

Huteng Dai

University of Michigan
611 Tappan Ave, 462 Lorch Hall
Ann Arbor, MI 48109 USA

Email: huteng@umich.edu
Personal website: hutengdai.com

Employment

2024 Assistant Professor, Dept. of Linguistics, University of Michigan, Ann Arbor.

Education

2024 PhD, Department of Linguistics, Rutgers University, New Brunswick.
Dissertation: *Phonological Learning in the Presence of Exceptions*
Committee: Adam Jardine (Chair), Bruce Tesar, Adam McCollum, Richard Futrell
2022 Certificate in Cognitive Science, Rutgers Center for Cognitive Science.
2018 BA, Capital Normal University, Beijing.

NON-DEGREE

2024 Deep Learning Specialization certificate, Coursera & DeepLearningAI.
2023 Machine Learning Specialization certificate, Coursera & DeepLearningAI.
2019 LSA Linguistics Institute, University of California, Davis.
2017 LSA Linguistics Institute, University of Kentucky.
2017 Department of Linguistics, University of California, Santa Barbara.

Publications

JOURNAL ARTICLES

In prep. **Huteng Dai**. Structured learning in noisy data: case studies of Navajo and Quechua.
submitted **Huteng Dai**. Towards a two-stage phonotactic-alternation learning model. *Language*.
to appear **Huteng Dai**. An exception-filtering approach to phonotactic learning. *Phonology*.
2019 Zihe Li, Jingjing Pan, and **Huteng Dai**. Nasal sound changes in Naish languages. *Bulletin of Linguistic Studies*.

PAPER-REVIEWED CONFERENCE PROCEEDINGS

- 2023 **Huteng Dai**, Connor Mayer, and Richard Futrell. Rethinking representations: A log-bilinear model of phonotactics. *Proceedings of the Society for Computation in Linguistics*.
- 2021 **Huteng Dai** and Richard Futrell. Simple induction of (deterministic) probabilistic finite-state automata for phonotactics by stochastic gradient descent. In *Proceedings of the 18th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*.

ABSTRACT-REVIEWED CONFERENCE PROCEEDINGS

- 2022 **Huteng Dai** and Richard Futrell. Learning phonotactics in a differentiable framework of subregular languages. In *Proceedings of the Annual Meetings on Phonology*, volume 9.
- 2021 **Huteng Dai**. Learning Nonlocal Phonotactics in a Strictly Piecewise Probabilistic Phonotactic Model. In *Proceedings of the Annual Meetings on Phonology*.
- 2021 **Huteng Dai**. Gradient similarity in Lezgian laryngeal harmony: representation and computation. In *Proceedings of the 38th West Coast Conference on Formal Linguistics*. Cascadilla Press.
- 2021 Wenyue Hua, Adam Jardine, and **Huteng Dai**. Learning Underlying Representations and Input-Strictly-Local Functions. In *Proceedings of the 37th West Coast Conference on Formal Linguistics*. Cascadilla Press.
- 2020 Tajudeen Mamadou Yacoubou, Mariapaola D’Imperio, **Huteng Dai**, and David Kleinschmidt. Global and local pitch level in caabe is not predicted by the frequency code. *The Journal of the Acoustical Society of America*.
- 2020 **Huteng Dai** & Richard Futrell. Information-theoretic characterization of the sub-regular hierarchy. In *Proceedings of the Society for Computation in Linguistics*.

Selected Grant, Fellowships, and Awards

- Submitted **NSF CompCog: Learning Phonotactics by Gradient Descent with Formal Restrictions** (collaboration with Connor Mayer and Richard Futrell). National Science Foundation.
- Fall 2023 **Louis Bevier Dissertation Completion Fellowship**. Rutgers University.
- May 2018 **Best BA Thesis**. Capital Normal University.

Selected Presentations

INVITED TALKS

- 2025 **Huteng Dai.** Phonological learning from noisy data. Invited talk at the LSA 2025 Annual Meeting, Workshop on Computational Models of Learnability and Acquisition of Morphology and Phonology, Philadelphia.
- 2022 **Huteng Dai.** Neo-Trubetzkoyan Learner—a categorical approach to robust phonotactic learning. Invited Talk for Experimental/Computational Linguistics group at Department of Linguistics and Philosophy, Massachusetts Institute of Technology, online
- 2022 **Huteng Dai.** Learning phonotactics in the presence of exceptions with a categorical approach. Invited Talk for the QuantLang Lab at the Department of Language Science, University of California, Irvine, CA.

