

## Roberta Rocca, PhD, MSc

### Profile

I am a researcher in neuroinformatics with an interdisciplinary background and a strong interest in AI and data science. My work focuses on developing methods and open-source software for psychology and neuroscience, and on making machine learning practices more available to these research communities. As a Data Fellow at the United Nations' Centre for Humanitarian Data, I have also done research on the potential of predictive modeling to inform intervention in humanitarian contexts. I am passionate about deep learning and NLP, and I am looking for opportunities to further develop my skills in these sectors, while contributing to increasing the fairness and positive societal impact of current models and technologies.

✉ [rbrcc@gmail.com](mailto:rbrcc@gmail.com)  
 🌐 [rbroc](#)  
 🌐 [rbroc.github.io](https://rbroc.github.io)  
 🐦 [RockBerta](#)

## Experience

### Postdoctoral Researcher

*December 2019 – ongoing*

Psychoinformatics Lab, University of Texas at Austin, USA

**Focus area:** Neuroinformatics, Deep Learning

#### Main projects

- Contributions to NeuroScout (<https://neuroscout.org>), an open-source platform for re-analysis of neuroimaging data
  - Implementation (and validation) of ML/NLP pipelines for feature extraction using *pliers*
  - Implementation (and validation) of extractors for auditory features
  - Scientific validation of the platform through a series of large-scale meta-analyses (preprint forthcoming)
  - QA testing, documentation, outreach
- Development of context-aware transformers for user-sensitive text encoding
  - Creation of a large-scale dataset of Reddit submissions grouped by user id
  - Design, implementation, and training of three novel DistilBERT-based architectures – a batch encoder, a hierarchical encoder, and a biencoder – supporting context-aware MLM pretraining
  - Evaluation on masked language modeling and triplet-based author discrimination
- Methods development research on ML-based improvements to model evaluation in psychology

### Postdoctoral Researcher

*December 2020 – ongoing*

Interacting Minds Centre, Aarhus University, Denmark

**Focus area:** Computational Social Science, Applied NLP

#### Main responsibilities

- Built a large-scale dataset of multilingual Twitter data to map cross-national attitudes towards the European Union
- Deployed transformer-based topic modeling techniques and sentiment pipelines for data analysis
- Designed large-scale online study aimed at identifying factors mediating political identity in a text-driven fashion

### Predictive Analytics Data Fellow / Consultant

*June - August 2021*

Centre for Humanitarian Data, United Nations

**Focus area:** Complex Systems, Predictive Modeling

#### Main responsibilities

- Researched desirability, feasibility and impact of “complex systems” approaches to predictive modeling of humanitarian crises
- Identified data requirements, modeling techniques, and partnerships for a pilot
- Wrote a [blog](#) and a [an extended report](#) outlining outcomes and recommendations

## Education

### PhD, Cognitive Neuroscience

*September 2016 – October 2019*

Aarhus University, Denmark

### MSc, Cognitive Science

*September 2014 – July 2016*

University of Trento, Italy

## Additional research experience

**Visiting Researcher**, Department of Applied Mathematics and Computer Science; Technical University of Denmark

**Visiting Researcher**, Institute of Cognitive Science and Technologies; National Research Council, Italy

**Research Assistant**, Department of Experimental Psychology, University College London, United Kingdom

**Research Assistant**, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

## Skills

**Programming languages:** Python; R; Matlab; SQL (SQLite); bash; PERL

**Technologies:** Git/Github; Tensorflow/Keras; Docker; Scikit-learn; numpy; scipy; pandas; PyTorch (basic)

**Methods:** Machine Learning; Data Analysis; Stats; Complex Systems (Agent-based simulations, Network Modeling)

**Languages:** Italian (native); English (excellent); Danish (excellent); French (intermediate)

## Teaching experience

Learning Session *Language as Data*, Centre for Humanitarian Data, United Nations (June 2021)  
 Guest Lecturer, *Advanced Cognitive Neuroscience*, MSc Cognitive Science (Fall 2019)  
 Full-course teaching assistant, *Introduction to Cognitive Neuroscience*, BSc Cognitive Science (Spring 2018)  
 Guest lecturer, *Social and Cultural Dynamics*, BSc Cognitive Science (Spring 2018)  
 Full-course lecturer and instructor, *Cognitive Science*, MA Cognitive Semiotics, (Fall 2017)  
 Instructor, *Workshop on Statistical Analyses in Language Sciences Using R*, staff course at Aarhus University (Fall 2017)

---

## Invited talks (selected)

Introducing Neuroscout, a platform for flexible (re)analysis of naturalistic datasets  
*September 2021*  
 Embodied Computation Group  
 Aarhus University

Complex systems modeling for humanitarian action: exploring opportunities and impact  
*July 2021*  
 Data Fellows Programme Final Showcase  
 Centre for Humanitarian Data, United Nations

Opportunities for integrating complex systems modeling in humanitarian planning and response  
*June 2021*  
 Webinar: Complex Systems Modelling: An Opportunity to Better Understand and Anticipate Humanitarian Needs?  
 Centre for Humanitarian Data, United Nations

