

Mariya Toneva

+49 (151) 21159581
mtoneva@mpi-sws.org
<http://mtoneva.com>

Positions Held

Max Planck Institute for Software Systems

Tenure-track Faculty (W2)
Visiting Researcher

started Sept 2022
2021–2022

Princeton University

Postdoctoral Researcher
Advisors: Ken Norman, Uri Hasson

2021–2022

Education

Carnegie Mellon University

Ph.D. in Machine Learning and Neural Computation
Thesis title: Bridging Language in Machines with Language in the Brain
Advisors: Tom Mitchell, Leila Wehbe

2014–2021

Carnegie Mellon University

Masters of Science in Machine Learning

2018

Yale University

Bachelor of Science in Computer Science, Cognitive Science

2014

Publications in Journals and Conference Proceedings

1. Training language models for deeper understanding improves brain alignment 2023
K.L. Aw and **M. Toneva**
(ICLR 2023) *International Conference on Learning Representations* [pdf]
[top 25% notable paper (Spotlight)]
2. A Roadmap to Reverse Engineering Real-world Generalization by Combining Naturalistic Paradigms, Deep Sampling, and Predictive Computational Models 2023
P. Herholz, E. Fortier, **M. Toneva**, N. Farrugia, L. Wehbe, V. Borghesani
Neurons, Behavior, Data Science, and Theory [pdf]
3. Combining Computational Controls with Natural Text Reveals New Aspects of Meaning Composition 2022
M. Toneva, T. Mitchell, and L. Wehbe
Nature Computational Science [pdf]
4. Same Cause; Different Effects in the Brain 2022
M. Toneva*, and J. Williams*, A. Bollu, C. Dann, and L. Wehbe
(CLEaR 2022) *Causal Learning and Reasoning* [pdf]

5. Single-trial MEG Data Can Be Denoised Through Cross-Subject Predictive Modeling 2021
S. Ravishankar, **M. Toneva**, and L. Wehbe
Frontiers in Computational Neuroscience 2021 [\[pdf\]](#)
6. Modeling Task Effects on Meaning Representation in the Brain 2020
via Zero-Shot MEG Prediction
M. Toneva*, O. Stretcu*, B. Poczós, L. Wehbe, and T. Mitchell
(NeurIPS 2020) *Neural Information Processing Systems* [\[pdf\]](#)
7. Interpreting and Improving Natural-Language Processing (in Machines) with 2019
Natural Language-Processing (in the Brain)
M. Toneva and L. Wehbe
(NeurIPS 2019) *Neural Information Processing Systems* [\[pdf\]](#)
8. Inducing Brain-relevant Bias in Natural Language Processing Models 2019
D. Schwartz, **M. Toneva**, and L. Wehbe
(NeurIPS 2019) *Neural Information Processing Systems* [\[pdf\]](#)
9. An Empirical Study of Example Forgetting during Deep Neural Network Learning 2019
M. Toneva*, A. Sordoni*, R. Tachet des Combes*, A. Trischler, Y. Bengio, and G. Gordon
(ICLR 2019) *International Conference on Learning Representations* [\[pdf\]](#)
10. Applying Artificial Vision Models to Human Scene Understanding 2015
E. M. Aminoff, **M. Toneva**, A. Shrivastava, X. Chen, I. Misra, A. Gupta, and M. J. Tarr
Frontiers in Computational Neuroscience 2015 [\[pdf\]](#)
11. Exploration of Social Grouping: Effects of Behavioral Mimicry, Appearance, and Eye Gaze 2014
A. Nawroj, **M. Toneva**, H. Admoni, B. Scassellati
(CogSci 2014) *Conference of the Cognitive Science Society* [\[with Oral presentation\]](#) [\[pdf\]](#)
12. The Physical Presence of a Robot Tutor Increases Cognitive Learning Gains 2012
D. Leyzberg, S. Spaulding, **M. Toneva**, and B. Scassellati
(CogSci 2012) *Conference of the Cognitive Science Society* [\[pdf\]](#)
13. Robot Gaze Does Not Reflexively Cue Human Attention 2011
H. Admoni, C. Bank, J. Tan, **M. Toneva**, and B. Scassellati
(CogSci 2011) *Conference of the Cognitive Science Society* [\[pdf\]](#)

