

William Merrill

<https://lambdaviking.com/>

willm[æt]nyu.edu

Last updated June 10, 2022

RESEARCH INTERESTS

Broad

NLP, deep learning, formal languages and automata, formal semantics, computational complexity, historical linguistics

Specific

- Neural nets for NLP: their implicit computational model and inductive biases
- Processing language with bounded resources: counter automata, circuits, streaming algorithms, etc.
- Theory of learning linguistic meaning from unlabeled data
- Robustness and interpretability of modern NLP systems

PROFESSIONAL EXPERIENCE

Google	2022	Student Researcher , Speech & Language Algos
AI2	2019–2021	Predoc. Young Investigator , AllenNLP
Google	2018	Software Engineering Intern <i>“Exceeds expectations” rating; return offer</i>
Boston College	2017	Research Intern , Language Learning Lab
New York University	2013–2015	Research Intern , Morphology Lab

EDUCATION

New York University	2021–	Ph.D. at Center for Data Science
Yale University	2015–2019	B.Sc. with distinction in Computer Science B.A. with distinction in Linguistics Thesis: Sequential neural networks as automata <i>Cum laude; note of excellence on thesis</i>

MENTORS AND GROUP AFFILIATIONS

<i>Tal Linzen</i>	ML ² , NYU	2021–
<i>Noah A. Smith, Yoav Goldberg, Roy Schwartz</i>	AllenNLP, (AI) ²	2019–2021
<i>Robert Frank, Dana Angluin</i>	CLAY, Yale	2016–2019
<i>Joshua Hartshorne, Sven Dietz</i>	L ³ , Boston College	2017
<i>Alec Marantz, Phoebe Gaston</i>	MorphLab, NYU	2013–2015

PUBLICATIONS

More recent publications appear first

- [1] William Merrill and Nikolaos Tsilivis. Extracting finite automata from RNNs using state merging, 2022.
- [2] William Merrill, Ashish Sabharwal, and Noah A. Smith. Saturated transformers are constant-depth threshold circuits, 2021.
- [3] Matt Gardner, William Merrill, Jesse Dodge, Matthew Peters, Alexis Ross, Sameer Singh, and Noah A. Smith. Competency problems: On finding and removing artifacts in language data. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 1801–1813, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [4] William Merrill, Yoav Goldberg, Roy Schwartz, and Noah A. Smith. Provable Limitations of Acquiring Meaning from Ungrounded Form: What Will Future Language Models Understand? *Transactions of the Association for Computational Linguistics*, 9:1047–1060, 09 2021.
- [5] William Merrill, Vivek Ramanujan, Yoav Goldberg, Roy Schwartz, and Noah A. Smith. Effects of parameter norm growth during transformer training: Inductive bias from gradient descent. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 1766–1781, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [6] William Merrill, Gail Weiss, Yoav Goldberg, Roy Schwartz, Noah A. Smith, and Eran Yahav. A formal hierarchy of RNN architectures. In *Proceedings of*

the 58th Annual Meeting of the Association for Computational Linguistics, pages 443–459, Online, July 2020. Association for Computational Linguistics.

