

Marcell Richard Fekete

✉ mrfe@cs.aau.dk | 📞 +4525311995

SUMMARY

My main research interests include multilinguality, language variation, and typology, parameter-efficient fine-tuning methods, computational linguistics and interpretability, and the comparison of human and AI language knowledge. My work is supported by my competence in natural language processing and my strong background in theoretical linguistics.

EXPERIENCE

PhD Fellow at Aalborg University Sep 2022 – Present

PhD funded by the Carlsberg Foundation project *Multilingual Modelling for Resource-Poor Languages*, supervised by Prof. Johannes Bjerva and Heather Lent. Also supervising bachelors and masters students in Software Design.

Junior Machine Learning Engineer at TAUS Sep 2021 – Jun 2022

Research projects *Dimensionality Reduction of Multilingual Sentence Embeddings Using Autoencoders* and *Cross-lingual Transfer of Multilingual Language Models Using Stacked Language Adapters*.

NLP Intern at Underlined Mar 2021 – Sep 2021

Research into automatic topic modelling using Latent Dirichlet Allocation.

EDUCATION

2020 – 2022 MA (Human Language Technology) at **Vrije Universiteit Amsterdam** *cum laude*

2015 – 2018 BA (Linguistics) at **University of Cambridge** Upper Second-Class degree

GRANTS

DFE International Postdoc Grant (2.140.857 DKK) Nov 2026 – Oct 2028

Personal grant awarded by the Independent Research Fund Denmark (Danmarks Frie Forskningsfond).

Otto Mønstedts Fonden Conference Participation (7,500 DKK) Sep 2025

OM Fonden funding for conference attendance at ACL in Vienna, Austria.

UniDive COST Action Research Visit (2,000 EUR) Feb 2025 – Apr 2025

UniDive funded research visit to the **HUN-REN Hungarian Research Centre for Linguistics**.

Otto Mønstedts Fonden Conference Participation (7,500 DKK) Jun 2025

OM Fonden funding for conference attendance at NAACL in Albuquerque, New Mexico.

Otto Mønstedts Fonden Research Stay (13,500 DKK) Mar 2024 – Jun 2024

OM Fonden funded research visit to the **University of Edinburgh**.

DISSEMINATION

Neural Networks (Workshop) Aug 2025

Introducing high school students to the basic principles of neural networks in Diósjenő, Hungary.

- Analysing Language Model Knowledge using Linguistic Theory (Invited Talk)** Feb 2025
Talk at the Department of Computer Science of the University of Göttingen, invited by Prof. Dr. Lisa Beinborn.
- Do Language Models Dream With Linguistics? (Presentation)** Dec 2024
Presentation at the AAU NLP Symposium 2024.
- Machine Intelligence (Lecture)** Nov 2024
Guest lecture about large language modelling for P5 Software Design at Aalborg University.
- Chatbots (Lecture)** Nov 2024
Introducing high school students to shortcomings of chatbots at FalconNXT.
- Chatbots and Large Language Models (Workshop)** Aug 2024
Activity for high school students regarding chatbots and the inner workings of large language models in Diósjenő, Hungary.
- Language Modelling (Lab)** Nov 2023
Guest lecture and lab about large language modelling for P5 Software Design at Aalborg University.

PUBLICATIONS

- Marcell Fekete**, Johannes Bjerva (2025). “Linguistically Grounded Analysis of Language Models using Shapley Head Values”. In: *Findings of the Association for Computational Linguistics: NAACL 2025*. Albuquerque, New Mexico: Association for Computational Linguistics, pp. 850–865. ISBN: 979-8-89176-195-7. DOI: [10.18653/v1/2025.findings-naacl.49](https://doi.org/10.18653/v1/2025.findings-naacl.49).
- Marcell Fekete**, Nathaniel Romney Robinson, Ernests Lavrinovics, Djeride Jean-Baptiste, Raj Dabre, Johannes Bjerva, Heather Lent (2025). “Limited-Resource Adapters Are Regularizers, Not Linguists”. In: *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Vienna, Austria: Association for Computational Linguistics, pp. 222–237. ISBN: 979-8-89176-252-7. DOI: [10.18653/v1/2025.acl-short.19](https://doi.org/10.18653/v1/2025.acl-short.19).
- Tomer Sagi, Moran Zaga, Sinai Rusinek, **Marcell Fekete**, Johannes Bjerva, Katja Hose (2025). “Utilizing phonetic similarity for cross-source and cross-language toponym matching: a benchmark and prototype”. In: *Language Resources and Evaluation* 59.3, pp. 2427–2451. DOI: [10.1007/s10579-025-09812-9](https://doi.org/10.1007/s10579-025-09812-9).
- Marcell Fekete**, Ernests Lavrinovics, Nathaniel Romney Robinson, Heather Lent, Raj Dabre, Johannes Bjerva (2024). “Leveraging Adapters for Improved Cross-lingual Transfer for Low-Resource Creole MT”. In: *Proceedings of the Fourth Workshop on Multilingual Representation Learning (MRL 2024)*. Ed. by Jonne Sälevä and Abraham Owodunni. Miami, Florida, USA: Association for Computational Linguistics, pp. 212–215. DOI: [10.18653/v1/2024.mrl-1.17](https://doi.org/10.18653/v1/2024.mrl-1.17). URL: <https://aclanthology.org/2024.mrl-1.17/>.
- Heather Lent, Kushal Tatariya, Raj Dabre, Yiyi Chen, **Marcell Fekete**, Esther Ploeger, Li Zhou, Ruth-Ann Armstrong, Abee Eijansantos, Catriona Malau, Hans Erik Heje, Ernests Lavrinovics, Diptesh Kanojia, Paul Belony, Marcel Bollmann, Loïc Grobol, Miryam de Lhoneux, Daniel Hershcovich, Michel DeGraff, Anders Søgaard, Johannes Bjerva (2024). “CreoleVal: Multilingual Multitask Benchmarks for Creoles”. In: *Transactions of the Association for Computational Linguistics* 12, pp. 950–978. DOI: [10.1162/tacl_a_00682](https://doi.org/10.1162/tacl_a_00682). URL: <https://aclanthology.org/2024.tacl-1.53/>.
- Marcell Fekete**, Johannes Bjerva (2023). “Gradual Language Model Adaptation Using Fine-Grained Typology”. In: *Proceedings of the 5th Workshop on Research in Computational Linguistic Typology and Multilingual NLP*. Dubrovnik, Croatia: Association for Computational Linguistics, pp. 153–158. DOI: [10.18653/v1/2023.sigtyp-1.19](https://doi.org/10.18653/v1/2023.sigtyp-1.19).
- Marcell Fekete**, Johannes Bjerva, Lisa Beinborn (2023). “Typological Challenges for the Application of Multilingual Language Models in the Digital Humanities”. In: *Multilingual Digital Humanities*. Ed. by Lorella Viola and Paul Joseph Spence. Routledge. Chap. 9, pp. 145–164. DOI: [10.4324/9781003393696-13](https://doi.org/10.4324/9781003393696-13).