Preprints and Non-Proceeding Publications

- Large language models can segment narrative events similarly to humans 2023
S. Michelmann, M. Kumar, K.A. Norman, and **M. Toneva**
(arXiv 2023) [\[pdf\]](#)
- Joint processing of linguistic properties in brains and language models 2022
S.R. Oota, M. Gupta, and **M. Toneva**
(arXiv 2022) [\[pdf\]](#)

- Language models and brain alignment: beyond word-level semantics and prediction 2022
 G. Merlin and **M. Toneva**
 (arXiv 2022) [\[pdf\]](#)
- Memory for long narratives 2022
M. Toneva, V. Vo, J. Turek, S. Jain, S. Michelmann, M. Capotă, A. Huth, U. Hasson,
 and K. Norman
 (CEMS 2022) *Context and Episodic Memory Symposium*
- The Courtois Neuromod project: a deep, multi-domain fMRI dataset to build individual 2022
 brain models
 J. Boyle*, B. Pinsard*, V. Borghesani, M. Saint-Laurent, F. Lespinasse, F. Paugam,
 P. Sainath, S. Rastegarnia, A. Boré, J. Chen, A. Cyr, E. Dessureault, E. DuPre, Y. Harel,
M. Toneva, S. Belleville, S. Brambati, J. Cohen-Adad, A. Fuente, M. Hebart, K. Jerbi,
 P. Rainville, L. Wehbe, and P. Bellec
 (HBM 2022) *Human Brain Mapping* [\[with Oral presentation\]](#)
- Does Injecting Linguistic Structure into Language Models Lead to Better Alignment 2021
 with Brain Recordings?
 M. Abdou, A.V. González, **M. Toneva**, D. Hershovich, and A. Søgaard
 (arXiv 2021) [\[pdf\]](#)
- Investigating Different Alignment Methods Between Natural and Artificial Neural Networks 2020
 for Language Processing
 A. Bollu, **M. Toneva**, and L. Wehbe
 (SNL 2020) *Society for the Neurobiology of Language*
- Investigating Task Effects on Brain Activity During Stimulus Presentation in MEG 2019
M. Toneva*, O. Stretcu*, B. Poczos, and T. Mitchell
 (HBM 2019) *Human Brain Mapping*
- Word Length Processing in Left Lateraloccipital through Region-to-Region Connectivity: 2018
 an MEG Study
M. Toneva, and T. Mitchell
 (HBM 2018) *Human Brain Mapping*
- MEG Representational Similarity Analysis Implicates Hierarchical Integration 2018
 in Sentence Processing
 N. Rafidi*, D. Schwartz*, **M. Toneva***, S. Jat, and T. Mitchell
 (HBM 2018) *Human Brain Mapping*
- Scene-Space Encoding within the Functional Scene-Selective Network 2015
 E. M. Aminoff, **M. Toneva**, A. Gupta, and M. J. Tarr
 (VSS 2015) *Vision Sciences Society*
- Towards a Model for Mid-level Feature Representation of Scenes 2014
M. Toneva, E. M. Aminoff, A. Gupta, and M. Tarr
 (VSS 2014) *Vision Sciences Society*

Grants and Fellowships

DFG Research Unit 5368 KI-FOR Abstract Representations in Neural Architectures Bridging Levels of Abstraction in Brains and Natural Language Processing Machines	2023-2027
National Institutes of Health T32 Training Grant Funded one year of postdoctoral research in Quantitative and Computational Neuroscience	2021-2022
C.V. Starr Fellowship Funded one year of postdoctoral research in computational neuroscience at Princeton University	2021-2022
National Science Foundation Graduate Research Fellowship Funded three years of interdisciplinary graduate research in machine learning and neuroscience	2016-2019
Grace Hopper Celebration Scholarship Funded attendance at the 2014 Grace Hopper Celebration of Women in Computing	2014
Mellon Forum Undergraduate Research Grant Funded submission and attendance at the 2014 Vision Sciences Society conference	2014
Robin Berlin Fellowship Funded neural modeling research at Laboratory of Computational Neuroscience, EPFL	2013