PEER-REVIEWED CONFERENCE PRESENTATIONS

- 2023 **Huteng Dai**, Connor Mayer, and Richard Futrell. Rethinking representations: A log-bilinear model of phonotactics. Talk presented at Society for Computation in Linguistics (SCiL), UMass Amherst, MA.
- 2022 **Huteng Dai.** Phonotactic learning in the presence of exceptions with a categorical approach. Talk presented at Annual Meeting on Phonology (AMP), University of California, Los Angeles, CA.
- 2021 **Huteng Dai** and Richard Futrell. Learning Phonotactics in a Differentiable Framework of Subregular Languages. Talk presented at Annual Meeting on Phonology (AMP), online.
- 2021 **Huteng Dai** and Richard Futrell. Simple induction of (deterministic) probabilistic finite-state automata for phonotactics by stochastic gradient descent. Talk presented at ACL Workshop on Computational Research in Phonetics, Phonology, and Morphology (SIGMORPHON), online.
- 2021 **Huteng Dai.** Learning nonlocal phonotactics in Strictly Piecewise phonotactic model: recent progress. Talk presented at Society for Computation in Linguistics (SCiL), online.
- 2020 **Huteng Dai.** Structure matters: Learning nonlocal phonotactics in Strictly Piecewise phonotactic model. Talk presented at Annual Meeting on Phonology (AMP), online.
- 2020 **Huteng Dai.** Gradient similarity in Lezgian laryngeal harmony: representation

and computation. Talk presented at West Coast Conference in Formal Linguistics (WCCFL), University of British Columbia, Vancouver, BC Canada.

2020 Wenye Hua, Adam Jardine, and **Huteng Dai**. A subregular approach to the problem of learning underlying representations. Poster presented at West Coast Conference in Formal Linguistics (WCCFL), University of British Columbia, Vancouver, Vancouver, BC Canada.

2020 **Huteng Dai**. Gradient similarity in Lezgian laryngeal harmony: representation and computation. Talk presented at Berkeley Linguistics Society (BLS) Workshop “Phonological Representations: at the crossroad between gradience and categoricity”, University of California, Berkeley, CA.

Teaching & Fieldwork Experience

2025 Instructor, UNIVERSITY OF MICHIGAN
 Winter/Spring Language & Information (grad-level; cross-listed with EECS and Information);

2025 Instructor, UNIVERSITY OF MICHIGAN
 Winter/Spring ABC’s of Python;

2024 Fall Instructor, UNIVERSITY OF MICHIGAN
 The Linguistics of Speech Recognition;

2023 Spring Instructor, RUTGERS UNIVERSITY
 Introduction to Linguistic Theory (Honors College);

2021 Fall-2022 Instructor, RUTGERS UNIVERSITY
 Spring Linguistic Perspectives on Language: Invented Languages;

2021 Spring Instructor, RUTGERS UNIVERSITY
 Introduction to Linguistic Theory;

2020 Fall Teaching Assistant, RUTGERS UNIVERSITY
 Introduction to Linguistic Theory.

2018.3-2018.6 Teaching Assistant, CAPITAL NORMAL UNIVERSITY
 The Phonetics and Phonology of Chinese;

2018.7-2018.8 Teaching Assistant, LINGUISTICS INSTITUTE OF CHINA
 Phonology II (Lian-Hee Wee); Semantics II (Haihua Pan); Syntax II (Wei-Tien Dylan Tsai);

2017.1 Fieldwork in Jiangxi Linchuan, Linchuan dialect

2016.7 Fieldwork in Yunnan Lijiang, Naish language

Reviewing

- 2024 Area Chair (Speech and Phonology), Conference on Computational Natural Language Learning (CoNLL)
- 2021 Society for Computation in Linguistics

Services & Leadership

- 2024 Lab Director, Computation, Acquisition, Processing (CAP) Lab, U-M
- 2024 Graduate Admission Committee, Department of Linguistics, U-M
- 2023 Graduate Mentor, Rutgers Honors College
- 2021 Volunteer and Session Chair, The 53rd Annual Meeting of the North East Linguistic Society at Rutgers
- 2019-2023 Organizer, Rutgers Phonology and Phonetics Research Group (PHONX)
- 2019-2020 Organizer, Rutgers Mathematical and Computational Linguistics Group
- 2020.1-2020.7 Tenure-track job search committee (Phonology)
- 2018-2020 Editor, Facebook Page, Department of Linguistics, Rutgers University
- 2018 Founder & Organizer, BEIJING P-SIDE READING GROUP

Skills

NATURAL LANGUAGES: Mandarin Chinese (Native), English (Fluent);

PROGRAMMING LANGUAGES: Python; R;

MACHINE LEARNING: numpy, pytorch, tensorflow, scikit-learn, cuda;

OTHER TOOLS: L^AT_EX; git; HTML, CSS, jsPsych, PsychoPy, Pavlovia (Web-based Experiment); EyeLink (Eyetracking), FSL (Neuroimaging); Praat, ELAN.

Memberships

ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL)

LINGUISTIC SOCIETY OF AMERICA (LSA)