

# Sarah Moeller

University of Florida  
4131 Turlington Hall P.O. Box 115454  
Gainesville, FL 32611-5454 USA

Email: [smoeller@ufl.edu](mailto:smoeller@ufl.edu)  
WEB: [sarahmoeller.github.io](https://sarahmoeller.github.io)

## Education

- 2021 PhD Linguistics and Cognitive Science, *University of Colorado Boulder*
- 2020 MA Linguistics, *University of Colorado Boulder*
- 2010 MA Applied Linguistics, *Dallas International University*
- 2002 BA History, *Thomas Edison State College*

## Certifications and Professional Development

- 2021 Certificate in College Teaching, University of Colorado Boulder
- 2018 Lisbon Machine Learning School (LxMLS)
- 2002 graduate work in Russian Linguistics, Lomonosov Moscow State University

## Academic Appointments

- 2021-present Assistant Professor, *Department of Linguistics, University of Florida*
- 2020 Instructor, *Department of Linguistics, University of Colorado Boulder*
- 2017-2021 Research Assistant, *Computational Language and Education Research (CLEAR), University of Colorado Boulder*
- 2016-2019 Graduate Teaching Assistant, *Department of Linguistics, University of Colorado Boulder*
- 2011-2016 Associate Instructor, *Department of Applied Linguistics, Dallas International University*
- 2009-10 Graduate Teaching Assistant, *Dept. of Applied Linguistics, Dept. of Applied Anthropology, Dallas International University*

## Industry Positions

- 2018 Project Manager Intern, Language Software Development (LSDev), SIL International
- 2011-2015 Linguistics Knowledge Engineer, Language Computer Corporation (LCC)
- 2002-2008 English as a Foreign Language Teacher & Russian-English Interpreter, Self-employed

## Grants, Honors & Awards

- 2024-2026 Principal Investigator, "Reanimating African American Oral Histories of the Gulf South: Enhancing Public Engagement", American Council of Learned Societies. \$99,956
- 2025 Publishing in AI Teaching and Learning Award. AI<sup>2</sup> Center, University of Florida
- 2025 Applicant, Sponsorship of ComputEL-8, Google LLC. \$5,000

2021-2025	Principal Investigator, “Reanimating African American Oral Histories of the Gulf South”, National Endowment for Humanities (NEH), Humanities Collections and Reference Resources (PW-277433-21). \$349,990
2024	Applicant, Unrestricted gift for conference support, Google LLC. \$10,000
2020	Principal Investigator, Individual Research Grant, International Literacy and Development (ILAD). \$20,000
2019	Graduate Research Grant. Center for Advancement of Teaching in Social Sciences (CARTSS), University of Colorado.
2019	Supplemental Award for Undergraduate Research Experience. Center for Advancement of Teaching in Social Sciences (CARTSS), University of Colorado.
2017 & 2019	Summer research grant. Department of Linguistics, University of Colorado Boulder.
2016	Department Fellowship. Department of Linguistics, University of Colorado Boulder.

## Publications

### MANUSCRIPTS

in progress	Sarah Moeller and Zoey Liu. Optimal Minimal Data Amounts for Low Resource Language Annotation. Sarah Moeller. Causes and Costs of the Annotation Bottleneck: Report on IGT Creation Survey. Marcus Sweeting, James Cuenod, Wayne Joubert, <b>Sarah Moeller</b> . Morphological Segmentation of Low-Resource Languages with Large Language Models.
under review	Sarah Moeller, Godfred Agyapong, Jarrod Cruz, Mans Hulden, Alexis Palmer. Computational Tools for Language Documentation and Description. <i>Annual Reviews in Linguistics</i> .
in revision	Sarah Moeller. Post-verbal subjects in Lezgi. <i>Rodnoy Yazyk</i> . Matt Simmons, Anna Hamilton, Alexis Davis, William Dyer, Christina Jean Philippe, Jessica Heady, Pryce Houck, Adolfo Romero, Rebekah Anne Cordova, Deborah Hendrix, Paul Ortiz, <b>Sarah Moeller</b> , Kevin Tang. A Meeting of Two Worlds, Oral History and Linguistics: Partnerships, Perplexities, and Potentialities in Researching African American Language. <i>The Oral History Review</i> .
to appear	Kriukova, Olga, Katherine Schmirler, Sarah Moeller, Olga Lovick, Inge Genée, Antti Arppe and Alexandra Smith. AI for Interlinearization and POS-tagging: Teaching Linguists to Fish. <i>ComputEL-8</i> . 2025.

