Roberta Rocca

Skt Pauls Gade 41, 8000 Aarhus C, Denmark

☎ +45 93 92 03 66 | ✉ rbrrcc@gmail.com | ◉ https://rbroc.github.io | ⊙ rbroc

WORK EXPERIENCE

Tenure-track Assistant Professor (Cognitive Science, Data Science & NLP)	12/2022 – Present
Department of Culture, Cognition, and Computation, Aarhus University	*Aarhus, Denmark*
Postdoctoral Researcher	03/2022 – 12/2022
Department of Culture, Cognition, and Computation, Aarhus University	*Aarhus, Denmark*
Predictive Analytics Data Fellow	06/2021 – 08/2021
Centre for Humanitarian Data, United Nations	*The Hague, The Netherlands (Remote)*
Postdoctoral Researcher	11/2019 – 03/2022
Psychoinformatics Lab, University of Texas at Austin	*Austin, TX*

EDUCATION

PhD in Cognitive Science	09/2016 – 10/2019
Department of Linguistics, Cognitive Science, and Semiotics, Aarhus University	*Aarhus, Denmark*
MSc in Cognitive Science (Language & Multimodal Interaction track)	09/2014 – 07/2016
Centre for Mind/Brain Sciences, University of Trento	*Trento, Italy*
BA in Philosophy	10/2010 – 03/2014
School of Humanities, University of Palermo	*Palermo, Italy*

ADDITIONAL RESEARCH EXPERIENCE

Visiting Researcher	06/2024 – 07/2024
MilaNLP group, Department of Computing Sciences, Bocconi University	*Milan, Italy*
Visiting Researcher	08/2018 – 03/2019
Department of Applied Mathematics and Computer Science, Technical University of Denmark	*Copenhagen, Denmark*
Visiting Researcher	06/2019 – 07/2019
Department of Psychology, University of East Anglia	*Norwich, United Kingdom*
Visiting Researcher	01/2018 and 10/2018
Institute of Cognitive Sciences and Technologies, National Research Council	*Rome, Italy*
Research Assistant	03/2016 – 07/2016
Department of Psychology, University College London	*London, United Kingdom*
Research Assistant	08/2015 – 03/2016
Department of Language and Cognition, Max Planck Institute for Psycholinguistics	*Nijmegen, The Netherlands*

TEACHING AND SUPERVISION

Data Science, Prediction, and Forecasting | *MSc Cognitive Science, Aarhus University*	Spring 2024
Natural Language Processing | *MSc Cognitive Science, Aarhus University*	Fall 2023
Applied Cognitive Science | *BSc Cognitive Science, Aarhus University*	Spring 2023, Spring 2024
Cognitive Neuroscience [Teaching Assistant] | *BSc Cognitive Science, Aarhus University*	Spring 2018
Cognitive Science | *MA Cognitive Semiotics, Aarhus University*	Fall 2017
BSc (4x) and **MSc (3x) thesis supervision** | *Cognitive Science, Aarhus University*	Fall 2023, Spring 2024
PhD co-supervision: Line Elgaard Kruse Danielsen | *Aarhus University*	Fall 2021 - Spring 2024

SKILLS

Expertise: atural Language Processing; Machine Learning; Cognition; Data Science; Statistics; Experimental Methods
Programming: Python, R, SQL, Matlab, bash, PERL, Git/GitHub, CI/CD, Docker
Main Libraries: Torch, TensorFlow/Keras, Transformers, Scikit-Learn, NumPy, SciPy, StatsModels, Pandas
Languages: English (fluent), Italian (native), Danish (fluent), French (intermediate)

PUBLICATIONS (SELECTED)

Full list available on Google Scholar. Hyperlinks are included for each paper (hover over the title).

ML/LLMs for Mental Health:

Rocca, R., Pistilli, G., Maheshwari, K., Fusaroli, R. (forthcoming), Introducing ELLIPS: an ethics-centered approach to research on LLM-based inference of psychiatric conditions, *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society.* Preprint available on ArXiv: arXiv:2409.15323

Hansen, L., **Rocca, R.**, Simonsen, A., Olsen, L., ... & Fusaroli, R. (2023), Speech- and text-based classification of neuropsychiatric conditions in a multidiagnostic setting. *Nature Mental Health*

NLP Methods and Model Evaluation:

Rocca, R., Yarkoni, T. (2022), Language as a fingerprint: Self-supervised learning of user encodings using transformers, *Findings of the Association for Computational Linguistics: EMNLP 2022*

Rocca, R., de la Vega, A. (2022), Evaluating the role of non-lexical markers in GPT-2's language modeling behavior, *Proceedings of the 3rd Workshop on Evaluation and Comparison of NLP Systems (Eval4NLP)*

