

R. Thomas McCoy

Department of Cognitive Science
141 Krieger Hall
Johns Hopkins University
3400 N. Charles Street
Baltimore, MD 21218-2685

Email: tom.mccoy@jhu.edu
Last updated: January 28, 2022

RESEARCH INTERESTS

I study computational linguistics using techniques from cognitive science, machine learning, and natural language processing. My research focuses on how to create models of language that generalize robustly to linguistic structures that they have not encountered before, as this remains one of the main areas where current AI systems fall short and one of the most impressive components of language processing in humans. In particular, I study which **linguistic inductive biases** and which **representations of linguistic structure** enable robust generalization, since these are two of the major components that determine how language learners generalize to novel types of input. Underlying much of this work is an additional theme of **analyzing and interpreting** neural network models: how can these models—which are based on vectors of continuous numbers—handle natural language—which has traditionally been viewed as a symbolic system?

EDUCATION

- | | |
|--------------|--|
| 2017–present | Johns Hopkins University: Ph.D. in Cognitive Science. GPA: 4.0.
<i>Advisors:</i> Tal Linzen, Paul Smolensky |
| 2013–2017 | Yale University: B.A. in Linguistics, <i>summa cum laude</i> , distinction in the major. GPA: 4.0.
<i>Advisor:</i> Robert Frank |
| Summer 2016 | Institute on Collaborative Language Research (CoLang), University of Alaska Fairbanks |
| Summer 2015 | Linguistic Summer Institute, University of Chicago |

EMPLOYMENT

- | | |
|-------------|--|
| Summer 2020 | Microsoft Research intern
Supervisor: Asli Celikyilmaz
<i>Evaluation methods for neural text generation systems</i> |
| Summer 2018 | Jelinek Summer Workshop on Speech and Language Technology (JSALT) sentence representations team
Team leaders: Sam Bowman, Ellie Pavlick
<i>Analysis techniques for learned sentence representations.</i> |

- Summer 2017 Carnegie Mellon University team for the DARPA Low Resource Languages for Emergent Incidents (LORELEI) project
Team leader: Patrick Littell
Finite-state morphological analyzer for Oromo.
- Summer 2017 Chirila project
Supervisor: Claire Bovern
Developed automatic semantic processing techniques for an online database of Australian languages.
- Summer 2016 Grammar Boot Camp
Supervisor: Claire Bovern
Wrote a sketch grammar of Kuwarra.
- Summer 2015 Yale Grammatical Diversity Project
Supervisors: Laurence Horn, Jim Wood, Raffaella Zanuttini, Jason Zentz
Edited web pages about regional grammatical phenomena.
- Summer 2014 Irish lip rounding research
Supervisor: Ryan Bennett
Collected lip rounding measurements from images of Irish speakers.
- Summer 2014 Linguistic Core Multi-University Research Initiative
Team leaders: Chris Dyer, Lori Levin
English-to-Malagasy tree-to-string transducer.
- Summer 2013 Linguistic Core Multi-University Research Initiative
Team leaders: Chris Dyer, Lori Levin
Finite state morphological analyzer for Kinyarwanda.

TEACHING

Across the five Johns Hopkins courses for which I have been a lab instructor or teaching assistant, 57 students have completed course evaluations rating my teaching on a 5-point scale. 79% scored my teaching as “Excellent” (the top score), 14% scored my teaching as “Good” (a score of 4/5), and the remaining 7% scored my teaching as “Satisfactory” (a score of 3/5). None have given scores of “Weak” or “Poor.”

- Spring 2020 Johns Hopkins University
Role: Teaching Assistant
Course: Foundations of Cognitive Science
Lecture Instructor: Paul Smolensky
Led one seminar discussion and graded assignments.
- Fall 2019 Johns Hopkins University
Role: Teaching Assistant, Lab Instructor
Course: Computational Psycholinguistics
Lecture Instructor: Tal Linzen
Led lab sessions and graded assignments.