### ARTICLES

2025	Rahul Porwal, Alice Rozet, Jotsna Gowda, Pryce Houck, Kevin Tang, and Sarah Moeller. Analysis of LLM as a grammatical feature tagger for African American English. In Luis Chiruzzo, Alan Ritter, and Lu Wang, editors, <i>Findings of the Association for Computational Linguistics: NAACL 2025</i> , pages 7744–7756, Albuquerque, New Mexico, April 2025. Association for Computational Linguistics
2024	Wilermine Previlon, Alice Rozet, Jotsna Gowda, Bill Dyer, Kevin Tang, and Sarah Moeller. <a href="#">Leveraging Syntactic Dependencies in Disambiguation: The Case of African American English</a> . In <i>Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)</i> , pages 10403–10415, Torino, Italia, May 2024. ELRA and ICCL Sarah Moeller and Antti Arppe. Machine-in-the-loop with documentary and descriptive linguists. In <i>Proceedings of the Seventh Workshop on the Use of Computational Methods in the</i>

- 2023 *Study of Endangered Languages*, pages 27–32, St. Julians, Malta, March 2024. Association for Computational Linguistics
- Michael Ginn, Sarah Moeller, Alexis Palmer, Anna Stacey, Garrett Nicolai, Mans Hulden, and Mikka Silfverberg. [Findings of the SIGMORPHON 2023 Shared Task on Interlinear Glossing](#). In *Proceedings of the 20th SIGMORPHON workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 186–201, Toronto, Canada, July 2023. Association for Computational Linguistics
- Shengyu Liao, Sarah Moeller, and Beth Bryson. [Morphological Data Generation from FLEx](#). In *Proceedings of the Sixth Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 51–59, Remote, March 2023. Association for Computational Linguistics
- 2022 Harrison Santiago, Joshua Martin, Sarah Moeller, and Kevin Tang. [Disambiguation of morpho-syntactic features of African American English – the case of habitual be](#). In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 70–75, Dublin, Ireland, May 2022. Association for Computational Linguistics
- 2021 Jens EL Van Gysel, Meagan Vigus, Jayeol Chun, Kenneth Lai, Sarah Moeller, Jiarui Yao, Tim O’Gorman, Andrew Cowell, William Croft, and Chu-Ren Huang. [Designing a uniform meaning representation for natural language processing](#). *KI-Künstliche Intelligenz*, 35(3):343–360, 2021. Publisher: Springer
- Sarah Moeller and Mans Hulden. [Integrating automated segmentation and glossing into language documentation and description](#). In *Proceedings of the Fourth Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL-4)*, Honolulu, HI, 2021
- Sarah Moeller. [Computational Morphology for Language Documentation and Description](#). *Colorado Research in Linguistics*, 25, 2021
- Sarah Moeller, Ling Liu, and Mans Hulden. [To POS tag or not to POS tag: The impact of POS tags on morphological learning in low-resource settings](#). In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 966–978, 2021
- 2020 Sarah Moeller, Ling Liu, Changbing Yang, Katharina Kann, and Mans Hulden. [IGT2P: From Interlinear Glossed Texts to Paradigms](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5251–5262, 2020
- Graham Neubig, ... Sarah Moeller ..., and Patrick Littell. [A Summary of the First Workshop on Language Technology for Language Documentation and Revitalization](#). In *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)*, pages 342–351, Marseille, France, 2020
- Sarah Moeller, Irina Wagner, Martha Palmer, Kathryn Conger, and Skatje Myers. [The Russian PropBank](#). In *Proceedings of the 12th Language Resources and Evaluation Conference (LREC)*, pages 5997–6004, Marseille, France, 2020
- 2019 Sarah R. Moeller. [Customizing the LinGO Grammar Matrix Morphology](#). *Colorado Research in Linguistics (CRiL)*, 24, 2019
- Kevin Stowe, Sarah Moeller, Laura Michaelis, and Martha Palmer. [Linguistic Analysis Improves Neural Metaphor Detection](#). In *Proceedings of the 23rd conference on computational natural language learning (CoNLL)*, pages 362–371, Hong Kong, China, 2019
- Sarah Moeller, Ghazaleh Kazeminejad, Andrew Cowell, and Mans Hulden. [Improving Low-Resource Morphological Learning with Intermediate Forms from Finite State Transduc-](#)