Language models as person(al)ity encoders: a text-driven approach to modeling individual differences  
*May 2021*  
 Department of Linguistics, Cognitive Science and Semiotics  
 Aarhus University

Introducing benchmarks in psychology: rethinking model evaluation with an eye on prediction and the real world  
*February 2021*  
 Joint online seminar series  
 McGill University / University of Toronto / Karolinska Institute

---

## Organization of scientific events

Organizer of Natural Language Processing Workshop @ Interacting Minds Centre, Aarhus University (Fall 2019)  
 Co-organizer of SALC7, Conference of the Scandinavian Association for Language and Cognition (Spring 2019)  
 Co-organizer of Symposium “From fieldwork to modelling: Explaining the variability of linguistic spatial referencing systems”, ICSC 2018 (Fall 2018)

---

## Reviewing services

Journal of Experimental Psychology  
 Journal of Memory and Language  
 Psychology of Language and Communication  
 F1000Research

---

## Publications

de la Vega, A., **Rocca, R.**, Blair, R., Mentsch, J., Markiewicz, C., Ghosh, S., Poldrack, R., Yarkoni, T., Neuroscout: a platform for large-scale reproducible fMRI research (in preparation)

Tylén, K., **Rocca R.**, Cognitive diversity promotes collective creativity: an agent-based simulation (in preparation)

**Rocca, R.**, de la Vega, A. Evaluating the role of non-lexical markers in GPT-2 language modeling behavior (submitted)

**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*

Todisco, E., **Rocca, R.**, Wallentin, M. (2021). The semantics of spatial demonstratives in Spanish: a demonstrative choice task study. *Language and Cognition*

- Rocca, R.**, Wallentin, M. (2020). Demonstrative reference and semantic space: a large-scale demonstrative choice task (DCT) study. *Frontiers in Psychology*
- Wallentin, M., **Rocca, R.** (2020). The semantics of spatial demonstratives. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*
- 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., Vesper, C., & Tylén, K. (2019). This is for you: social modulations of proximal vs. distal space in collaborative interaction, *Scientific Reports*
- Rocca, R.**, Tylén, K., Wallentin, M. (2019), *This shoe, that tiger*: Semantic properties reflecting manual affordances of the referent modulate demonstratives use, *PlosOne*.
- Wallentin, M., **Rocca, R.**, Stoustrup S. (2019), Grammar, gender and demonstratives in lateralized imagery for sentences, *Journal of Psycholinguistic Research*
- Rocca, R.**, Wallentin, M., Vesper, C. & Tylén, K. (2018). *This and that* back in context: grounding demonstrative reference in manual and social affordances, *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*
- Wallentin, M., **Rocca, R.**, Stoustrup, S. (2018), Lateralized imagery for sentence content: Testing grammar, gender and demonstratives, *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*
- Carapezza, M., & **Rocca, R.** (2017). In-seguire la Regola: Giochi Linguistici e Arti Performative. *Rivista Italiana di Filosofia del Linguaggio*, 11(2). (in Italian)
- Rocca, R.** & Augustin, M. (2016), Gesturing in L2: Evidence for cross-linguistic transfer in the visual modality, in *TwistX Proceedings of the 10th Linguistics Student Conference, Leiden 22<sup>nd</sup> April 2016*
- Carapezza, M. & **Rocca, R.** (2014), Ceci n'est pas une ontologie: a contribution to a quasi-resolute reading of the *Tractatus Logico-Philosophicus*, in Rinhofner-Kreidl S. & Wilsche H. (eds), *Analytical and Continental Philosophy: Methods and Perspectives. Paper of the 37th Wittgenstein Symposium*, Austrian Ludwig Wittgenstein Society, Kirchberg am Wechsel

## Talks at conferences/workshops

- Rocca, R.**, Milano, L., (October 2021). Exploring dynamic approaches to needs modeling in humanitarian action. Conference on Complex Systems #CCS2021, Lyon – Satellite Meeting “Complex Systems for the Most Vulnerable”
- Rocca, R.**, Wallentin, M., (November 2019). Representing word meaning using cognitive features. NLP workshop at Interacting Minds Centre, Aarhus University
- Rocca, R.**, Staib, M., Tylén, K., Coventry, K.R., Wallentin, M. (July 2019). Here and there, what and where. Bilateral network of dorsal regions supports comprehension of spatial demonstratives in naturalistic language processing. *Deictic Communication: Theory and Application*, University of East Anglia
- Rocca, R.**, Staib, M., Tylén, K., Coventry, K.R., Wallentin, M. (May 2019). The neural correlates of spatial deictics: A fast fMRI study using naturalistic auditory stimuli. SALC7 – Conference of the Scandinavian Association for Language and Cognition, Aarhus University
- Rocca, R.**, Tylén, K., Wallentin, M. (May 2019). Crocodiles, peace and harmonicas: (how) does word semantics modulate demonstrative use? SALC7 – Conference of the Scandinavian Association for Language and Cognition, Aarhus University
- Rocca, R.**, Rabe, F., Madsen, KH, Siebner H., Lehmann, S., Konvalinka, I. (November 2018), Interpersonal coordination mechanisms in a real-world social network, *Social Cognition Workshop*, The Interacting Minds Centre, Aarhus University
- Rocca, R.**, Wallentin, M., Vesper, C. & Tylén, K. (July 2018). This and that back in context: grounding demonstrative reference in manual and social affordances; Oral presentation at *40th<sup>th</sup> Annual Meeting of the Cognitive Science Society*, Madison, Wisconsin, USA.
- Augustin, M., & **Rocca, R.** (April 2016), Gesturing in L2: Evidence for cross-linguistic transfer in the visual modality, Oral presentation: *TWISTx Conference*, Leiden.
- Rocca, R.** & Augustin, M. (March 2016), A cross-cultural study on gestural transfer in Italian, Dutch and English, oral presentation: Annual Juniorendag of the Dutch Society for Applied Linguistics (Anéla), Amsterdam
- Rocca, R.** & Carapezza, M. (August 2014), *Ceci n'est pas une ontologie*: a contribution to a quasi-resolute reading of the *Tractatus Logico-Philosophicus*, oral presentation: 37th ALWS Conference, Kirchberg am Wechsel