[under review] Kardós, M., Kostkan, J., Vermillet, A.Q., Nielbo, K., Enevoldsen, K., **Rocca, R.**, S^3 – Semantic Signal Separation, Preprint available on ArXiv: arXiv:2406.09556

[forthcoming] Luo, X., Rechardt, A., Sun, ..., **Rocca, R.**, ..., & Love, B. C.. Large language models surpass human experts in predicting neuroscience results

Research Methods in Cognitive Science:

de la Vega, A., **Rocca, R. (co-first)**, Blair, R., Mentch, J., Markiewicz, C., Ghosh, S., Poldrack, R., Yarkoni, T. (2022), Neuroscout: a unified platform for generalized and reproducible fMRI research, *eLife*

Rocca, R., Yarkoni, T. (2021). Putting psychology to the test: rethinking model evaluation through benchmarking and prediction, *Advances in Methods and Practices in Psychological Science*

Applied NLP & Society:

Rocca, R., Tamagnone, N., Contla, X., Bove, J.B., Rekabsaz, N. (2023), Natural language processing for humanitarian action: challenges, opportunities, and the path towards a humanitarian NLP community, *Frontiers in Big Data*

Rocca, R., Lawall, K., Tsakiris, M., Cram, L. (2024), Communicating Europe: a computational analysis of the evolution of the European Commission's communication on Twitter. *Journal of Computational Social Science*, 1-52.

Applied NLP & Cognition:

[under review] **Rocca, R.** & Tylén, K., The effect of diversity and social interaction on cognitive search: An agent-based simulation. Preprint available on PsyArXiv: https://doi.org/10.31234/osf.io/n3t6j

Fusaroli, R., Weed, E., **Rocca, R.**, Fein, D., Naigles, L. (2023), Caregiver linguistic alignment to autistic and typically developing children: A natural language processing approach illuminates the interactive components of language development, *Cognition*

Human Language Processing:

Coventry, K.R., Gudde, H. B., Diessel, H., Collier, J., ..., **Rocca, R.**, ... & Incel, O.D., (2023), Spatial communication systems across languages reflect universal action constraints. *Nature Human Behavior*

Rocca, R., Coventry, K. R., Tylén, K., Staib, M., Lund, T. E., Wallentin, M. (2020). Language beyond the language system: dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI, *NeuroImage*

Rocca, R., Wallentin, M. (2020). The semantics of spatial demonstratives. *Proceedings of the 42th Annual Meeting of the Cognitive Science Society*

Rocca, R., Wallentin, M., Vesper, C., Tylén, K. (2019). This is for you: social modulations of proximal vs. distal space in collaborative interactions, *Scientific Reports*

Blog posts and technical reports

Rocca, R. (2024), Interpreting R^2: a Narrative Guide for the Perplexed, *Towards Data Science*

Rocca, R. (2021) Complex Systems Modeling for Humanitarian Action: Methods and Opportunities, *Research Report for the United Nations Centre for Humanitarian Data*

Open-source contributions

Co-developer of *Turftopic*, a Python library for scikit-learn compatible transformer-based topic modeling: https://github.com/x-tabdeveloping/turftopic

Co-developer and **maintainer** of *pliers*, Python library for automated multimodal feature extraction using machine learning: https://github.com/PsychoinformaticsLab/pliers

Co-developer and **maintainer** of the *Neuroscout* ecosystem, a suite of tools for fast and reproducible analysis of naturalistic fMRI data: https://github.com/neuroscout

Contributor to the Massive Multilingual Text Embedding Benchmark (MMTEB), a big science initiative for multilingual and multi-task benchmarking of LLM embeddings: https://github.com/embeddings-benchmark/mteb

Code to **reproduce my research papers** is available on GitHub, see for example:

- Rocca & Yarkoni, 2022: https://github.com/rbroc/contrastive-user-encoders
- Rocca & Yarkoni, 2021: https://github.com/rbroc/benchmarks_paper
- Rocca & Tylén, 2024: https://github.com/rbroc/simcat

Talks, Grants and Awards (Selected)

Shortlisted for ERC Starting Grant Interview	May 2024
Recruitment Grant by Aarhus University Research Foundation: 1.949.000 DKK	December 2023
Invited Talk at MIT/Harvard Speech and Language Biomarkers Interest Group	December 2022
Invited Talk at Humanitarian Networks and Partnerships Week (HNPW), Geneva	April 2023
Invited Talk at Joint seminar series: University of Toronto, McGill University, Karolinska Institute	May 2021
University of Trento Merit Award	July 2016