- 2018 ers. In *Proceedings of the Workshop on Computational Methods for Endangered Languages (ComputEL-3)*, volume 1, Honolulu, HI, 2019
- Sarah Moeller, Ghazaleh Kazeminejad, Andrew Cowell, and Mans Hulden. [A Neural Morphological Analyzer for Arapaho Verbs Learned from a Finite State Transducer](#). In *Proceedings of the Workshop on Computational Modeling of Polysynthetic Languages*, pages 12–20, Santa Fe, New Mexico, USA, 2018
- Sarah Moeller and Mans Hulden. [Automatic Glossing in a Low-Resource Setting for Language Documentation](#). In *Proceedings of the Workshop on Computational Modeling of Polysynthetic Languages*, pages 84–93, Santa Fe, New Mexico, USA, 2018
- Nathan Schneider, Jena D. Hwang, ... Sarah R. Moeller ..., and Omri Abend. [Comprehensive Supersense Disambiguation of English Prepositions and Possessives](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (ACL)*, pages 185–196, Melbourne, Australia, 2018
- 2011 Sarah Ruth Moeller. [Word Stress in Lezgian in Optimality Theory](#). *Occasional Papers in Applied Linguistics (OPAL)*, 10:1–9, 2011
- Mekoulnodji Ndjerareou, Christy Melick, and Sarah Moeller. [A Brief Grammatical Sketch of Ngambay](#). *GIALens: Electronic Notes Series*, 4(2), 2011

#### SYSTEM DESCRIPTIONS

- 2024 Firoz Ahmed, Nitin Venkateswaran, and Sarah Moeller. The Bangla/Bengali seed dataset submission to the WMT24 open language data initiative shared task. In Barry Haddow, Tom Kocmi, Philipp Koehn, and Christof Monz, editors, *Proceedings of the Ninth Conference on Machine Translation*, pages 556–566, Miami, Florida, USA, November 2024. Association for Computational Linguistics
- Jim Su, Justin Ho, George Broadwell, Sarah Moeller, and Bonnie Dorr. A comparison of fine-tuning and in-context learning for clause-level morphosyntactic alternation. In *Proceedings of the 4th Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP 2024)*, pages 179–187, 01 2024
- 2018 Ling Liu, Ilamvazhuthy Subbiah, Adam Wiemerslage, Jonathan Lilley, and Sarah Moeller. [Morphological Reinflection in Context: CU Boulder’s Submission to CoNLL-SIGMORPHON 2018 Shared Task](#). In *Proceedings of the CoNLL SIGMORPHON 2018 Shared Task: Universal Morphological Reinflection*, pages 86–92, Brussels, 2018

#### BOOKS

- 2016 Brenda H. Boerger, Sarah Ruth Moeller, Will Reiman, and Stephen Self. *Language and Culture Documentation Manual*. Leanpub, June 2016

#### EDITED WORKS

- to appear Jordan Lachler, Godfred Agyapong, Antti Arppe, Aditi Chaudhary, Sarah Moeller, Shruti Rijhwani, Daisy Rosenblum (editors). *Proceedings of the Eighth Workshop on the Use of Computational Methods in the Study of Endangered Languages*, Honolulu, Hawaii, March 2025. Association for Computational Linguistics.
- 2024 Sarah Moeller, Godfred Agyapong, Antti Arppe, Aditi Chaudhary, Shruti Rijhwani, Christopher Cox, Ryan Henke, Alexis Palmer, Daisy Rosenblum, and Lane Schwartz, editors. *Proceedings*

- 2023 of the *Seventh Workshop on the Use of Computational Methods in the Study of Endangered Languages*, St. Julians, Malta, March 2024. Association for Computational Linguistics
- 2022 Atticus Harrigan, Aditi Chaudhary, Shruti Rijhwani, Sarah Moeller, Antti Arppe, Alexis Palmer, Ryan Henke, and Daisy Rosenblum, editors. *Proceedings of the Sixth Workshop on the Use of Computational Methods in the Study of Endangered Languages*. Association for Computational Linguistics, Online, March 2023
- 2021 Sarah Moeller, Antonios Anastasopoulos, Antti Arppe, Aditi Chaudhary, Atticus Harrigan, Josh Holden, Jordan Lachler, Alexis Palmer, Shruti Rijhwani, and Lane Schwartz, editors. *Proceedings of the Fifth Workshop on the Use of Computational Methods in the Study of Endangered Languages*. Association for Computational Linguistics, Dublin, Ireland, May 2022
- Antti Arppe, Jeff Good, Atticus Harrigan, Mans Hulden, Jordan Lachler, Sarah Moeller, Alexis Palmer, Miikka Silfverberg, and Lane Schwartz, editors. *Proceedings of the 4th Workshop on the Use of Computational Methods in the Study of Endangered Languages Volume 1 (Papers)*. Association for Computational Linguistics, Online, March 2021