## Poster presentations

- Rocca, R.**, Staib, M., Tylén, K., Coventry, K., Lund, T.E., Wallentin, M. (August 2019), Here and there, what and where: investigating the role of the dorsal visuospatial stream in linguistic encoding, Annual Conference of the Society for the Neurobiology of Language, Helsinki, Finland
- Rocca, R.**, Staib, M., Tylén, K., Coventry, K., Lund, T.E., Wallentin, M. (June 2019), Investigating the neural correlates of spatial demonstratives using naturalistic fast fMRI, Organization for Human Brain Mapping Annual Meeting, Rome
- Rocca, R.**, Lund, T.E., Staib, M., Tylén, K., Coventry, K.R., Wallentin, M. (August 2018), Spatial attention and perceptual-motor representations of demonstratives: a fast fMRI study using naturalistic auditory stimuli, *SNL 2018 – Conference of the Society for the Neurobiology of Language*, Québec City, Canada.
- Wallentin, M., Lund, T. E., Andersen, C. M., & **Rocca, R.** (August 2018). Fast phonotopic mapping with oscillation-based fMRI– Proof of concept, *SNL 2018 – Conference of the Society for the Neurobiology of Language*, Québec City, Canada.
- Wallentin, M., **Rocca, R.**, Stoustrup, S. (July 2018), Lateralized imagery for sentence content: Testing grammar, gender and demonstratives, Poster presentation at *40th<sup>th</sup> Annual Meeting of the Cognitive Science Society*, Madison, Wisconsin, USA.
- Rocca, R.**, Staib, M., Tylén, K., Coventry, K., Lund, T.E., Wallentin, M., *Here and There* around the brain: The role of extra-linguistic spatial representations in the comprehension of spatial deictics
- Brain Prize Meeting 2017, Helsingørsk Castle – Middelfart (Denmark), October/November 2017;
  - 7th IMPRS/ICN *NeuroCom* Summer School, University College London, 12-14 July 2017;
  - *Language as a Form of Action*, National Research Council (CNR), Rome (Italy), 21-23 June 2017.

---

## Non-academic publications (selected)

- Algoritmi di classe, *Doppiozero*, 22/01/2021
- Chi ha paura dei data scientists? Numeri e pandemia, *Doppiozero*, 05/11/2020
- Dati, miti, stati, *L'identità di Clio*, 07/07/2020
- Montanelli, le statue e il colore della storia, *Lavoro culturale*, 06/07/2020
- Ricerca, democrazia e Unione Europea, *Doppiozero*, 03/06/2019
- Chiamatemi Israele: storia di un monte, un tempio e il suo eremita, *Indiscreto*, 16/03/2018
- Recreating space with “Oui” @ Aarhus Kunsthall, *Jutland Station*, 06/06/2017
- Art exhibition: Exploring the power of representation in contemporary media, *Jutland Station*, 28/03/2017
- Art exhibition: Signe Klejs creates Hesitation Through Light, *Jutland Station*, 25/01/2017
- Capitali coraggiose, *Indiscreto*, 18/01/2017
- Music review: Who Killed Bambi & Ida Wenoe @ Aarhus Filmbyen, *Jutland Station*, 14/01/2017
- Art exhibition: “No man is an island – The Satanic verses” @ Aros, *Jutland Station*, 03/12/2016
- Art exhibition: PROCES by Holm/Ibjerg @ PLADS Artspace, *Jutland Station*, 16/11/2016
- Guttuso, la forza delle cose, *Indiscreto*, 21/09/2016
- In risposta all'Assoluto: Alberto Giacometti e Yves Klein, *Indiscreto*, 29/08/2016
- Brulicamenti (ovvero: cartoline dalla nuova Tate), *Indiscreto*, 04/07/2016
- Bosch a Den Bosch, *Lavoro culturale*, 08/04/2016
- 

## Other interests

Culture journalism; Creative writing; Music; Art-house cinema