

Jennifer Hu

Contact

Email: jenniferhu@fas.harvard.edu
Website: jennhu.github.io

Academic Positions

Eff. July 2025 **Johns Hopkins University**
Assistant Professor, Department of Cognitive Science

2024–Present **Johns Hopkins University**
Assistant Research Professor, Department of Cognitive Science

2023–Present **Harvard University**
Research Fellow, Kempner Institute for the Study of Natural and Artificial Intelligence

Education

2018–2023 **Massachusetts Institute of Technology**
Ph.D. in Cognitive Science
Dissertation: “Neural language models and human linguistic knowledge”
Advisor: Roger Levy
Committee: Joshua Tenenbaum (chair), Christopher Potts, Evelina Fedorenko

2014–2018 **Harvard University**
B.A. in Mathematics and Linguistics
Secondary Field in Germanic Languages & Literatures
Magna cum laude with Highest Honors

Experience

2022 **Allen Institute for Artificial Intelligence**
Summer Research Intern (Mosaic Team)
Advisor: Prithviraj Ammanabrolu

2017 **Stanford University Center for the Study of Language and Information**
Summer Research Intern
Advisor: Christopher Potts

2017 **Harvard University Program for Research in Science and Engineering**
Research Fellow (Department of Computer Science)
Advisor: Stuart Shieber

Publications

Journal Articles

- [1] Jennifer Hu, Felix Sosa, and Tomer Ullman. “Re-evaluating Theory of Mind evaluation in large language models”. *Philosophical Transactions of the Royal Society B* (2025). To appear.
- [2] Jennifer Hu, Felix Sosa, and Tomer Ullman. “Shades of Zero: Distinguishing impossibility from inconceivability”. *Journal of Memory and Language* (2025). To appear.
- [3] Jennifer Hu, Ethan Wilcox, Siyuan Song, Kyle Mahowald, and Roger Levy. “What can string probability tell us about grammaticality?” *Transactions of the Association for Computational Linguistics* (2025). To appear.

- [4] Jennifer Hu, Kyle Mahowald, Gary Lupyan, Anna Ivanova, and Roger Levy. “Language models align with human judgments on key grammatical constructions”. *Proceedings of the National Academy of Sciences* (2024).
- [5] Jennifer Hu, Roger Levy, Judith Degen, and Sebastian Schuster. “Expectations over unspoken alternatives predict pragmatic inferences”. *Transactions of the Association for Computational Linguistics* (2023).
- [6] Jennifer Hu, Hannah Small, Hope Kean, Atsushi Takahashi, Leo Zelekman, Daniel Kleinman, Elizabeth Ryan, Alfonso Nieto-Castañón, Victor Ferreira, and Evelina Fedorenko. “Precision fMRI reveals that the language-selective network supports both phrase-structure building and lexical access during language production”. *Cerebral Cortex* (2022).

Book Chapters

- [1] Ethan Wilcox, Jon Gauthier, Jennifer Hu, Peng Qian, and Roger Levy. “Learning syntactic structures from string input”. *Algebraic Structures in Natural Language*. Ed. by Shalom Lappin and Jean-Philippe Bernardy. Taylor & Francis, 2023.