## MISCELLANEOUS

- 2014 Sarah Ruth Moeller. [Review of SayMore, a tool for Language Documentation Productivity](#). *Language Documentation & Conservation*, 8:66–74, 2014

## Presentations

### INVITED TALKS

- 2024 “Large Language Models in a World of ‘Small’ Languages” in *The Promises and Pitfalls of Large Language Models for Science and Society* Scientific Session, American Association for the Advancement of Science (AAAS) Annual Meeting, February 15-17, 2024, Denver, Colorado.
- “Getting Started with Computational Linguistics”, Special Interest Group on Endangered Languages (SIGEL) Speaker Series, February 6, 2024 (remote).
- 2023 “AI and Language Documentation”, AMAZONICAS IX, Universidad de los Andes, Bogotá, Colombia (remote).
- Panelist, “Big Bot on Campus”, University of Florida College of Liberal Arts & Sciences Public Forum, February 7, 2023
- 2022 Panelist, “Women in High Power Computing (HPC) & AI Panel and Luncheon”, University of Florida, December 2, 2022
- “[Sharing the experiences of African American elders](#).” From Florida, episode 15, 2022 April 13. Brown, Nicci (host/producer), Brooke Adams (producer), James L. Sullivan (producer), Paul Ortiz (contributor), Sarah Moeller (contributor).
- 2021 “The Future of Interlinear Glossed Texts in Computational and Descriptive Linguistics”, Linguistics Colloquium, Australian National University, March 2021.
- 2017 “Digital tools for a metadata entry and collection management software”, NSF workshop *Developing Standards for Data Citation and Attribution for Reproducible Research in Linguistics*, Alcanena, Portugal. October 2017.

### PEER REVIEWED PRESENTATIONS (ABSTRACTS)

- 2025 Changbing Yang, Godfred Agyapong, Patricia Anderson, Sarah Moeller. “Adaptability of NLP Annotation Tools for Linguistics”, 8th Workshop on the Use of Computational Methods in

- the Study of Endangered Languages (ComputEL-8), University of Hawaii at Manoa, March 3-4, 2025.
- Sarah Moeller. "Causes and Costs of the Annotation Bottleneck", 9th International Conference on Language Documentation and Conservation (ICLDC9), University of Hawaii at Manoa, March 5-9, 2025.
- Sarah Moeller and Zoey Liu. "Leveraging AI for Language Documentation: How Much Data?", 9th International Conference on Language Documentation and Conservation (ICLDC9), University of Hawaii at Manoa, March 5-9, 2025.
- 2024 Kevin Tang, Alice Rozet, Sarah Moeller. "Automatic Detection of Morphosyntactic Dialect Features in African American English Oral Histories". MODIFIED 2024: Morphosyntactic Dialect Feature Detection Workshop, Leiden University, Netherlands, June 20-21 2024.
- 2022 Changbing Yang, Andrew Cowell, Sarah Moeller. "How Language Data Comes to Be", 5th Workshop on Computational Methods for the Study of Endangered Languages (ComputEL-5), June 2022, Online.
- 2019 Sarah Moeller. "SayMore X" Tech Showcase at 6th International Conference on Language Documentation and Conservation, University of Hawaii at Manoa, February 28-March 3, 2019.
- 2015 Sarah Moeller. "Developments in SayMore: The language documentation tool for citizen scientists", 4th International Conference on Language Documentation and Conservation, University of Hawaii at Manoa, February 27-March 1, 2015.
- 2011 Sarah Moeller. "Ups and Downs: Bringing Students and Language Communities Together", 2nd International Conference on Language Documentation and Conservation, University of Hawaii at Manoa, February 11-13, 2011, and 3rd Dallas-Ft. Worth Metroplex Linguistics Conference, University of North Texas (UNT), September 10, 2011.
- 2010 Sarah Moeller. "A preliminary analysis of Crimean Tatar written discourse", 2nd Dallas-Ft. Worth Metroplex Linguistics Conference, Graduate Institute of Applied Linguistics, November 5, 2010.