Awards

Ph.D. Dissertation Award, Honorable Mention Society for the Neurobiology of Language	2021
Machine Learning Student Leadership Award Awarded for exemplary efforts and their significant impact on life in the Machine Learning Department	2020
Top Reviewer NeurIPS 2022, ICLR 2022, ICML 2022, NeurIPS 2018	2018-2022
Citadel Datathon Runner-up Analyzed a genomics dataset to predict age-related differences in disease-related gene expression	2017
Machine Learning Teaching Assistant Award Awarded for outstanding performance as a TA in 10-725 Convex Optimization	2017
BrainHub Neurohackathon Winner Reduced need for human supervision by classifying diffusion MRI tracks into anatomical bundles	2016

Conference and Invited Talks

▪ Why Do Large Language Models Align with Human Brains? CIMeC, University of Trento; Host: Raffaella Bernardi	2022
▪ Why Do Large Language Models Align with Human Brains? TALEP group, Aix-Marseille University; Hosts: Abdellah Fourtassi, Carlos Ramisch	2022
▪ Bridging Language in Machines with Language in the Brain CMMRS Summer School, MPI for Software Systems	2022

- **NLP Systems as Model Organisms for Human Language Comprehension** 2022
IMPRS NeuroCom Summer School at the MPI for Human Cognitive and Brain Sciences
- **Deep Neural Networks as Model Organisms for Human Language Comprehension** 2022
ML, Abstract Thought, and the Expanding Reach of AI: Ethical and Conceptual Frontiers
- **Same Cause; Different Effects in the Brain** 2022
MIT, Host: Evelina Fedorenko
- **Bridging Language in Machines with Language in the Brain** 2022
Distinguished Speakers in Language Science Colloquium, Saarland University
- **Data-Driven Transfer of Insight between Brains and AI Systems** 2021
MIT, Host: Roger Levy
- **NLP Systems as Model Organisms for Human Language Comprehension** 2021
Computational Neuroscience Symposium
- **Data-Driven Transfer of Insight between Brains and AI Systems** 2021
MIT
Yale University
Boston University
University of Southern California
University of Utah
Max Planck Institute for Software Systems
Aarhus University
Toyota Technological Institute at Chicago
IST Austria
University of Chicago
University of Notre Dame
University of North Florida
University of Liverpool
- **NLP Systems as Model Organisms for Human Language Comprehension** 2021
Courtois NeuroMod Group, Host: Pierre Bellec
- **Data-Driven Direct Transfer of Insight between Brains and AI Systems** 2020
SFB-TRR 161 Lecture Series (University of Stuttgart, University of Konstanz, Ulm University, and the LMU Munich), Host: Lewis Chuang
- **Nonlinear Models for Scientific Discovery about Language in the Brain** 2020
Invited speaker and panelist, *Cognitive Computational Neuroscience (CCN) workshop*
Is it that simple? The use of linear models in cognitive neuroscience
- **Modeling Task Effects on Meaning Representation in the Brain** 2020
Carnegie Mellon University, brAln seminar
- **Composition of Context- and Task-dependent Meaning** 2020
UT Austin, Host: Alexander Huth
- **Towards a Model for Mid-level Feature Representation of Scenes** 2014
Oral presentation, *Women in Machine Learning (WiML) workshop at NeurIPS*
- **Exploration of Social Grouping: Effects of Behavioral Mimicry, Appearance, and Eye Gaze** 2014
Oral presentation, *Conference of the Cognitive Science Society (CogSci)*

