Caleb Belth

2913 Languages and Communications Building
Salt Lake City, UT
Email: caleb.belth@utah.edu

Phone: 260-494-7633
Website: https://cbelth.github.io/

Positions

Assistant Professor Department of Linguistics, University of Utah, Salt Lake City, UT 2023-Present

Education

PhD, Computer Science, University of Michigan, Ann Arbor, MI 2018-2023
Advisors: Andries Coetzee, Danai Koutra
Thesis: *Towards an Algorithmic Account of Phonological Rules and Representations*
Committee: Andries Coetzee (LING), Danai Koutra (CS), Charles Yang (LING, UPenn), Richard Lewis (Psychology), Lu Wang (CS)

M.S., Computer Science, University of Michigan, Ann Arbor, MI 2018-2019

B.S., Computer Science, Purdue University, West Lafayette, IN 2014-2018
Minors: Philosophy, Mathematics
Research Advisors: Jennifer Neville, Dan Goldwasser, Daisuke Kihara

Research Interests

Phonology, Language Acquisition, Computational Linguistics, Linguistic Representations, Psycholinguistics, Linguistic Variation

Awards and Honors

Rackham Graduate School Travel Award 2022

Weinberg Institute of Cognitive Science Travel Award 2022

Rackham Graduate School Travel Award 2022

Weinberg Institute of Cognitive Science Travel Award 2021

Richard F. and Eleanor A. Towner Prize for Distinguished Academic Achievement 2021
Awarded to the outstanding graduate student in each degree program

Best paper candidate, IEEE ICDM 2020

NSF Graduate Research Fellowship 2020

NDSEG Fellowship (declined for NSF GRF) 2020

Rackham Graduate School Travel Award 2019

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) Travel Award 2019

Dean's List, Purdue 2015-2018

Semester Honors, Purdue 2015-2018

Journal Publications

3. **Caleb Belth**. In Press. A Learning-Based Account of Phonological Tiers. *Linguistic Inquiry*.
2. **Caleb Belth**. In Press. A Learning-Based Account of Local Phonological Processes. *Phonology*.

1. **Caleb Belth,** Alican Büyükçakır, and Danai Koutra. 2022. A Hidden Challenge of Link Prediction: Which Pairs to Check? *Knowledge and Information Systems.* 64(3), 743-771.

Peer-Reviewed Proceedings Articles

10. **Caleb Belth**. 2024. Meaning-Informed Low-Resource Segmentation of Agglutinative Morphology. *Society for Computation in Linguistics.*
9. **Caleb Belth**. 2023. Towards a Learning-Based Account of Underlying Forms: A Case Study in Turkish. *Society for Computation in Linguistics*.
8. **Caleb Belth**, Sarah Payne, Deniz Beser, Jordan Kodner, and Charles Yang. 2021. The Greedy and Recursive Search for Morphological Productivity. *CogSci*.
7. **Caleb Belth,** Alican Büyükçakır, and Danai Koutra. 2020. A Hidden Challenge of Link Prediction: Which Pairs to Check? *IEEE International Conference on Data Mining (ICDM).*
   Selected as one of the best papers at ICDM'20. Invited for publication at the KAIS Journal, Springer.
6. **Caleb Belth,** Xinyi Zheng, Danai Koutra. 2020. Mining Persistent Activity in Continually Evolving Networks. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*.
5. **Caleb Belth,** Xinyi Zheng, Jilles Vreeken, and Danai Koutra. 2020. What is Normal, What is Strange, and What is Missing in a Knowledge Graph: Unified Characterization via Inductive Summarization. *ACM The Web Conference (WWW).*
4. **Caleb Belth**, Xinyi Zheng, Danai Koutra. 2020. Mining Persistent Activity in Continually Evolving Networks. *ACM SIGKDD Workshop on Mining and Learning with Graphs (MLG).*
3. Tara Safavi, **Caleb Belth**, Lukas Faber, Davide Mottin, Emmanuel Muller, and Danai Koutra. 2019. Personalized Knowledge Graph Summarization: From the Cloud to Your Pocket. *IEEE International Conference on Data Mining (ICDM).*
2. **Caleb Belth**, Fahad Kamran, Donna Tjandra, and Danai Koutra. 2019. When to remember where you came from: Node representation learning in higher-order networks. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).*
1. **Caleb Belth**, Fahad Kamran, Donna Tjandra, and Danai Koutra. 2019. When to remember where you came from: Node representation learning in higher-order networks. *ACM SIGKDD Workshop on Mining and Learning with Graphs (MLG).*

Peer-Reviewed Abstracts with Proceedings

2. **Caleb Belth**. 2023. A Learning-Based Account of Non-Productivity in Dutch Voicing Alternations. *Boston University Conference on Language Development.*
1. **Caleb Belth**. 2022. Learning Non-Local Phonological Alternations via Automatic Creation of Tiers. 2022. *Cognitive Modeling and Computational Linguistics workshop at ACL.*