Conference Papers

- [1] Sonia K. Murthy, Tomer Ullman, and Jennifer Hu. “One fish, two fish, but not the whole sea: Alignment reduces language models’ conceptual diversity”. *Proceedings of the 2025 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. 2025.
- [2] Junyi Chu, Jennifer Hu, and Tomer Ullman. “The Task Task: Creative problem generation in humans and language models”. *Proceedings of the Cognitive Science Society*. 2024.
- [3] Jennifer Hu and Michael C. Frank. “Auxiliary task demands mask the capabilities of smaller language models”. *Proceedings of the Conference on Language Modeling*. **Outstanding Paper Award**. 2024.
- [4] Jennifer Hu, Felix Sosa, and Tomer Ullman. “Shades of Zero: Distinguishing impossibility from inconceivability”. *Proceedings of the Cognitive Science Society*. 2024.
- [5] Daniel Fried, Nicholas Tomlin, Jennifer Hu, Roma Patel, and Aida Nematzadeh. “Pragmatics in Grounded Language Learning: Phenomena, Tasks, and Modeling Approaches”. *Findings of the Association for Computational Linguistics: EMNLP 2023*. 2023.
- [6] Jennifer Hu, Sammy Floyd, Olessia Jouravlev, Evelina Fedorenko, and Edward Gibson. “A fine-grained comparison of pragmatic language understanding in humans and language models”. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. 2023.
- [7] Jennifer Hu and Roger Levy. “Prompting is not a substitute for probability measurements in large language models”. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 2023.
- [8] Pei Zhou, Andrew Zhu, Jennifer Hu, Jay Pujara, Xiang Ren, Chris Callison-Burch, Yejin Choi, and Prithviraj Ammanabrolu. “I Cast Detect Thoughts: Learning to Converse and Guide with Intent and Theory-of-Mind in Dungeons and Dragons”. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. 2023.
- [9] Irene Zhou, Jennifer Hu, Roger Levy, and Noga Zaslavsky. “Teasing apart models of pragmatics using optimal reference game design”. *Proceedings of the Cognitive Science Society*. 2022.
- [10] Jennifer Hu, Noga Zaslavsky, and Roger Levy. “Competition from novel features drives scalar inferences in reference games”. *Proceedings of the Cognitive Science Society*. 2021.
- [11] Yiwen Wang, Jennifer Hu, Roger Levy, and Peng Qian. “Controlled Evaluation of Grammatical Knowledge in Mandarin Chinese Language Models”. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. 2021.
- [12] Jon Gauthier, Jennifer Hu, Ethan Wilcox, Peng Qian, and Roger Levy. “SyntaxGym: An Online Platform for Targeted Evaluation of Language Models”. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Online: Association for Computational Linguistics, July 2020, pp. 70–76.

- [13] Jennifer Hu, Sherry Yong Chen, and Roger Levy. “A closer look at the performance of neural language models on reflexive anaphor licensing”. *Proceedings of the Society for Computation in Linguistics*. Vol. 3. 2020, pp. 382–392.
- [14] Jennifer Hu, Jon Gauthier, Peng Qian, Ethan Wilcox, and Roger Levy. “A Systematic Assessment of Syntactic Generalization in Neural Language Models”. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, July 2020, pp. 1725–1744.
- [15] Ethan Wilcox, Jon Gauthier, Jennifer Hu, Peng Qian, and Roger Levy. “On the predictive power of neural language models for human real-time comprehension behavior”. *Proceedings of the Cognitive Science Society*. 2020.
- [16] Jennifer Hu, James Traer, and Josh H. McDermott. “Separating object resonance and room reverberation in impact sounds”. *Proceedings of the Cognitive Science Society*. 2019.
- [17] Will Monroe, Jennifer Hu, Andrew Jong, and Christopher Potts. “Generating Bilingual Pragmatic Color References”. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. New Orleans, Louisiana: Association for Computational Linguistics, June 2018, pp. 2155–2165.

Workshop Papers

- [1] Jennifer Hu and Michael Franke. *Deep and shallow thinking in a single forward pass*. Workshop on Behavioral Machine Learning @ NeurIPS 2024. 2024.
- [2] Jennifer Hu, Roger Levy, and Sebastian Schuster. *Predicting scalar diversity with context-driven uncertainty over alternatives*. ACL Workshop on Cognitive Modeling and Computational Linguistics. 2022.
- [3] Jennifer Hu, Roger Levy, and Noga Zaslavsky. *Scalable pragmatic communication via self-supervision*. ICML Workshop on Self-Supervised Learning for Reasoning and Perception. 2021.

