

Huteng Dai

University of Michigan

611 Tappan Ave, 462 Lorch Hall

Ann Arbor, MI 48109 USA

Email: huteng@umich.edu

Web: hutengdai.com

Employment

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

Education

- 2024 **PhD**, Dept. 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 (CNU), Beijing.
Advisor: Zihe Li

NON-DEGREE

- 2024 Deep Learning Specialization cert., Coursera & DeepLearningAI.
- 2023 Machine Learning Specialization cert., Coursera & DeepLearningAI.
- 2019 LSA Linguistics Institute, UC Davis.
- 2017 LSA Linguistics Institute, Univ. of Kentucky.
- 2017 Dept. of Linguistics, UC Santa Barbara.

Publications

JOURNAL ARTICLES

- 2025 **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

- 2025 Chi-Yun Chang, Xueyang Huang, Humaira Nasir, Shane Storks, Olawale Akingbade, and **Huteng Dai**. Mind the gap: How BabyLMs learn filler-gap dependencies. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025.
- 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** and Richard Futrell. Information-theoretic characterization of the subregular hierarchy. In *Proceedings of the Society for Computation in Linguistics*.

Selected Grants, Fellowships, and Awards

- In prep. **NSF: Learning Phonotactics by Gradient Descent with Formal Restrictions** (with Connor Mayer and Richard Futrell).

Fall 2023 **Louis Bevier Dissertation Completion Fellowship.** Rutgers Univ.

Selected Presentations

INVITED TALKS

- 2026 **Huteng Dai.** Interpreting Language Models Trained on Cognitively Plausible Corpora (Open House Colloquium). Dept. of Linguistics, U-M.
- 2025 **Huteng Dai.** Machine Learning for Linguists. Peking University (PKU), Beijing.
- 2025 **Huteng Dai.** Machine Learning for Linguists. The Education University of Hong Kong (EdUHK).
- 2025 **Huteng Dai.** Mind Matters: Computational Linguistics for Cognitive Scientists. Weinberg Inst. for Cognitive Science, U-M.
- 2025 **Huteng Dai.** Mind the gap: How BabyLMs learn filler-gap dependencies. Human-Centered AI Symposium. School of Information (SI), U-M.
- 2025 **Huteng Dai.** Phonological learning from noisy data. Workshop on Computational Models of Learnability and Acquisition of Morphology and Phonology, LSA 2025 Annual Meeting, Philadelphia.
- 2022 **Huteng Dai.** Neo-Trubetzkoyan Learner—a categorical approach to robust phonotactic learning. Experimental/Computational Linguistics group, Dept. of Ling. & Phil., MIT (online).
- 2022 **Huteng Dai.** Learning phonotactics in the presence of exceptions with a categorical approach. QuantLang Lab, Dept. of Lang. Sci., UC Irvine.

PEER-REVIEWED CONFERENCE PRESENTATIONS

- 2023 **Huteng Dai**, Connor Mayer, and Richard Futrell. Rethinking representations: A log-bilinear model of phonotactics. Talk at SCiL, UMass Amherst.
- 2022 **Huteng Dai.** Phonotactic learning in the presence of exceptions with a categorical approach. Talk at AMP, UCLA.
- 2021 **Huteng Dai** and Richard Futrell. Learning phonotactics in a differentiable framework of subregular languages. Talk at AMP (online).
- 2021 **Huteng Dai** and Richard Futrell. Simple induction of (deterministic) probabilistic finite-state automata for phonotactics by stochastic gradient descent. Talk at SIG-MORPHON (online).
- 2021 **Huteng Dai.** Learning nonlocal phonotactics in Strictly Piecewise phonotactic

model: recent progress. Talk at SCiL (online).

2020 **Huteng Dai**. Structure matters: Learning nonlocal phonotactics in Strictly Piece-wise phonotactic model. Talk at AMP (online).

2020 **Huteng Dai**. Gradient similarity in Lezgian laryngeal harmony: representation and computation. Talk at WCCFL, UBC, Vancouver.

2020 Wenyue Hua, Adam Jardine, and **Huteng Dai**. A subregular approach to the problem of learning underlying representations. Poster at WCCFL, UBC, Vancouver.

2020 **Huteng Dai**. Gradient similarity in Lezgian laryngeal harmony: representation and computation. Talk at BLS Workshop “Phonological Representations: at the crossroad between gradience and categoricity”, UC Berkeley.

Teaching & Fieldwork

2026 W/Sp Instructor, U-M
Language & Information (grad-level; cross-listed with CSE and SI)

2026 W/Sp Instructor, U-M
Phonology (grad-level)

2025 W/Sp Instructor, U-M
Language & Information (grad-level; cross-listed with CSE and SI)

2025 W/Sp Instructor, U-M
ABC’s of Python

2024 Fall Instructor, U-M
The Linguistics of Speech Recognition

2023 Sp Instructor, RUTGERS UNIV.
Intro. to Linguistic Theory (Honors College)

2021 F–22 Sp Instructor, RUTGERS UNIV.
Linguistic Perspectives on Language: Invented Languages

2021 Sp Instructor, RUTGERS UNIV.
Intro. to Linguistic Theory

2020 Fall TA, RUTGERS UNIV.
Intro. to Linguistic Theory

2018 TA, CNU
The Phonetics and Phonology of Chinese

2018 TA, LINGUISTICS INSTITUTE OF CHINA

Phonology II (Lian-Hee Wee); Semantics II (Haihua Pan); Syntax II (Wei-Tien Dylan Tsai)

2017 Fieldwork in Jiangxi Linchuan, Linchuan dialect

2016 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 (SCiL)

Service & Leadership

2024- PI, Computation, Acquisition, Processing (CAP) Lab, U-M

2025–2026 Tenure-track Job Search Comm., Dept. of Linguistics, U-M

2024–2025 Grad. Admissions Comm., Dept. of Linguistics, U-M

2023 Grad. Mentor, Rutgers Honors College

2021 Volunteer and Session Chair, NELS 53 at Rutgers

2019–2023 Organizer, Rutgers Phonology and Phonetics Research Group (PHONX)

2019–2020 Organizer, Rutgers Math. & Comp. Linguistics Group

2020 Tenure-track job search comm. (Phonology)

2018–2020 Editor, Facebook Page, Dept. of Linguistics, Rutgers Univ.

2018 Founder & Organizer, BEIJING P-SIDE READING GROUP

Skills

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

PROGRAMMING: Python; R; NumPy, PyTorch, TensorFlow

OTHER TOOLS: L^AT_EX; Git; HTML/CSS; jsPsych, PsychoPy; Praat; ELAN

Memberships

ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL)

LINGUISTIC SOCIETY OF AMERICA (LSA)