

Michael Ginn

www.michaelginn.com
michael.ginn@colorado.edu

University of Colorado
Department of Computer Science
Institute of Cognitive Science

Education

B.S. Computer Science, Washington University in St. Louis, 2022.

M.A. Linguistics¹, University of Colorado, 2024.

Ph.D. Computer Science, University of Colorado, 2022-present. Supervised by Profs. Alexis Palmer and Mans Hulden.

Simultaneous with Combined Ph.D., Institute of Cognitive Science.

Awards and Honors

Best Paper Award, TeachingNLP Workshop @ ACL, 2024.

Best Published Work, Department of Computer Science Research Expo, University of Colorado. 2025.

Ralph Slutz Excellence Award, University of Colorado. 2024.

Valedictorian, Washington University. 2022.

Summa Cum Laude, Washington University. 2022.

Dean's List, Washington University. 2018-2022.

Antoinette Frances Dames Award for Productive Scholarship, Washington University. 2020.

Engineering Outstanding Junior Award, Washington University. 2020.

Tau Beta Pi Honors Society. 2020-2022.

Apple Worldwide Developer Conference Scholarship. 2015, 2019.

James M. McKelvey Undergraduate Research Award, Washington University. 2018.

Alexander S. Langsdorf Fellowship, Washington University. 2018.

Publications

* denotes shared first authorship

Massively Multilingual Joint Segmentation and Glossing

Michael Ginn, Lindia Tjuatja, Enora Rice, Ali Marashian, Maria Valentini, Jasmine Xu, Graham Neubig, and Alexis Palmer. 2026. Forthcoming at ACL 2026. [\[Link\]](#)

¹Non-terminal degree obtained upon completing Ph.D. comprehensive exam.

Neural Induction of Finite-State Transducers

Michael Ginn, Alexis Palmer, and Mans Hulden. 2026. Forthcoming at ACL 2026 Findings. [\[Link\]](#)

Is linguistically-motivated data augmentation worth it?

Ray Groshan*, Michael Ginn*, and Alexis Palmer. Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics. 2025. [\[Link\]](#)

LLM Dependency Parsing with Symbolic Rules

Michael Ginn and Alexis Palmer. Proceedings of The 1st Joint Workshop on Large Language Models and Structure Modeling. 2025. [\[Link\]](#)

Can we teach language models to gloss endangered languages?

Michael Ginn, Mans Hulden, and Alexis Palmer. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024. [\[Link\]](#)

GlossLM: A Massively Multilingual Corpus and Pretrained Model for Interlinear Glossed Text

Michael Ginn*, Lindia Tjuatja*, Taiqi He, Enora Rice, Graham Neubig, Alexis Palmer, and Lori Levin. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024. [\[Link\]](#)

BELT: Building Endangered Language Technology

Michael Ginn, David Ricardo Saavedra Beltran, Camilo Alberto Robayo Romero, and Alexis Palmer. Proceedings of the Fifth Workshop on Teaching NLP. 2024. [\[Link\]](#)

PyFoma: a Python finite-state compiler module

Mans Hulden, Michael Ginn, Miikka Silfverberg, and Michael Hammond. Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics. 2024. [\[Link\]](#)

Resisting the Lure of the Skyline: Grounding Practices in Active Learning for Morphological Inflection

Saliha Muradoğlu, Michael Ginn, Miikka Silfverberg, and Mans Hulden. Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2024. [\[Link\]](#)

On the Robustness of Neural Models for Full Sentence Transformation

Michael Ginn, Ali Marashian, Bhargav Shandilya, Claire Benet Post, Enora Rice, Juan Vásquez, Marie McGregor, Matt Buchholz, Mans Hulden, and Alexis Palmer. Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP). 2024. [\[Link\]](#)

Decomposing Fusional Morphemes with Vector Embeddings

Michael Ginn and Alexis Palmer. Proceedings of the 21st SIGMORPHON workshop on Computational Research in Phonetics, Phonology, and Morphology. 2024. [\[Link\]](#)

Robust Generalization Strategies for Morpheme Glossing in an Endangered Language Documentation Context

Michael Ginn and Alexis Palmer. Proceedings of the 1st GenBench Workshop on (Benchmarking) Generalisation in NLP. 2023. [\[Link\]](#)

Findings of the SIGMORPHON 2023 Shared Task on Interlinear Glossing

Michael Ginn, Sarah Moeller, Alexis Palmer, Anna Stacey, Garrett Nicolai, Mans Hulden, and Miikka Silfverberg. Special Interest Group on Computational Morphology and Phonology Workshop. 2023. [\[Link\]](#)

Employment

Research Assistant, Amazon Alexa AI (*Dr. Alessandro Moschitti*). 2025-present.

Developed neurosymbolic synthetic data generation method for information-seeking dialogs using directed graphs. Designed evaluation procedure motivated by Gricean discourse theory.

Research Assistant, University of Colorado (*Prof. Alexis Palmer*). 2022-present.

Worked under NSF grant FOLTA: From One Language to Another. Researched NLP for rare languages, multilinguality, and large language models. Wrote grant applications for Microsoft Research Fellowship and Schmidt Science Humanities and AI grant.

Research Assistant, University of Colorado (*Prof. Mans Hulden*). 2022-present.

Researched automata theory and computational morphology. Maintained PyFoma open-source library for automata manipulation.

Machine Learning Intern, Apple (*Developer Tools AI*). 2025.

Trained 3b parameter LLM for code edit prediction on 10M synthetic data examples, achieving > 95% accuracy on real coding tasks. Built Xcode prototype leveraging speculative decoding with < 1s response time.

Teaching Assistant, University of Colorado. 2024-2025.

TA for CSCI 2270: Data Structures and LING 3100: Language Sound Structures

Machine Learning Intern, Apple (*Input Experience*). 2024.

Finetuned 30B MM1 multimodal LLM for image-grounded story generation. Implemented multimodal training in internal LLM framework.

Machine Learning Student Researcher, xMentium. 2023-2024.

Machine Learning Engineering Intern, Apple (*iCloud*). 2023.

Built early prototypes of Apple Intelligence features for LLM-based email summary and composition.

Language Data Consortium Liaison, University of Colorado. 2022-2024.

Software Engineering Intern, Apple (*Localization Engineering*). 2021, 2022.

Research Assistant, Washington University (*Prof. Kristin Van Engen*). 2021-2022.

Teaching Assistant, Washington University. 2018-2022.

Data Science Intern, Splice Machine. 2020.

Workshops

BELT: Building Endangered Language Technology. *AMAZONICAS* in Bogota, Colombia, 2023.

Led four-day workshop with Amazonian linguists where participants created simple language tools such as spellcheckers and language learning games.

Talks

NLP for Endangered Languages

Google Machine Learning Group (Boulder). 2025. Oral presentation

Advances and Challenges in Automated Interlinear Glossing

ACL Special Interest Group on Endangered Languages (SIGEL) Speaker Series. 2024. Oral presentation

Toward Robust and Human-Like IGT Generation Models for Real-World Usage

Open Text Collections Consumer Workshop. 2023. Oral presentation

Death, Morality, and Verb Construction in *Odes 2.3*

The Classical Organization of the Middle West and South. 2021. Poster presentation. [\[Link\]](#)

Service

Research mentor for five M.S. students

Reviewer for ACL ARR, LREC-COLING, SIGMORPHON, ComputEL, ACL Student Research Workshop

EMNLP 2023 Student Volunteer

Last updated: April 9, 2026