

William Sheffield

sheffieldw@utexas.edu | [GitHub](#) | [webpage](#)

EDUCATION

- Ph.D. *Computational Linguistics*, **The University of Texas at Austin**, August 2022-Expected Fall 2026
Advisor: Jessie Li
Committee Members: David Beaver, Kanishka Misra, Alex Warstadt
- B.S. *Mathematics*, with Honors, **The University of Texas at Austin**, August 2019-May 2022
Minor in Russian Language
- B.A. *Linguistics*, with High Honors, **The University of Texas at Austin**, August 2019-May 2022
Elements of Computing Certificate

PEER-REVIEWED PUBLICATIONS

Referred Conference + Journal Articles

1. Daniel Brubaker, **William Sheffield**, Junyi Jessie Li, and Kanishka Misra. 2026. [Wugnectives: Novel Entity Inferences of Language Models from Discourse Connectives](#). In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6109–6127, Rabat, Morocco. Association for Computational Linguistics.
2. **William Sheffield**, Kanishka Misra, Valentina Pyatkin, Ashwini Deo, Kyle Mahowald, and Junyi Jessie Li. 2025. [Is It JUST Semantics? A Case Study of Discourse Particle Understanding in LLMs](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 21704–21715, Vienna, Austria. Association for Computational Linguistics.
3. Yating Wu*, **William Sheffield***, Kyle Mahowald, Junyi Jessie Li. 2023. [Elaborative Simplification as Implicit Questions Under Discussion](#). In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
4. Ashwin Devaraj, **William Sheffield**, Byron C. Wallace, and Junyi Jessie Li. 2022. [Evaluating Factuality in Text Simplification](#). In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. **Outstanding Paper Award**. pp. 7331–7345.

Peer-reviewed Workshop Articles

5. Junyi Jessie Li, Yang Janet Liu, Kanishka Misra, Valentina Pyatkin, and **William Sheffield**. 2026. [Which course? Discourse! Teaching Discourse and Generation in the Era of LLMs](#). In *Proceedings of the Seventh Workshop on Teaching Natural Language Processing (TeachNLP 2026)*, pages 45–58, Rabat, Morocco. Association for Computational Linguistics.

*Denotes equal contribution

RESEARCH APPOINTMENTS

Graduate Research Assistant, Department of Linguistics - UT Austin

August 2024 – June 2025

Investigating the naturalness gap between human-centered language and existing robot instruction datasets, with a focus on implicit information:

- To establish a human-centered benchmark, created 32 high-quality rewritten versions of existing robot instruction dataset entries
- Designed and deployed a crowdsourcing interface to collect implicitness ratings across task-completion conditions
- Applied Question Under Discussion theory to systematically analyze implicit information structures in robot instructions

Advised by Dr. Jessie Li (Linguistics), Dr. Joydeep Biswas (Computer Science), and Dr. Amy Zhang (Electrical and Computer Engineering), UT Austin

Graduate Research Assistant, Department of Linguistics - UT Austin

June 2023 – August 2023

Contributing to "Is It JUST Semantics? A Case Study of Discourse Particle Understanding in LLMs" [2]:

- Built a novel dataset of 90 sentences targeting fine-grained sense distinctions of the English discourse particle "just" to expose gaps in LLM semantic understanding
- Designed two evaluation methodologies to probe model performance on these distinctions, then implemented and benchmarked them across multiple open-source LLMs (Llama 3, 3.2, 3.3; Gemini 2; Mistral-v0.3; Olmo, Olmo2), revealing only partial understanding of senses in the larger models

Undergraduate Research Assistant, Department of Linguistics - UT Austin

June 2021 – May 2022

Contributing to "Elaborative Simplification as Implicit Questions Under Discussion" [3] and "Evaluating Factuality in Text Simplification" [4]:

- Engineered a Python/boto3 interface for Amazon Mechanical Turk to crowdsource discourse annotation data, enabling scalable data collection for the elaborative simplification project
- Annotated 200 elaboration samples for discourse relations and authored qualitative analysis sections for both papers
- Co-developed and stress-tested an annotation scheme for crowdsourcing factuality error measures, then analyzed error trends across 500+ data points in model outputs and datasets

HONORS AND AWARDS

- | | |
|---|---------------------------|
| ● Outstanding Paper Award | ACL 2022 |
| ● UT Austin University Honors | Fall 2019 – Spring 2022 |
| ● UT Austin College of Liberal Arts Dean’s Honor List | Spring 2020 – Spring 2022 |
| ● Eagle Scout | September 2014 |

SERVICE

- Reviewer for:**
- | | |
|---|--------------------------|
| ● ACL Rolling Review (ARR) | October 2025, March 2026 |
| ● CogSci2026 | March 2026 |
| ● Linguistic Annotation Workshop (LAW) XX | April 2026 |

- Committee Positions as a Graduate Student:**
- | | |
|---|-------------------------|
| ● Associated Students of Linguistics Executive Committee Co-chair | Fall 2024 – present |
| ● UT Austin Linguistics Department Curriculum Committee | Fall 2023 – Spring 2024 |
| ● Associated Students of Linguistics Outreach Committee Co-chair | Fall 2023 – Spring 2024 |
| ● Texas Linguistics Society 2023 Organizing Committee Member | Fall 2022 – Spring 2023 |

SKILLS

Coding Languages: Advanced R, Python and PyTorch, Transformers packages; Intermediate Bash.

Natural Languages: Intermediate Spanish; Basic Russian and Mandarin Chinese; Advanced written Latin; Basic Hawaiian.

TEACHING

- Teaching Assistant**
- UT Austin - Linguistics
- | | | |
|---------------|--|-------------|
| UGS303 | <i>Ethics of Artificial Intelligence: Data & Language</i> | Spring 2026 |
| LIN353D/CS378 | <i>Computational Discourse and Natural Language Generation</i> | Fall 2025 |
| LIN365 | <i>Bias</i> | Spring 2024 |
| LIN313 | <i>Language and Computers</i> | Fall 2023 |
| UGS303 | <i>Ethics of Artificial Intelligence: Data & Language</i> | Spring 2023 |
| LIN353C | <i>Introduction to Computational Linguistics</i> | Fall 2022 |

UNDERGRADUATE WORK EXPERIENCE

Tutor, Sanger Learning Center - UT Austin

August 2021 - May 2022

Provide free, 1-hr appointment Tutoring for UT Austin undergraduates in:

- CS303E: Elements of Computing and Programming (Python),
- CS313E: Elements of Software Design (Python),
- M362K: Probability I,
- M427J: Advanced Calculus for Applications (Vector Calculus),
- M325K: Discrete Mathematics (Intro to Proof)

PRIOR RESEARCH PROJECTS

- **Bioinformatics Research (BIO321G):**

January 2020 - December 2020

- Compared genetic mutation detection tools on blood cancer Next-Gen RNA sequencing data by using PCA and clustering analysis.
- Established a standardized pipeline for benchmarking genetic mutation detectors on RNA data.
- Skills used: Python, Bash, R, TACC Supercomputing Server.
- *Supervised by Dhivya Arasappan, Center for Biomedical Research Support, UT Austin*