

POSITIONS

Assistant Professor

August 2025 -

Khoury College of Computer Sciences
Northeastern University

Postdoctoral Researcher

September 2024 - August 2025

Faculty of Computer Science
Universität Wien (University of Vienna)

Visiting Researcher

March 2020 - March 2022

Facebook AI Research (FAIR), Seattle

Research Contractor

July 2018 - March 2020

Facebook AI Research (FAIR), Seattle

EDUCATION

University of Washington, Seattle, WA

PhD in Computer Science

Advisor: Luke Zettlemoyer

September 2017 - June 2024

Columbia University, New York, NY

BA in Computer Science, Cum Laude

Advisor: Kathy McKeown

September 2013 - May 2017

PUBLICATIONS

Varvara Arzt, Allan Hanbury, and **Terra Blevins**. *Left-Branching Transformers Excel at Right-Branching Languages: Data Shapes Word Order Preferences in Language Models*. To appear in Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2026.

Terra Blevins, Stephen Mayhew, Marek Suppa, Hila Gonen, Shachar Mirkin, Vasile Pais, Kaja Dobrovoljc, Voula Giouli, Jun Kevin, Enes Yilandiloğlu, Eugene Jang, Eungseo Kim, Jeongyeon Seo, Xenophon Gialis, and Yuval Pinter. *Universal NER v2: Towards a Massively Multilingual Named Entity Recognition Benchmark*. In Proceedings of the 2026 International Conference on Language Resources and Evaluation (LREC), 2026.

Terra Blevins, Susanne Schmalwieser, Benjamin Roth. *Do language models accommodate their users? A study of linguistic convergence*. In Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2026.

Varvara Arzt, Allan Hanbury, Michael Wiegand, Gábor Recski, **Terra Blevins**. *Relation Extraction or Pattern Matching? Unravelling the Generalisation Limits of Language Models for Biographical RE*. In Proceedings of the International Joint Conference on Natural Language Processing & Asia-Pacific Chapter of the Association for Computational Linguistics (IJCNLP-AACL), 2025.

Loris Schoenegger, Lukas Thoma, **Terra Blevins**, and Benjamin Roth. *Influence-driven Curriculum Learning for Pre-training on Limited Data*. In Proceedings of the 3rd BabyLM Challenge Workshop, 2025.

Alessio Tosolini and **Terra Blevins**. *Analyzing the Linguistic Priors of Language Models with Synthetic Languages*. In Proceedings of the 7th Workshop on Research in Computational Linguistic Typology and Multilingual NLP, 2025.

Hila Gonen, **Terra Blevins**, Alisa Liu, Luke Zettlemoyer, and Noah A. Smith. *Does Liking Yellow Imply Driving a School Bus? Semantic Leakage in Language Models*. In Proceedings of the

2025 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2025.

Terra Blevins, Tomasz Limisiewicz, Suchin Gururangan, Margaret Li, Hila Gonen, Noah A. Smith, and Luke Zettlemoyer. *Breaking the Curse of Multilinguality with Cross-lingual Expert Language Models*. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024.

C.M. Downey, **Terra Blevins**, Dhvani Serai, Dwija Parikh, and Shane Steinert-Threlkeld. *Targeted Multilingual Adaptation for Low-resource Language Families*. In Findings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024.

Tomasz Limisiewicz, **Terra Blevins**, Hila Gonen, Orevaoghene Ahia, and Luke Zettlemoyer. *MYTE: Morphology-Driven Byte Encoding for Better and Fairer Multilingual Language Modeling*. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL), 2024.

Stephen Mayhew, **Terra Blevins**, Shuheng Liu, Marek Šuppa, Hila Gonen, Joseph Marvin Imperial, Börje F. Karlsson, Peiqin Lin, Nikola Ljubešić, LJ Miranda, Barbara Plank, Arij Riabi, Yuval Pinter. *Universal NER: A Gold-Standard Multilingual Named Entity Recognition Benchmark*. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2024.

Akari Asai, Sneha Kudugunta, Xinyan Velocity Yu, **Terra Blevins**, Hila Gonen, Machel Reid, Yulia Tsvetkov, Sebastian Ruder, Hannaneh Hajishirzi. *BUFFET: Benchmarking Large Language Models for Few-shot Cross-lingual Transfer*. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2024.