Manuscripts

4. {**Caleb Belth**, Sarah Payne}, Jordan Kodner, and Charles Yang. In prep. *An Adequate Discovery Procedure.*
3. **Caleb Belth**. In prep. *Experimental Evidence that Learning Morphophonological Alternations Starts Local.*
2. **Caleb Belth**. In prep. *A Computational Comparative Study of the Acquisition of Germanic Noun Voicing Alternations: The cases of Dutch, German, and Yiddish.*
1. **Caleb Belth**. In prep. *A Learning-Based Account of the Categories Underlying Stochastic Phonological Knowledge.*

Talks and Posters

(all presented by me or in tandem with coauthor(s) unless otherwise stated)

22. **Caleb Belth**. 2024. *Experimental Evidence that Learning of Morphophonological Alternations Starts Local*. Talk at the 54th Annual Meeting of the North East Linguistics Society (NELS). Jan 27.
21. **Caleb Belth**. 2023. *A Learning-Based Account of Non-Productivity in Dutch Voicing Alternations*. Poster at the 48th Annual Boston University Conference on Language Development. Nov 3.

20. **Caleb Belth**. 2023. *A Learning-Based Account of Non-Productivity in Dutch Voicing Alternations*. Poster accepted, but unable to present due to schedule conflict.
19. **Caleb Belth**. 2023. *Towards a Learning-Based Account of Underlying Forms: A Case Study in Turkish*. Poster at the 6th Annual Meeting of the Society for Computation in Linguistics. Held Jun 16-17.
18. **Caleb Belth**. 2023. *A Learning-Based Account of Phonological Tiers*. Talk at the 47th Annual Penn Linguistics Conference. Held Mar 18-19.
17. **Caleb Belth**. 2023. *Learning Non-Local Phonological Alternations via Automatic Creation of Tiers*. Talk at the 97th Annual Meeting of the Linguistic Society of America. Held Jan 5-8.
16. **Caleb Belth**. 2022. *How a Proclivity for Adjacency can Drive the Learning of Non-Local Alternations.* Talk at MidPhon 27. Oct 8.
15. **Caleb Belth**. 2022. *Learning Non-Local Phonological Alternations via Automatic Creation of Tiers*. Poster at the ACL workshop on Cognitive Modeling and Computational Linguistics. May 22.
14. Sarah Payne, **Caleb Belth,** Jordan Kodner, and Charles Yang. 2022. *Searching for Morphological Productivity.* Talk given at the 96th Annual Meeting of the Linguistic Society of America. Held Jan 6-9.
13. **Caleb Belth**, Sarah Payne, Deniz Beser, Jordan Kodner, and Charles Yang. 2021. *The Greedy and Recursive Search for Morphological Productivity*. Poster at the 43rd Annual Meeting of the Cognitive Science Society. Held Jul 26-29.
12. {**Caleb Belth,** Sarah Payne}, Jordan Kodner, and Charles Yang. 2021. *Searching for Morphological Productivity*. Talk by Sarah Payne at the 46th Annual Boston University Conference on Language Development. Held Nov 4-7.
11. Sarah Payne, **Caleb Belth**, Jordan Kodner, and Charles Yang. 2021. *The Recursive Search for Morphological Productivity*. Poster at the 5th Annual American International Morphological Meeting. Held Aug 26-29.
10. **Caleb Belth,** Alican Büyükçakır, and Danai Koutra. 2020. *A Hidden Challenge of Link Prediction: Which Pairs to Check?*. Talk at the IEEE International Conference on Data Mining (ICDM). Held Nov 17-20.
9. **Caleb Belth,** Xinyi Zheng, Danai Koutra. 2020. *Mining Persistent Activity in Continually Evolving Networks.* Talk at the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). Held Aug 23-27.
8. **Caleb Belth,** Xinyi Zheng, Jilles Vreeken, and Danai Koutra. 2020. *What is Normal, What is Strange, and What is Missing in a Knowledge Graph: Unified Characterization via Inductive Summarization.* Talk at ACM The Web Conference (WWW). Held Apr 20-24.
7. **Caleb Belth**, Xinyi Zheng, Danai Koutra. 2020. *Mining Persistent Activity in Continually Evolving Networks.* Talk at the ACM SIGKDD Workshop on Mining and Learning with Graphs (MLG). Aug 24.
6. Tara Safavi, **Caleb Belth**, Lukas Faber, Davide Mottin, Emmanuel Muller, and Danai Koutra. 2019. *Personalized Knowledge Graph Summarization: From the Cloud to Your Pocket*. Talk by Tara Safavi at the IEEE International Conference on Data Mining (ICDM). Held Nov 8-11.
5. **Caleb Belth**, Fahad Kamran, Donna Tjandra, and Danai Koutra. 2019. *When to remember where you came from: Node representation learning in higher-order networks*. Talk at the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). Held Aug 27-30.
4. **Caleb Belth**, Fahad Kamran, Donna Tjandra, and Danai Koutra. 2019. *When to remember where you came from: Node representation learning in higher-order networks.* Poster at the ACM SIGKDD Workshop on Mining and Learning with Graphs (MLG). Aug 5.
3. **Caleb Belth**, Jilles Vreeken, and Danai Koutra. 2019. *What is Normal, What is Strange, and What is Missing in a Knowledge Graph: Unified Characterization via Inductive Summarization.* Poster at the University of Michigan MIDAS Symposium. Nov 15.
2. **Caleb Belth**, Fahad Kamran, Donna Tjandra, and Danai Koutra. 2019. *When to remember where you came from: Node representation learning in higher-order networks*. Poster at University of Michigan AI Symposium. Oct 19.
1. **Caleb Belth** and Daisuke Kihara. 2017. *Deep Learning for Protein Binding Ligand Prediction*. Poster at the Purdue University Undergraduate Research & Poster Symposium. Apr 11.