- Spring 2019 Johns Hopkins University
 Role: Teaching Assistant
 Course: Syntax I
 Lecture Instructor: Géraldine Legendre
Led review sessions and graded assignments.
- Fall 2018 Johns Hopkins University
 Role: Teaching Assistant
 Course: Introduction to Computational Cognitive Science
 Lecture Instructor: Tal Linzen
Created educational simulations, tutorials, and homeworks in Javascript and Jupyter and taught lectures using these resources.
- Spring 2018 Johns Hopkins University
 Role: Fieldwork Instructor
 Course: World of Language
 Lecture Instructor: Géraldine Legendre
Led two sections of weekly fieldwork sessions complementing lectures.
- Summer 2015 Linguistic Society of America Summer Institute
 Role: Workshop Co-Instructor
 Course: Linguistic Enigmatography
 Co-Instructor: Lori Levin
Developed and co-taught a one-week workshop on creating linguistic puzzles.

AWARDS

1. Fellowships

- 2018–2021 NSF Graduate Research Fellowship
Project title: Assessing the capacity of computational models to make linguistic generalizations
- 2021 Sweitzer Fellow
Fellowship awarded by the Johns Hopkins Department of Cognitive Science to one graduate student.
- 2020 Finalist: Facebook Fellowship
One of four finalists in the Natural Language Processing category; two of the four finalists received fellowships.
- 2017–2020 Owen Scholars Fellowship
Fellowship for outstanding incoming Johns Hopkins PhD students in the natural sciences.
- 2017 Finalist: Rhodes Scholarship
- 2017 Finalist: Marshall Scholarship

2. Prizes

- 2017 Alpheus Henry Snow Prize
Award for the graduating Yale senior who is “adjudged by the faculty to have done the most for Yale by inspiring in his or her classmates an admiration and love for the best traditions of high scholarship.”
- 2016 Hart Lyman Prize
Award for the Yale junior who “has made through his/her own efforts the best record intellectually and socially.”
- 2016 Phi Beta Kappa
One of 13 Yale students admitted as juniors.
- 2013 World champion team at the International Linguistics Olympiad
Member of the four-person U.S. team selected by the North American Computational Linguistics Olympiad.
- 2013 United States Presidential Scholar
One of two for Pennsylvania.

3. Grants

- 2019 NeurIPS Travel Grant
Grant to fund travel to present work at the NeurIPS workshop on Context and Compositionality in Biological and Artificial Neural Systems.
- 2019 ICLR Travel Grant
Grant to fund travel to present two projects at the 2019 ICLR conference.
- 2018–2019 Johns Hopkins University Center for Educational Resources Technology Fellowship Grant
Co-Grantee: Tal Linzen
Grant to develop interactive visualizations of concepts in computational cognitive science.

PEER-REVIEWED PUBLICATIONS

- 2022 Paul Smolensky, **R. Thomas McCoy**, Roland Fernandez, Matthew Goldrick, and Jianfeng Gao. Neurocompositional computing: From the central paradox of cognition to a new generation of AI systems. Accepted to *AI Magazine*.
- 2021 **R. Thomas McCoy**, Jennifer Culbertson, Paul Smolensky, and Géraldine Legendre. [Infinite use of finite means? Evaluating the generalization of center embedding learned from an artificial grammar](#). In *Proceedings of the 43rd Annual Conference of the Cognitive Science Society*.