Haoqiang Kang*, **Terra Blevins***, and Luke Zettlemoyer. *Translate to Disambiguate: Zero-shot Multilingual Word Sense Disambiguation with Pretrained Language Models*. In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2024.

Weijia Shi, Anirudh Ajith, Mengzhou Xia, Yangsibo Huang, Daogao Liu, **Terra Blevins**, Danqi Chen, Luke Zettlemoyer. *Detecting Pretraining Data from Large Language Models*. In the International Conference on Learning Representations (ICLR), 2024.

Hila Gonen, Srini Iyer, **Terra Blevins**, Noah A. Smith, and Luke Zettlemoyer. *Demystifying Prompts in Language Models via Perplexity Estimation*. In Findings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.

C.M. Downey, **Terra Blevins**, Nora Goldfine, and Shane Steinert-Threlkeld. *Embedding structure matters: Comparing methods to adapt multilingual vocabularies to new languages*. In Proceedings of the 3rd Multilingual Representation Learning (MRL) Workshop, 2023. **Awarded Best Paper**.

Terra Blevins, Hila Gonen, and Luke Zettlemoyer. *Prompting Language Models for Linguistic Structure*. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL), 2023.

Terra Blevins, Hila Gonen, and Luke Zettlemoyer. *Analyzing the Mono- and Cross-Lingual Pretraining Dynamics of Multilingual Language Models*. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022.

Terra Blevins and Luke Zettlemoyer. *Language Contamination Helps Explain the Cross-lingual Capabilities of English Pretrained Models*. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022.

Terra Blevins, Mandar Joshi, and Luke Zettlemoyer. *FEWS: Large-Scale, Low-Shot Word Sense Disambiguation with the Dictionary*. In Proceedings of the 16th Conference of the European Chapter

of the Association for Computational Linguistics (EACL), 2021.

Terra Blevins and Luke Zettlemoyer. *Moving Down the Long Tail of Word Sense Disambiguation with Gloss Informed Bi-encoders*. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020.

Terra Blevins and Luke Zettlemoyer. *Better Character Language Modeling Through Morphology*. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019.

Terra Blevins, Omer Levy, and Luke Zettlemoyer. *Deep RNNs Encode Soft Hierarchical Syntax*. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL), 2018.

Terra Blevins, Robert Kwiatkowski, Jamie C MacBeth, Kathleen McKeown, Desmond Patton and Owen Rambow, *Automatically Processing Tweets from Gang-Involved Youth: Towards Detecting Loss and Aggression*. In Proceedings of the International Conference on Computational Linguistics (COLING), 2016.

Or Biran, **Terra Blevins** and Kathleen McKeown, *Mining Paraphrasal Typed Templates from a Plain Text Corpus*. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL), 2016.

PREPRINTS

Sanchit Ahuja and **Terra Blevins**. *Parameter Alignment Mitigates Catastrophic Forgetting in Multilingual Expert Language Models*. Preprint, 2026.

Terra Blevins. *Accommodation Goes Both Ways: Studying Linguistic Convergence Between Humans and Language Models*. Preprint, 2026.

Yuxi Xia, Dennis Ulmer, **Terra Blevins**, Yihong Liu, Hinrich Schütze, and Benjamin Roth. *Calibration Is Not Enough: Evaluating Confidence Estimation Under Language Variations*. Preprint, 2026.

Tyler A Chang, Catherine Arnett, ..., **Terra Blevins**, ..., Baber Abbasi, Stella Biderman, David Stap, Duygu Ataman, Fabian Schmidt, Hila Gonen, Jiayi Wang, and David Ifeoluwa Adelani. *Global PIQA: Evaluating Commonsense Reasoning Across 100+ Languages and Cultures*. Preprint, 2025.

Haoqiang Kang, **Terra Blevins**, and Luke Zettlemoyer. *Comparing Hallucination Detection Metrics for Multilingual Generation*. Preprint, 2024.

Mandar Joshi, **Terra Blevins**, Mike Lewis, Daniel S. Weld, and Luke Zettlemoyer. *Few-shot Mining of Naturally Occurring Inputs and Outputs*. Preprint, 2022.

FELLOWSHIPS AND AWARDS

EECS Rising Stars 2024
Selected as a participant in the EECS Rising Stars Workshop, which is “an intensive workshop for graduate students and postdocs with historically marginalized or underrepresented genders who are interested in pursuing academic careers in electrical engineering and computer science, artificial intelligence, and decision-making.”