Extended Abstracts

- [1] Jennifer Hu, Roger Levy, and Sebastian Schuster. *Predicting scalar diversity with context-driven expectations*. *Proceedings of the Experimental Pragmatics Conference (XPRAG)*. 2022.
- [2] Yiwen Wang, Jennifer Hu, Roger Levy, and Peng Qian. *Facilitative Effect Induced by Classifier-Noun Mismatch in Mandarin Chinese*. *The 35th Annual Conference on Human Sentence Processing*. 2022.
- [3] Noga Zaslavsky, Jennifer Hu, and Roger Levy. *A Rate–Distortion view of human pragmatic reasoning*. *Proceedings of the Society for Computation in Linguistics*. 2021.
- [4] Irene Zhou, Jennifer Hu, Roger Levy, and Noga Zaslavsky. *Empirical support for a Rate–Distortion account of pragmatic reasoning*. *Proceedings of the Cognitive Science Society*. Member abstract. 2021.
- [5] Jennifer Hu, Hannah Small, Hope Kean, Atsushi Takahashi, Leo Zekelman, Daniel Kleinman, Elizabeth Ryan, Victor Ferreira, and Evelina Fedorenko. *Distributed and overlapping neural mechanisms for lexical access and syntactic encoding during language production*. *Proceedings of the Society for the Neurobiology of Language*. 2020.
- [6] Ethan Wilcox, Jon Gauthier, Jennifer Hu, Peng Qian, and Roger Levy. *Benchmarking neural networks as models of human language processing*. *Proceedings of the 26th Architectures and Mechanisms for Language Processing Conference*. 2020.
- [7] Ethan Wilcox, Jon Gauthier, Peng Qian, Jennifer Hu, and Roger Levy. *Evaluating the effect of model inductive bias and training data in predicting human reading times*. *Proceedings of the 33rd Annual CUNY Human Sentence Processing Conference*. 2020.
- [8] Noga Zaslavsky, Jennifer Hu, and Roger Levy. *Emergence of pragmatic reasoning from least-effort optimization*. *Proceedings of Evolution of Language International Conferences*. 2020.
- [9] Jennifer Hu. *A graph-theoretic approach to comparing typologies in Parallel OT and Harmonic Serialism*. *Proceedings of the 92nd Annual Meeting of the Linguistic Society of America*. Salt Lake City, UT, 2018.

Awards

- 2024 Outstanding Paper Award, COLM 2024
- 2024 Harvard University Hodgson Memorial Fund
- 2021 National Science Foundation Doctoral Dissertation Research Improvement Grant
- 2019 Computationally-Enabled Integrative Neuroscience Training Program
- 2019 National Science Foundation Graduate Research Fellowship
- 2018 Thomas T. Hoopes Prize
- 2018 Friends of Harvard Mathematics Prize
- 2017 Harvard College Research Program Grant
- 2017 Robert Fletcher Rogers Prize
- 2015 Detur Book Prize
- 2015 John Harvard Scholarship

Invited Talks

- 2025 *Title TBD*
Experimental Pragmatics Conference
- 2025 “Large language models and human linguistic knowledge”
Symposium: What Big Data Can (and Can’t!) Tell Us About How Language Works
Annual Meeting of the American Association for the Advancement of Science
- 2024 “Cognitive evaluation of language models”
Tutorial: Experimental Design and Analysis for AI Researchers
NeurIPS 2024
- 2024 “How to know what language models know”
NYU NLP and Text-as-Data Speaker Series
- 2024 “How to know what language models know”
University of Oxford NLP Group
- 2024 “How to know what language models know”
Stanford NLP Group
- 2023 “Using artificial language models to test linguistic theories: Case studies and caveats”
Harvard Language and Cognition Reading Group
- 2023 “Neural language models and human linguistic knowledge”
Harvard Department of Psychology Cognition, Brain, and Behavior Seminar Series
- 2023 “Neural language models and human linguistic knowledge”
International Interdisciplinary Computational Cognitive Science Summer School
- 2023 “Cognitive benchmarking of neural language models: A case study in pragmatics”
Workshop: Advancing Cognitive Science and AI with Cognitive-AI Benchmarking
CogSci 2023
- 2022 “A targeted evaluation of human-like linguistic knowledge in neural language models”
Brown University BigAI Group
- 2022 “Investigating ad-hoc scalar implicatures”
University of Tübingen Department of Linguistics
- 2021 “Competition from novel features drives scalar inferences in reference games”
Harvard Language and Cognition Reading Group
- 2020 “Benchmarking neural networks as models of human language processing”
Google DeepMind

