

# Huteng Dai

Rutgers University  
18 Seminary Place  
New Brunswick, NJ 08901-1184 USA

Email: [huteng.dai@rutgers.edu](mailto:huteng.dai@rutgers.edu)  
Personal website: [hutengdai.com](http://hutengdai.com)

## Research Interests

Computational Linguistics, Phonology, Phonetics, Acquisition, Psycholinguistics

## Education

- 2018-2024 (expected) PhD student, Department of Linguistics, Rutgers University, New Brunswick.  
Dissertation: *Phonological Learning in the Presence of Exceptions*  
Committee: Adam Jardine (Chair), Bruce Tesar, Adam McCollum, Colin Wilson (Johns Hopkins University)
- 2018-2022 Certificate in Cognitive Science, Rutgers Center for Cognitive Science.  
Qualifying paper: *Lexical and Frequency Biases in the Perceptual Learning of Mandarin Tonal Categories*  
Committee: Adam McCollum (Chair), Mariapaola D'Imperio, Joseph Casillas
- 2014-2018 BA, Capital Normal University, Beijing.  
Thesis: A Constraint-Based Analysis of Disyllabic Tone Sandhi in Linchuan.  
Advisor: Zihe Li (now Peking University)

## NON-DEGREE

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

## Grant/Fellowships/Awards

- Fall 2023 **Louis Bevier Dissertation Completion Fellowship.** Rutgers University.
- Fall 2022 **School of Graduate Studies Travel Grant.** Rutgers University.
- Fall 2019 **Cognitive Science Center Travel Grant.** Rutgers University.

- 2018-2020 **Graduate Excellence Fellowship.** Rutgers University.  
 May 2018 **Best BA Thesis.** Capital Normal University.

## Publications

### JOURNAL ARTICLES

**Huteng Dai.** An exception-filtering approach to phonotactic learning. Revised & Resubmitted to *Phonology*.

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*, 148(4), 2020.

Zihe Li, Jingjing Pan, and **Huteng Dai**. Nasal sound changes in Naish languages. *Bulletin of Linguistic Studies* 语言研究集刊, 23, 2019.

### PAPER-REVIEWED CONFERENCE PROCEEDINGS

**Huteng Dai**, Connor Mayer, and Richard Futrell. Rethinking representations: A log-bilinear model of phonotactics. *Proceedings of the Society for Computation in Linguistics*, 6(1), 2023.

**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*, 2021.

### ABSTRACT-REVIEWED CONFERENCE PROCEEDINGS

**Huteng Dai** and Richard Futrell. Learning phonotactics in a differentiable framework of subregular languages. In *Proceedings of the Annual Meetings on Phonology*, volume 9, 2022.

**Huteng Dai.** Learning Nonlocal Phonotactics in a Strictly Piecewise Probabilistic Phonotactic Model. In *Proceedings of the Annual Meetings on Phonology*, volume 8, 2021.

**Huteng Dai.** Gradient similarity in Lezgian laryngeal harmony: representation and computation. In *Proceedings of the 38th West Coast Conference on Formal Linguistics*. Cascadia 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, 2021.

**Huteng Dai** & Richard Futrell. Information-theoretic characterization of the sub-regular hierarchy. In *Proceedings of the Society for Computation in Linguistics*, volume 3, 2020.

## Selected 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**. 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.
- 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 Tajudeen Mamadou Y, **Huteng Dai**, Yair Haendler and Mariapaola D’Imperio. The Role of Global and Local Pitch Levels in the Perception of Questions in Ede Chaabe. Poster presented at Phonetics and Phonology in Europe (PaPE), 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 Piece-

wise 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 Wenyue 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.

## Professional Experience

- 2023 Spring Instructor, RUTGERS UNIVERSITY  
Introduction to Linguistic Theory (Honor 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); Syntax II (Wei-Tien Dylan Tsai); Semantics II (Haihua Pan)
- 2017.1 Fieldwork in Jiangxi Linchuan, Linchuan dialect
- 2016.7 Fieldwork in Yunnan Lijiang, Naish language

## Reviewing

- 2021 Society for Computation in Linguistics

## Services and Leadership

- 2023 Graduate Mentor, Rutgers Honor College
- 2021 Volunteer and Session Chair, The 53rd Annual Meeting of the North East Linguistic Society at Rutgers
- 2019-present 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, cuda;

OTHER TOOLS: L<sup>A</sup>T<sub>E</sub>X; Praat; ELAN; git; PsychoPy, Pavlovia (Experimental design); EyeLink (Eyetracking), FSL (Neuroimaging); HTML, CSS (Web development);

## Memberships

Rutgers Computational Linguistics Lab  
UCI Language Processing Group

LINGUISTIC SOCIETY OF AMERICA (LSA)