Teaching

Independent Study: Computational Approaches to Phonological Learning — Spring 2024
Independent study on computational phonology, focusing on phonological learning

LING 5981/6080: Python for Linguists — Fall 2023
Head Instructor; teaching programming in Python to linguistics students

LING 111: Lenses into Language — Winter/Spring 2023
Graduate Student Instructor for undergraduate, introductory linguistics course

LING 347 / PSYCH 349: Talking Minds — Fall 2022
Graduate Student Instructor for undergraduate course in psycholinguistics

International Summer School on Data Science — 2020
Tutorial Instructor

MIDAS Data Science Summer Camp for High School Students, University of Michigan — 2019
Instructor, week-long summer camp

Outreach and Service

Lead Developer and Maintainer of Python package *algophon* [link] — 2024-Present

Department of Linguistics Undergraduate Committee Member — 2023-Present

Department of Linguistics Search Committee Member — 2023-2024

Diversity Initiative in Graduate Applications, University of Michigan — 2022-2023
Founder, program to connect students from underrepresented communities to UM research groups

M-DICE, City of Detroit, World Economic Forum, The Knight Foundation — 2019-2021
Graduate student lead, project to make access to transportation more equitable

CSEG Wellness, University of Michigan — 2019-2021
Co-founder, organization to improve graduate student wellness

Explore Graduate Studies, University of Michigan — 2019
Volunteer, workshop to broaden participation in computer science graduate programs

Student Mentoring

Mohammed Al-Ariqy — 2023-Present
Dissertation Committee Member

Xueming Xu, Undergraduate, University of Michigan — 2020-2021
Now: Software Engineer, TikTok

Xinyi Zheng, Undergraduate, University of Michigan — 2019-2020
Now: PhD student, Carnegie Mellon University CS

Invited Talks

Stories from An Algorithmic Approach to Phonology, Brigham Young University — 2024

The Interaction Between Learning Algorithms and Formal Language Theory — 2024
LSA Conference Organized Session on Formal Language Theory in Morphology and Phonology

Historical Contingency in Language, College of Humanities, University of Utah — 2023

ThinkB1G: Your Roadmap to Landing a Role at a Startup, Purdue University — 2017

Guest Lectures

In Charles Yang's Graduate Seminar on Discovery Procedures, UPenn — 2023

In Kyle Gorman's Graduate Seminar on Computational Linguistics, CUNY — 2022

Reviewing

NSF Linguistics Program 2024
Ad hoc reviewer

Phonology 2024
Reviewer

Society for Computation in Linguistics (SCiL) 2024
Reviewer

CogSci Conference 2021
Reviewer

ACM The Web Conference (WWW) 2021
Subreviewer

ACM International Conference on Information and Knowledge Management (CIKM) 2020
PC member, posters and demos session

SIAM Workshop on Network Science (NS20) 2020
Subreviewer

ACM The Web Conference (WWW) 2020
Subreviewer

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) 2019
Subreviewer

IEEE International Conference on Data Science and Advanced Analysis (DSAA) 2019
Subreviewer

Grants

Rackham Graduate Student Research Grant 2023
$3,000

Facebook Research Award 2020
$50,000; Based on my research, and contributed to writing.

Industry Experience

Applied Science Intern, Amazon, Seattle, WA (Remote; COVID-19) Summer 2020
Created an approach for discovering product attributes

Software Engineer Intern, Sift, San Francisco, CA Summer 2018
Developed and deployed a gradient tree-boosting algorithm for automated fraud detection

Software Engineer Intern, Handshake, San Francisco, CA Summer 2017
Developed a web platform for university students to find their ideal employers

Software Engineer Intern, Iris, Owosso, MI Summer 2016
Developed Android code to run computer vision inference on mobile

Software Engineer Intern, Iris, Owosso, MI Summer 2015
Developed Android code

Software Development Intern, Enspire Software, Fort Wayne, IN Summer 2014
Developed Android code

Programming Languages in order of Proficiency

Python (expert), Java (proficient), C (proficient), C++ (proficient), Bash (proficient), Scala (familiar)

Professional Membership

Linguistic Society of America (LSA) Member	Present
Association of Computing Machinery (ACM) Student Member 2018	Graduate School
Institute of Electrical and Electronics Engineers (IEEE) Student Member	Graduate School

Other Projects

Machine Learning Text and Network Joint Embeddings, Purdue University 2017-2018
Researched jointly embedding text and social network nodes into the same embedding space

Deep Learning for Protein Binding Ligand Prediction, Purdue University 2015-2018
Used deep learning to predict protein binding ligands for drug design