Teaching

Teaching assistant positions

- 2021 Language in the Mind and Brain (9.S52), MIT
- 2020 Computational Psycholinguistics (9.19/9.190), MIT
- 2018 Paradoxes and Infinities (PDOX), Johns Hopkins University Center for Talented Youth
- 2016 Linear Algebra and Real Analysis II (MATH 23B), Harvard
- 2015 Linear Algebra and Real Analysis I (MATH 23A), Harvard
- 2015 Vectors: A Tool for Teaching Algebra, Geometry, and Trigonometry (MATH S-323), Harvard

Invited guest lectures

- 2024 “Neural language models and human linguistic knowledge”
University of California Irvine
- 2022 “What do language models know about meaning?”
The Science of Intelligence (9.58), MIT
- 2020 “Language understanding in minds and machines”
Language, Structure, and Cognition (LING 83), Harvard
- 2016 “Testing synchronous tree-adjoining grammar analyses of linguistic phenomena”
Topics in Computational Linguistics (LING 98A), Harvard

Service

Organizing

- 2026 Dagstuhl Seminar on Social Intelligence in AI Systems
- 2024 NeurIPS Tutorial on Experimental Design and Analysis for AI Researchers
- 2023 ICML Workshop on Theory of Mind in Communicating Agents
- 2022 NeurIPS Meaning in Context Workshop

Reviewing

Conference editorial responsibilities:

- Area Chair for COLM (2025), EMNLP (2024), NAACL (2024)

Ad-hoc journal reviewing:

- Glossa (2024)
- Computational Linguistics (2024)
- Philosophical Transactions of the Royal Society B (2024)
- Nature Human Behaviour (2024)
- Proceedings of the National Academy of Sciences (2024)
- Cognitive Science (2024)
- Mind and Language (2024)
- Journal of Memory and Language (2023)
- Open Mind (2022)
- Linguistics and Philosophy (2021)
- Language, Cognition and Neuroscience (2021)

Ad-hoc conference reviewing:

- ACL (2021)

- EMNLP (2022)
- ARR (Oct 2021, Nov 2021, Jan 2022, Apr 2022)
- CogSci (2020, 2021, 2022, 2023, 2024, 2025)

Ad-hoc workshop reviewing:

- BehavioralML Workshop (NeurIPS 2024)
- Workshop on Theory of Mind in Human-AI Interaction (CHI 2024)
- Workshop on Large Language Models and Cognition (ICML 2024)
- UnImplicit Workshop (NAACL 2022, EACL 2024)
- Workshop on Theory of Mind in Communicating Agents (ICML 2023)
- CoNLL (EMNLP 2020, 2021, 2022)

Ad-hoc grant proposal reviewing:

- National Science Foundation (2023)

Advocacy

2020-2021 Member of MIT School of Science Graduate Council
 2019-2021 Committee member of MIT Women’s Advisory Group
 2019-2021 Co-Chair of Graduate Women at MIT

Mentorship

Supervised undergraduates

2024 Jōsh Mysore, Harvard
 2024 Siyuan Song, UT Austin
 2024 Antara Bhattacharya, Harvard
 2020-2022 Irene Zhou, MIT
 2019 Eric Hong, MIT

Other mentorship

2023 Harvard Psychology PPREP Program
 2022 MIT-Harvard Women in AI
 2018-2019 Non-Resident Tutor at Mather House, Harvard University