- 2021 Paul Soulos, Sudha Rao, Caitlin Smith, Eric Rosen, Asli Celikyilmaz, **R. Thomas McCoy**, Yichen Jiang, Coleman Haley, Roland Fernandez, Hamid Palangi, Jianfeng Gao and Paul Smolensky. [Structural Biases for Improving Transformers on Translation into Morphologically Rich Languages](#). In *Proceedings of the 4th Workshop on Technologies for Machine Translation of Low Resource Languages (LoResMT2021)*.
- 2020 **R. Thomas McCoy**, Erin Grant, Paul Smolensky, Thomas L. Griffiths, and Tal Linzen. [Universal linguistic inductive biases via meta-learning](#). In *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*.
- 2020 **R. Thomas McCoy**, Robert Frank, and Tal Linzen. [Does syntax need to grow on trees? Sources of hierarchical inductive bias in sequence-to-sequence networks](#). *Transactions of the Association for Computational Linguistics (TACL)*.
- 2020 Michael Lepori and **R. Thomas McCoy**. [Picking BERT’s brain: Analyzing contextualized embeddings using Representational Similarity Analysis](#). In *Proceedings of the 28th International Conference on Computational Linguistics (COLING)*.
- 2020 Paul Soulos, **R. Thomas McCoy**, Tal Linzen, and Paul Smolensky. [Uncovering the compositional structure of vector representations with Role Learning Networks](#). In *BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*.
- 2020 Michael Lepori, Tal Linzen, and **R. Thomas McCoy**. [Representations of Syntax \[MASK\] Useful: Effects of Constituency and Dependency Structure in Recursive LSTMs](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*.
- 2020 Junghyun Min, **R. Thomas McCoy**, Dipanjan Das, Emily Pitler, and Tal Linzen. [Syntactic data augmentation increases robustness to inference heuristics](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*.
- 2020 **R. Thomas McCoy**, Junghyun Min, and Tal Linzen. [BERTs of a feather do not generalize together: Large variability in generalization across models with similar test set performance](#). In *BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*.
- 2019 Najoung Kim, Roma Patel, Adam Poliak, Alex Wang, Patrick Xia, **R. Thomas McCoy**, Ian Tenney, Alexis Ross, Tal Linzen, Benjamin Van Durme, Samuel R. Bowman, Ellie Pavlick. [Probing What Different NLP Tasks Teach Machines about Function Word Comprehension](#). In *Proceedings of the Eighth Joint Conference on Lexical and Computational Semantics (*SEM 2019)*.
Best paper award at *SEM 2019.
- 2019 **R. Thomas McCoy**, Tal Linzen, Ewan Dunbar, and Paul Smolensky. [RNNs implicitly implement tensor-product representations](#). *International Conference on Learning Representations (ICLR)*.

- 2019 Ian Tenney, Patrick Xia, Berlin Chen, Alex Wang, Adam Poliak, **R. Thomas McCoy**, Najoung Kim, Benjamin Van Durme, Samuel R. Bowman, Dipanjan Das, and Ellie Pavlick. [What do you learn from context? Probing for sentence structure in contextualized word representations](#). *International Conference on Learning Representations (ICLR)*.
- 2019 **R. Thomas McCoy**, Ellie Pavlick, and Tal Linzen. [Right for the Wrong Reasons: Diagnosing Syntactic Heuristics in Natural Language Inference](#). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*.
- 2019 **R. Thomas McCoy**. [Touch down in Pittsburghese](#). *Yale Working Papers in Grammatical Diversity*.
- 2019 Samuel R. Bowman, Ellie Pavlick, Edouard Grave, Benjamin Van Durme, Alex Wang, Jan Hula, Patrick Xia, Raghavendra Pappagari, **R. Thomas McCoy**, Roma Patel, Najoung Kim, Ian Tenney, Yinghui Huang, Katherin Yu, Shuning Jin, and Berlin Chen. [Can You Tell Me How to Get Past Sesame Street? Sentence-Level Pretraining Beyond Language Modeling](#). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*.
- 2019 **R. Thomas McCoy** and Tal Linzen. [Non-entailed subsequences as a challenge for natural language inference](#). In *Proceedings of the Society for Computation in Linguistics (SCiL) 2019*.
- 2018 **R. Thomas McCoy**, Robert Frank, and Tal Linzen. [Revisiting the poverty of the stimulus: hierarchical generalization without a hierarchical bias in recurrent neural networks](#). In *Proceedings of the 40th Annual Conference of the Cognitive Science Society*.
- 2018 Patrick Littell, **R. Thomas McCoy**, Na-Rae Han, Shruti Rijhwani, Zaid Sheikh, David Mortensen, Teruko Mitamura, and Lori Levin. [Parser combinators for Tigrinya and Oromo morphology](#). In *Language Resources and Evaluation Conference (LREC) 2018*.
- 2018 **R. Thomas McCoy** and Robert Frank. [Phonologically Informed Edit Distance Algorithms for Word Alignment with Low-Resource Languages](#). In *Proceedings of the Society for Computation in Linguistics (SCiL) 2018*.
- 2017 Jungo Kasai, Robert Frank, **R. Thomas McCoy**, Owen Rambow, and Alexis Nasr. [TAG parsing with neural networks and vector representations of supertags](#). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- 2017 Dan Friedman*, Jungo Kasai*, **R. Thomas McCoy***, Robert Frank, Forrest Davis, and Owen Rambow. [Linguistically Rich Vector Representations of Supertags for TAG Parsing](#). In *Proceedings of the 13th International Workshop on Tree Adjoining Grammars and Related Formalisms*.

