised the CDTalks event at my university, a monthly NLP seminar for CDT in NLP students. I also co-organised a writing retreat for PhD students.
- I was a Mental Health First Aider at my university, qualified from April 2021 to April 2024 by MHScot.



- I was a STEM ambassador in Scotland, and as an ambassador, I volunteered at [DataKirk](#) between September 2021 and June 2023. DataKirk is an organisation with the mission to introduce children and young adults of ethnic minorities to data science and coding to improve data literacy. During school seasons, I hosted the weekly [Data Literacy Club](#) for children aged 8 - 13, about programming in Python, data science and basic artificial intelligence.
- In December 2020, and throughout the year of 2024, I was a part of the "[I'm a Scientist](#)" initiative, which connects school classes with scientists for Q&A sessions.

Last updated: February 9, 2026