## Teaching & Mentoring

### UNIVERSITY OF FLORIDA

- 2025 Research Methods in Computational Linguistics, *undergraduate/graduate, Spring*
- 2024 Advances in NLP for Low-Resource Languages, *graduate seminar, Spring*
- 2023-2025 Language and Computers, *undergraduate general education (Quest 2), Spring*
- Methods for Language Documentation, *undergraduate/graduate (co-instructor), Fall*
- 2022 Computational Linguistics, *undergraduate/graduate, Spring*
- Language Documentation, *undergraduate/graduate, Spring*
- 2021-22, '24-25 Languages of the World, *undergraduate, Fall*
- 2021-22 Introduction to Python Programming for Linguists, *undergrad/graduate, Fall*

### UNIVERSITY OF COLORADO

- 2020 Python Programming for Linguistics, *undergraduate (instructor), Spring*
- 2019 Python Programming for Linguistics, *undergraduate (teaching assistant), Spring*
- 2017-2018 Languages of the World, *undergraduate (teaching assistant), 3 semesters*
- 2016 Introduction to Linguistics *undergraduate (teaching assistant), Fall*

## DALLAS INTERNATIONAL UNIVERSITY

2015	Introduction to Grammatical Analysis, <i>undergraduate (instructional designer), Fall</i>
2011-2015	Language and Culture Documentation, <i>graduate (co-instructor), Summer</i>

## WORKSHOPS

2023	“Machine-in-the-Loop of Language Documentation”, University of Alberta, Edmonton, Alberta, November 15-17, 2023. “Applying AI to the Study of Language”, Dallas International University, Texas, October 19, 2023.
2016	“Database Design, Annotating, and Archiving”, Institute on Collaborative Language Research (CoLang), Fairbanks, Alaska, June 1-6, 2023.
2015	“SayMore”, Transarch Workshop, Free University of Bolzano-Bozen, December 11, 2023.

## GUEST LECTURES

2023	Graham Civic Center, University of Florida
2020	“Applications of Interlinear Glossed Texts” Department of Linguistics, University of North Texas

## THESIS ADVISEES

### *PhD*

current	Godfred Agyapong
---------	------------------

### *Masters*

2024	Godfred Agyapong
2023	Ryan Cheves

### *Undergraduate*

2024	Wilermine Previlon (AI Scholar)
2026	Pryce Houck (AI Scholar)

## Professional Service

### UNIVERSITY

2024-present	MS in Computational Linguistics Committee, University of Florida
2022-present	Site Coordinator, North American Computational Linguistics Olympiad (NACLO)
2024-2025	Visiting Scholars Committee, Department of Linguistics, University of Florida
2023	Computational Linguistics Search Committee, Department of Linguistics, University of Florida
2022-2023	Linguistics Hiring Committee, Department of Linguistics, University of Florida
2022-2023	Coordinator, NLP@UF Working Group, University of Florida
2021-2024	Undergraduate Vitality Committee, Department of Linguistics, University of Florida
2010	Service Learning Leader, Dallas International University & Mississippi Band of Choctaw Indians

### PROFESSIONAL

2025-2027	Past-president, Association of Computational Linguistics Special Interest Group on Endangered Languages (SIGEL)
2025	Organizing Committee, 8th Workshop on Computational Methods for the Study of Endangered

	Languages (ComputEL-8)
2023-2025	President, Association of Computational Linguistics Special Interest Group on Endangered Languages (SIGEL)
2021-present	Board Member, Ethics Review Committee, International Literacy and Development (ILAD)
2023	Organizing Committee, ACL Special Interest Group on Morphology and Phonology (SIGMORPHON) Shared Task 2
2020-2022	Secretary, Association of Computational Linguistics Special Interest Group on Endangered Languages (SIGEL)
'20, '22, '24	Program Chair, Workshop on Computational Methods for the Study of Endangered Languages (ComputEL-4), (ComputEL-5), (ComputEL-7)
2023	Organizing Committee, 6th Workshop on Computational Methods for the Study of Endangered Languages (ComputEL-6)
2018-2021	Editorial Board, <i>Colorado Research in Linguistics</i> (CRiL)
2000-2013	Volunteer English as Second Language Teacher for refugees and immigrants

#### REVIEWING

SIGMORPHON 2024, 2025

Americas NLP 2023, 2024, 2025

ACL Academic Rolling Review for NAACL, 2024

*Algonquian Papers* 2024

The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (COLING-LREC)

European Chapter of the Association for Computational Linguistics (EACL) 2024

International Workshop on Designing Meaning Representations (DMR) 2019, 2020, 2024

North American Chapter of the Association for Computational Linguistics (NAACL) 2022, 2025

Empirical Methods in Natural Language Processing (EMNLP), 2021

Language Resources and Evaluation Conference (LREC) 2019, 2020

*Language Documentation & Conservation* (LD&C) 2019