*Equal contribution.

- 2017 **R. Thomas McCoy**. [English comparatives as degree-phrase relative clauses](#). In *Proceedings of the Linguistic Society of America 2*.

WORK IN PREPARATION

R. Thomas McCoy, Paul Smolensky, Tal Linzen, Jianfeng Gao, and Asli Celikyilmaz. [How much do language models copy from their training data? Evaluating linguistic novelty in text generation using RAVEN](#). arXiv preprint.

Aditya Yedetore, Tal Linzen, Robert Frank, and **R. Thomas McCoy**. How poor is the stimulus? Evaluating hierarchical generalization in neural networks trained on child-directed speech. In preparation to submit to ACL Rolling Review.

R. Thomas McCoy, Tal Linzen, and Paul Smolensky. DISCOVER: A framework for dissecting compositionality in vector representations. In preparation to submit to the *Journal of Artificial Intelligence Research*.

UNPUBLISHED CONFERENCE PRESENTATIONS

2020 **R. Thomas McCoy**, Tal Linzen, Ewan Dunbar, and Paul Smolensky. Tensor product decomposition networks: Uncovering representations of structure learned by neural networks. Poster presentation, *Society for Computation in Linguistics*, New Orleans, Louisiana, January 2.

2018 **R. Thomas McCoy**, Robert Frank, and Tal Linzen. Investigating hierarchical bias in the acquisition of English question formation with recurrent neural networks. Poster presentation, *2018 Legrain conference: Learning Language in Humans and in Machines*, Paris, France, July 6.

2018 Robert Frank, **R. Thomas McCoy**, and Tal Linzen. Neural network syntax in the age of deep learning: The case of question formation. Oral presentation, *Society for Computation in Linguistics*, Salt Lake City, Utah, January 5.

2017 Patrick Littell, **R. Thomas McCoy**, and Lori Levin. The North American Computational Linguistics Olympiad. Oral presentation, in Datablitz: Getting High School Students into Linguistics: Current Activities and Future Directions, *Linguistic Society of America Annual Meeting*, Austin, Texas, January 7.

INVITED TALKS

2022 Google NLX Conversation Group Meeting. January 27, 2022.
How much do language models copy from their training data? Evaluating linguistic novelty in text generation using RAVEN.

2021 Montreal Computational and Quantitative Linguistics Lab at McGill (MCQLL). October 26, 2021.
Discovering implicit compositional representations in neural networks

- 2021 Edinburgh Centre for Language Evolution. September 28, 2021.
How do neural networks represent compositional symbolic structure?
- 2021 USC ISI Natural Language Seminar. April 15, 2021.
Universal linguistic inductive biases via meta-learning.
- 2020 DeepMind language reading group. December 7, 2020.
Analyzing the syntactic inductive biases of sequence-to-sequence networks.
- 2020 Berkeley NLP Seminar. October 16, 2020.
Analyzing the syntactic inductive biases of sequence-to-sequence networks.
- 2020 NLP With Friends seminar series. August 12, 2020.
Universal linguistic inductive biases via meta-learning.
- 2019 Workshop on Gradient Symbolic Computation. Johns Hopkins University.
September 19, 2019.
Tensor product decomposition of continuous vector representations
- 2018 Microsoft Research, Redmond. December 11, 2018.
Discovering the compositional structure implicitly learned by neural networks

MENTORING

Master's students

- 2019–2020 Junghyun Min
Co-supervised with Tal Linzen.
- 2019–2020 Paul Soulos
Co-supervised with Paul Smolensky.