- [7] Lucy Lu Wang, Kyle Lo, Yoganand Chandrasekhar, Russell Reas, Jiangjiang Yang, Doug Burdick, Darrin Eide, Kathryn Funk, Yannis Katsis, Rodney Michael Kinney, Yunyao Li, Ziyang Liu, William Merrill, Paul Mooney, Dewey A. Murdick, Devvret Rishi, Jerry Sheehan, Zhihong Shen, Brandon Stilson, Alex D. Wade, Kuansan Wang, Nancy Xin Ru Wang, Christopher Wilhelm, Boya Xie, Douglas M. Raymond, Daniel S. Weld, Oren Etzioni, and Sebastian Kohlmeier. CORD-19: The COVID-19 open research dataset. In *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*, Online, July 2020. Association for Computational Linguistics.
- [8] William Merrill. On the linguistic capacity of real-time counter automata, 2020.
- [9] William Merrill. Sequential neural networks as automata. In *Proceedings of the Workshop on Deep Learning and Formal Languages: Building Bridges*, pages 1–13, Florence, August 2019. Association for Computational Linguistics.
- [10] William Merrill, Lenny Khazan, Noah Amsel, Yiding Hao, Simon Mendelsohn, and Robert Frank. Finding hierarchical structure in neural stacks using unsupervised parsing. In *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 224–232, Florence, Italy, August 2019. Association for Computational Linguistics.
- [11] William Merrill, Gigi Stark, and Robert Frank. Detecting syntactic change using a neural part-of-speech tagger. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, pages 167–174, Florence, Italy, August 2019. Association for Computational Linguistics.
- [12] Yiding Hao, William Merrill, Dana Angluin, Robert Frank, Noah Amsel, Andrew Benz, and Simon Mendelsohn. Context-free transductions with neural stacks. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 306–315, Brussels, Belgium, November 2018. Association for Computational Linguistics.

- [13] Jungo Kasai, Robert Frank, Pauli Xu, William Merrill, and Owen Rambow. End-to-end graph-based TAG parsing with neural networks. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 1 (Long Papers)*, pages 1181–1194, 2018.
- [14] William Merrill. A semantics of subordinate clauses using delayed evaluation. *Toronto Undergraduate Linguistics Conference*, 2018.

INVITED TALKS

- **Umeå University**, Foundations of Language Processing, 2022
Saturated Transformers are Constant-Depth Threshold Circuits
- **MILA**, ML for Code Seminar, 2022
Saturated Transformers are Constant-Depth Threshold Circuits
- **MIT**, CompLang Seminar, 2022
Language Models Have Implicit Entailment Semantics
- **NYU**, Semantics Seminar, 2022
Distributional Learnability of Entailment
- **Google**, Speech and Language Algorithms, 2022
Neural Networks as Automata
- **ArthurAI**, Journal Club, 2021
Competency Problems: On Finding and Removing Artifacts in Language Data
- **EMNLP**, ML Track, 2021
Competency Problems: On Finding and Removing Artifacts in Language Data
- **EMNLP**, ML Track, 2021
Parameter Norm Growth During Transformer Training: Inductive Bias From Gradient Descent

- **AI2**, All Hands, 2021
Provable Limitations of Acquiring Meaning from Ungrounded Form: What Will Future Language Models Understand?
- **UW**, Noah's ARK, 2020
Provable Limitations of Acquiring Meaning from Ungrounded Form: What Will Future Language Models Understand?
- **EMNLP**, Blackbox NLP, 2018
Context-Free Transductions with Neural Stacks
- **Packer Collegiate Institute**, Science Research Symposium, 2018
Neural networks, L2 Acquisition, and the Voynich
- **CodeHaven**, 2018
Programming, Language, and the Book of Thoth
- **UToronto**, TULCon, 2018
A Semantics of Subordinate Clauses Using Delayed Evaluation

TEACHING EXPERIENCE

- Lead TA for **Natural Language Processing**, Tal Linzen (NYU, Fall 2022)
- TA for **Artificial Intelligence**, Dragomir Radev (Yale, Spring 2019)
- TA for **Natural Language Processing**, Dragomir Radev (Yale, Fall 2018)
- TA for **Artificial Intelligence**, Dragomir Radev (Yale, Spring 2017)

SERVICE

Reviewing

ARR	April 2022	1 review
ARR	Jan 2022	2 review
ARR	Dec 2021	3 reviews
ARR	Nov 2021	1 review
CL	2021	1 review
ACL	2021	6 reviews
EACL	2021	4 reviews
EMNLP	2020	2 reviews
Neural Networks	2020	1 review

Outreach

NYC AI School	2022	Volunteer instructor
AllenNLP Hackathon	2021	Technical support
Yale Tangut Language Workshop	2018	Workshop facilitator
Yale NACLO	2017	Student volunteer
Yale Kitan Language Workshop	2016	Workshop facilitator
CodeHaven	2016–2018	Student volunteer
Splash at Yale	2016–2017	Student instructor