Industry Internships

Microsoft Research, Montreal

Research Intern 2018

Investigated the learning dynamics of neural networks as they train on single classification tasks, finding that certain examples are forgotten with high frequency, and some not at all, and that, based on these forgetting dynamics, a significant fraction of examples can be omitted from the training data set while still maintaining state-of-the-art generalization performance

Cognitive Computing Center, Thomson Reuters

Research Intern 2017

Investigated the use of a recurrent neural network encoder for unsupervised word-order sensitive hashing as a step towards improving ranking results

Research Visits

Carnegie Mellon University

Research Assistant; Advisor: Michael Tarr 2013–2014

Investigated mid-level scene representation in humans using computer vision techniques

École Polytechnique Fédérale de Lausanne (EPFL)

Summer Intern; Advisor: Wulfram Gerstner 2013

Worked towards improving the state-of-the-art calcium-based model of spike-timing dependent plasticity

Massachusetts Institute of Technology

Technical Trainee; Advisor: John Gabrieli 2012

Examined links between working memory capacity and various brain metrics through the analysis of resting state functional connectivity fMRI data

Mentorship and Supervision

Gabriele Merlin

PhD at CS@MaxPlanck Graduate Program 2021-

Tianai (Dota) Dong

Masters in Language Science and Technology at Saarland University 2021-2022

Subba Reddy Oota

Research Intern 2022

Khai Loong Aw

Research Intern 2022

Ruchit Rawal

Research Intern 2022

Anand Bollu

Masters at Department of Computer Science, CMU 2019-2021

Sydney Zheng

Undergraduate at Department of Computer Science, CMU 2019

Aditri Bhagirath

Undergraduate at Department of Computer Science, CMU 2019

Tara Pirnia

MD/PhD candidate, CMU and University of Pittsburgh 2015

Teaching

3370 Mathematical Neuroscience, University of Pittsburgh Teaching Assistant	2018
10-725 Convex Optimization, CMU Teaching Assistant Awarded Machine Learning TA award	2016
Machine Learning for Neuroscience, Multimodal Neuroimaging Training Program Instructor Created curriculum and instructed 4-week course; video recordings can be found on personal webpage	2016

Service

Organizer

Memory in Artificial and Real Intelligence Workshop, NeurIPS	2022
Deep Learning for Brain Encoding and Decoding Tutorial, Cognitive Sciences Society	2022
What can NLP systems teach us about language in the brain? Symposium, Society for the Neurobiology of Language	2021
How can findings about the brain improve AI systems? Workshop, ICLR	2021

Reviewer

ML: *NeurIPS* 2016, 2018(Top 30% Reviewer)-2022; *ICML* 2019,2021,2022(Top 10% Reviewer);
AAAI 2020-2021, *CoLLa* 2022, *ICLR* 2022-2023 (Highlighted Reviewer)
NLP: *ACL* 2019-2021; *NAACL* 2019-2021; *EMNLP* 2020-2021; *CoNLL* 2020-2021; *AAACL-IJCNLP* 2020;
EACL 2021
Neuro: *Nature Communications*; *Frontiers in Computational Neuroscience*;
Society for the Neurobiology of Language 2022; *Organization for Human Brain Mapping* 2018
Other venues: *CogSci* 2021

ML@CMU Blog

Chief Editor and Co-founder 2018–2020
Oversaw more than 30 research posts featuring recent ML research across 6 departments in the School of Computer Science as well as other CMU schools and departments, and more than 10 educational posts

University Leadership Student Advisory Council

Member 2015–2017
Advising senior leadership at Carnegie Mellon University on the strategic priorities of the university

Graduate Student Assembly

Representative for the Program of Neural Computation 2015–2018
Advocating for the needs of graduate students

Yale Review of Undergraduate Research in Psychology

Chief Editor 2013–2014
Reviewed 50 submissions from 31 universities, and edited 9 submissions for publication

Personal

Languages Bulgarian (Native), English (Fluent), German (Intermediate)
Citizenship United States, Bulgaria
Github profile <http://github.com/mtoneva>
Google Scholar profile <https://scholar.google.com/citations?user=a61sk-4AAAAJ>