Undergraduate students

- 2019–present Aditya Yedetore
Co-supervised with Tal Linzen.
- 2019–2020 Michael Lepori
Co-supervised with Tal Linzen.

OUTREACH AND CONTRIBUTIONS TO DIVERSITY

- 2020–present Johns Hopkins Cognitive Science Representation and Diversity Committee
Co-created and co-organized a program for giving feedback on PhD applications to prospective students who belong to underrepresented groups. Served as a mentor for 6 prospective students.

- 2020 Johns Hopkins Cognitive Science syllabus section for raising awareness about research and graduate school.
With one other graduate student, wrote a statement that faculty members added to their syllabi and course discussions describing how to pursue research opportunities and graduate school, in order to raise awareness of these opportunities among a broader group of undergraduates.
- 2020 Public talk for the National Museum of Language: *Language Squared: The Linguistics of Crosswords.*
- 2013–present North American Computational Linguistics Olympiad (NACLO).
Contest that introduces high school students to computational linguistics, with 1000 to 1500 students participating each year. Last year, 42% of participants were female, a high proportion for a computational initiative. National level: Co-Program Chair; problem writer (20 problems to date). Local level: Co-founder and co-organizer of the Yale contest site (2013–2017); co-organizer of the Johns Hopkins contest site (2017–2021); organizer of pre-contest practice sessions at both sites.
- 2018–2019 International Linguistics Olympiad (IOL): Problem writer.
- 2016 Yale Grammatical Diversity Project
Authored two webpages describing regional grammatical phenomena (all the further, subject contact relatives).
- 2013–2017 Linguistics teaching initiatives
Designed and taught a one-lecture linguistics class to high school students in New Haven in connection with the programs Splash, Sprout, and Math Mornings. Presented 8 times to groups ranging from 25 to 50 students.

SERVICE

- 2019–present Departmental representative for the Department of Cognitive Science in the Johns Hopkins Graduate Representative Organization.
- 2016–2017 Computational Linguistics at Yale (CLAY) reading group: Co-organizer.
- 2015–2017 Yale Undergraduate Linguistics Society: Co-founder (2015), president (2015–2016), treasurer (2016–2017).

REVIEWING

- 2021 Workshop reviewer: SCiL 2022
- 2021 Workshop reviewer: BlackboxNLP 2021.
- 2021 Journal reviewer: Natural Language Engineering.
- 2021 Conference reviewer: EMNLP 2021. Recognized as an outstanding reviewer.
- 2020 Conference reviewer: CoNLL 2020.
- 2020 Conference reviewer: EMNLP 2020. Recognized as an outstanding reviewer.
- 2020 Conference reviewer: ACL 2020.
- 2019 Journal reviewer: Language Acquisition.
- 2019 Conference reviewer: CoNLL 2019.

2018 Conference reviewer: CoNLL 2018.
2018 Conference reviewer: ACL 2018. Recognized as a top reviewer.

PROFESSIONAL MEMBERSHIPS

2015–present Linguistic Society of America (LSA).
2017–present Association for Computational Linguistics (ACL).
2018–present Cognitive Science Society.

SKILLS

Programming languages Python, PyTorch, JavaScript, Haskell, C, Java, R, Scheme.
Natural languages English (native), Bahasa Indonesia (basic conversation), Old English (basic reading ability), Old Norse (basic reading ability), Latin (basic reading ability).

COURSEWORK

Undergraduate GPA: 4.0 Graduate GPA: 4.0

Computational Linguistics: Language and Computation I, Language and Computation II, Formal Foundations of Linguistic Theories, Computing Meaning

Natural Language Processing: Natural Language Processing, Machine Learning: Linguistic and Sequence Modeling

Syntax: Syntax I, Syntax II, Grammatical Diversity in US English

Phonetics/Phonology: Phonetics, Phonology I, Phonology II, The Phonetics/Phonology Interface

Semantics: Semantics I, Semantics II

Computer Science: Data Structures and Programming Techniques, Computational Tools for Data Science

Mathematics: Multivariable Calculus, Discrete Mathematics, Probability and Statistics, Advanced Statistical Methods

Other relevant courses: Linguistic Field Methods, Foundations of Cognitive Science