NSF Graduate Research Fellowship 2018-2023
Fellowship awarded by the National Science Foundation (NSF) to “recruit and support individuals who demonstrate the potential to make significant contributions in STEM.”

College of Engineering Dean’s Fellowship 2017-2018
Fellowship awarded to select first-year graduate students within UW’s College of Engineering.

CRA Outstanding Undergraduate Researchers Award Finalist 2017
 Awarded by Computing Research Association to undergraduate students of North American institutions for outstanding research potential.

Theodore R. Bashkow Award 2017
 A prize presented to a computer science senior who has excelled in independent projects. This is awarded by Columbia University in honor of Professor Theodore R. Bashkow.

INVITED TALKS

Breaking the Curse of Multilinguality in Language Models
 HumanCLAIM Workshop, University of Göttingen, Germany 2026
 AI Seminar Series, University of Vienna, Austria 2024
 UFAL Seminar Series, Charles University, Czech Republic 2024
 Department of Computer Science, George Washington University 2024
 Computer Science and Engineering, Washington University in St. Louis 2024
 Department of Computer Science, Columbia University 2024
 Department of Computer Science, University of California, Irvine 2024
 Khoury College of Computer Sciences, Northeastern University 2024

Moving Down the Long Tail of Word Sense Disambiguation with Gloss Informed Bi-encoders
 West Coast NLP Summit (WeCNLP, Poster) 2020

Better Character Language Modeling Through Morphology
 UW CSE Women's Research Day (Poster) 2020
 UW CSE Industry Affiliates Day 2019

Deep RNNs Encode Soft Hierarchical Syntax
 GenDeep Workshop (Poster) 2018

TEACHING

CS 4120 Natural Language Processing Northeastern University
 Spring 2026, Fall 2026

CS 7170 Seminar in Artificial Intelligence Northeastern University
 Multilingual Language Modeling Fall 2025

Guest Lecturer
 Frontiers in Multilingual AI, University of Maryland, College Park 2026
 Large Language Model Seminar, Washington University in St. Louis 2024
 NLP Seminar, University of Toronto; UTech Ontario 2023

Teaching Assistant Computer Science and Engineering
 Winter 2020, Winter 2021 University of Washington
 TA for Natural Language Processing (CSEP 517 and CSE 447/517); held office hours, graded assignments, and developed homeworks for the class.

Teaching Assistant Computer Science Dept.
 Spring 2016 - Spring 2017 Columbia University
 TA'ed Introduction to Database Systems (cs4111; Spring 2016, Summer 2016, Spring 2017) and Advanced Database Systems (cs6111; Autumn 2016).

PROFESSIONAL SERVICE

ACL Workshop Officer 2025-2027

Conference Organizing Committees
 ACL Workshop Co-Chair 2025

Conference Program Committees

ICLR <u>Area Chair</u>	2025
ARR <u>Action Editor</u>	2024-
ACL Rolling Review (ARR)	2021-2024
EMNLP	2020-2023
ACL	2020-2021, 2023
COLING	2020, 2022
NAACL	2021
EACL	2021

Workshop Program Committees

ACL Student Research Workshop	2023
Workshop on Noisy User-generated Text	2022
Methods for Optimizing and Evaluating Neural Language Generation	2019

Journal Reviewing

Computational Linguistics (CL)	2021
--------------------------------	------

ADVISING

Sanchit Ahuja	September 2025 -
Northeastern Ph.D	
Varvara Arzt	October 2024 -
Technische Universität Wien Ph.D	
Eric Bennett	September 2026 -
Northeastern Ph.D	
Eugene Jang	September 2025 -
Northeastern Ph.D	

Marlene Albrecht	October 2024 - December 2025
Universität Wien MA	
Ksenia Dvorkina	October 2024 - September 2025
Universität Wien MA	
Harold Ngoupeyou	April 2025 - August 2025
École polytechnique BS	
Alessio Tosolini	March 2023 - June 2024
University of Washington CSE/Linguistics BS → McGill University PhD	
Haoqiang (Murray) Kang	September 2022 - June 2024
University of Washington CSE BS → UC San Diego PhD	
Shuheng Liu	January 2020 - June 2021
University of Washington CSE BS → GA Tech PhD	
Kaushal Mangipudi	January 2020 - March 2021
University of Washington CSE MS → Facebook ML Software Engineer	