SELECTED PUBLIC SOFTWARE

- **AllenNLP**: Open-source NLP framework (contributor)
- **The Book of Thoth**: Puzzle game with compositional spell casting in Middle Egyptian hieroglyphs
- **DraftNet**: Dota 2 drafting using neural networks
- **Voynich2Vec**: Word embedding analysis of the Voynich manuscript
- **StackNN**: Differentiable stacks, queues, and dequeues in PyTorch

BLOG POSTS

Research Content

- *A Formal Hierarchy of RNN Architectures* (2020)
- *Theory of Saturated Neural Networks* (2019)
- *The State of Interpretability in NLP* (2019, outdated!)
- *Word2vec Analysis of the Voynich Manuscript* (2018)
- *Review: Learning to Transduce with Unbounded Memory* (2018)
- *Capsule Networks for NLP* (2018)

Translations

- *The Wanderer* (Old English → English)
- *After Ragnarok* (Old Norse → English)
- *The Saga of Mary* (Old Norse → English)

AWARDS AND GRANTS

- **NSF Graduate Student Research Fellowship** (2022)
- **Student Travel Grant** to attend DELFOL workshop at ACL, presented by Naver Labs (2019)
- **Mellon Grant** for senior thesis work, presented by Benjamin Franklin College at Yale University (2019)
- **Grace Hopper Prize** for computer science finalist (2017)
- Yale College **freshman rap battle champion** (2016)
- **Rising Scientist Award** presented by the Child Mind Institute (2015)
- **National Merit Scholarship** letter of commendation (2013)

- **Study of American History Award** presented by the Society of Mayflower Descendants (2013)
- National Latin Exam *cum honore maximo egregio* (2010)

SELECTED COURSEWORK

Theoretical Computer Science and Formal Linguistics

- Formal Foundations of Linguistic Theory (Yale, 2015)
- Introduction to Computer Science (Yale, 2015)
- Computing Meanings (Yale, 2016)
- Design and Analysis of Algorithms (Yale, 2017)
- Computability and Logic (Yale, 2017)
- Computational Complexity Theory (Yale, 2018)
- Foundations of Machine Learning (NYU, 2022)

Deep Learning and Natural Language Processing

- Natural Language Processing (Yale, 2017)
- Deep Learning Theory and Applications (Yale, 2018)
- Neural Networks and Language (Yale, 2018)
- Computational Vision and Biological Perception (Yale, 2018)
- Seminar: Advanced Natural Language Processing (Yale, 2018)
- Seminar: Selected Topics in Neural Networks (Yale, 2019)
- Ph.D. Introduction to Data Science (NYU, 2021)
- Seminar: Scaling Laws, the Bitter Lesson, and AI Research (NYU, 2021)

Other Linguistics

- Old English (Yale, 2015)
- Semantics I (Yale, 2016)
- Medieval Latin Paleography (Yale, 2016)
- Seminar: Beowulf and the Northern Heroic Tradition (Yale, 2017)
- Syntax I (Yale, 2017)
- Indo-European Linguistics (Yale, 2018)
- The Voynich Manuscript (Yale, 2018)
- Phonology I (Yale, 2018)
- Hybrid Grammars: Language Contact and Change (Yale, 2019)

Other Computer Science

- Data Structures and Programming Techniques (Yale, 2016)
- Systems Programming Techniques and Computer Organization (Yale, 2017)
- Big Data (NYU, 2022)

Continuous Math

- MATH 230: Vector Calculus and Linear Algebra I (Yale, 2015)
- MATH 231: Vector Calculus and Linear Algebra II (Yale, 2016)
- Introduction to Analysis (Yale, 2017)

Reading Groups

- Deep Learning Theory (AI2, 2020)
- Nonlinear Dynamical Systems (AI2, 2021)

LANGUAGES

- **Coding:** Python, Java, C, Haskell, PyTorch, AllenNLP, *inter alias*
- **Modern:** English (Native), Icelandic (Intermediate)
- **Ancient:** Latin, Old Norse, Old English