

# Luke Gessler

<https://lgessler.com>

[lgessler@iu.edu](mailto:lgessler@iu.edu)

*Last updated August 31, 2024*

## Academic Appointments

<b>Indiana University Bloomington</b>	Assistant Professor Department of Linguistics	2024–present
<b>University of Colorado Boulder</b>	Postdoctoral Associate Department of Computer Science PI: Katharina von der Wense	2023–2024

## Education

<b>Ph.D., Georgetown University</b>	Linguistics Concentration: Computational Linguistics Advisor: Amir Zeldes Title: <i>Low-resource Transformer Language Models</i>	2018–2023
<b>M.S., Georgetown University</b>	Linguistics Concentration: Computational Linguistics	2018–2021
<b>B.A., University of Virginia</b>	Linguistics and Computer Science (with distinction, ΦBK)	2012–2016

## Other Employment

<b>Luminare Inc.</b>	Software Development Consultant	2021
<b>Epic Systems Corporation</b>	Software Developer	2016–2018
<b>Epic Systems Corporation</b>	Software Developer Intern	2015

## Publications

### Refereed Conference Papers

**Luke Gessler.** 2024. PrOnto: Language Model Evaluations for 859 Languages. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. Torino, Italia: ELRA & ICCL. 2024. 13243–13256. <https://aclanthology.org/2024.lrec-main.1159>.

- Enora Rice, Ali Marashian, **Luke Gessler**, Alexis Palmer and Katharina von der Wense. 2024. TAMS: Translation-Assisted Morphological Segmentation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Bangkok, Thailand: Association for Computational Linguistics. 2024. 6752–6765. <https://aclanthology.org/2024.acl-long.366>.
- Luke Gessler** and Nathan Schneider. 2023. Syntactic Inductive Bias in Transformer Language Models: Especially Helpful for Low-Resource Languages? In *Proceedings of the 27th Conference on Computational Natural Language Learning (CoNLL)*. Singapore: Association for Computational Linguistics. 2023. 238–253. <https://aclanthology.org/2023.conll-1.17>.
- Aryaman Arora, **Luke Gessler** and Nathan Schneider. 2020. Supervised Grapheme-to-Phoneme Conversion of Orthographic Schwas in Hindi and Punjabi. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 2020. 7791–7795. <https://aclanthology.org/2020.acl-main.696>.
- Luke Gessler**, Siyao Peng, Yang Janet Liu, Yilun Zhu, Shabnam Behzad and Amir Zeldes. 2020. AMAL-GUM – A Free, Balanced, Multilayer English Web Corpus. English. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association. 2020. 5267–5275. <https://aclanthology.org/2020.lrec-1.648>.
- Mitchell Abrams, **Luke Gessler** and Matthew Marge. 2019. B. Rex: a dialogue agent for book recommendations. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*. Stockholm, Sweden: Association for Computational Linguistics. 2019. 418–421. <https://aclanthology.org/W19-5948>.

## Refereed Workshop Papers

- Luke Gessler** and Katharina von der Wense. 2024. NLP for Language Documentation: Two Reasons for the Gap between Theory and Practice. In *Proceedings of the 4th Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP 2024)*. Mexico City, Mexico: Association for Computational Linguistics. 2024. 1–6. <https://aclanthology.org/2024.americasnlp-1.1>.
- Tatsuya Aoyama, Shabnam Behzad, **Luke Gessler**, Lauren Levine, Jessica Lin, Yang Janet Liu, Siyao Peng, Yilun Zhu and Amir Zeldes. 2023. GENTLE: A Genre-Diverse Multilayer Challenge Set for English NLP and Linguistic Evaluation. In *Proceedings of the 17th Linguistic Annotation Workshop (LAW-XVII)*. Toronto, Canada: Association for Computational Linguistics. 2023. 166–178. <https://aclanthology.org/2023.law-1.17>.
- Luke Gessler**. 2022. Closing the NLP Gap: Documentary Linguistics and NLP Need a Shared Software Infrastructure. In *Proceedings of the Fifth Workshop on the Use of Computational Methods in the Study of Endangered Languages*. Dublin, Ireland: Association for Computational Linguistics. 2022. 119–126. <https://aclanthology.org/2022.computel-1.15>.
- Luke Gessler**, Austin Blodgett, Joseph C. Ledford and Nathan Schneider. 2022. Xposition: An Online Multilingual Database of Adpositional Semantics. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. Marseille, France: European Language Resources Association. 2022. 1824–1830. <https://aclanthology.org/2022.lrec-1.194>.
- Luke Gessler**, Lauren Levine and Amir Zeldes. 2022. Midas Loop: A Prioritized Human-in-the-Loop Annotation for Large Scale Multilayer Data. In *Proceedings of the 16th Linguistic Annotation Workshop (LAW-XVI) within LREC2022*. Marseille, France: European Language Resources Association. 2022. 103–110. <https://aclanthology.org/2022.law-1.13>.
- Luke Gessler** and Amir Zeldes. 2022. MicroBERT: Effective Training of Low-resource Monolingual BERTs through Parameter Reduction and Multitask Learning. In *Proceedings of the 2nd Workshop*

on *Multi-lingual Representation Learning (MRL)*. Abu Dhabi, United Arab Emirates (Hybrid): Association for Computational Linguistics. 2022. 86–99. <https://aclanthology.org/2022.mrl-1.9>.

**Luke Gessler**, Shabnam Behzad, Yang Janet Liu, Siyao Peng, Yilun Zhu and Amir Zeldes. 2021. DisCoDisCo at the DISRPT2021 Shared Task: A System for Discourse Segmentation, Classification, and Connective Detection. In *Proceedings of the 2nd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2021)*. Punta Cana, Dominican Republic: Association for Computational Linguistics. 2021. 51–62. <https://aclanthology.org/2021.disrpt-1.6>.

**Luke Gessler** and Nathan Schneider. 2021. BERT Has Uncommon Sense: Similarity Ranking for Word Sense BERTology. In *Proceedings of the Fourth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*. Punta Cana, Dominican Republic: Association for Computational Linguistics. 2021. 539–547. <https://aclanthology.org/2021.blackboxnlp-1.43>.

**Luke Gessler**, Shira Wein and Nathan Schneider. 2020. Supersense and Sensibility: Proxy Tasks for Semantic Annotation of Prepositions. In *Proceedings of the 14th Linguistic Annotation Workshop*. Barcelona, Spain: Association for Computational Linguistics. 2020. 117–126. <https://aclanthology.org/2020.law-1.11>.

Graham Neubig, Shruti Rijhwani, Alexis Palmer, Jordan MacKenzie, Hilaria Cruz, Xinjian Li, Matthew Lee, Aditi Chaudhary, **Luke Gessler**, Steven Abney, Shirley Anugrah Hayati, Antonios Anastasopoulos, Olga Zamaraeva, Emily Prud’hommeaux, Jennette Child, Sara Child, Rebecca Knowles, Sarah Moeller, Jeffrey Micher, Yiyuan Li, Sydney Zink, Mengzhou Xia, Roshan Sharma and Patrick Littell. 2020. A Summary of the First Workshop on Language Technology for Language Documentation and Revitalization. English. 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)*. Marseille, France: European Language Resources association. 2020. 342–351. <https://aclanthology.org/2020.sltu-1.48>.

**Luke Gessler**. 2019. Developing without developers: choosing labor-saving tools for language documentation apps. In *Proceedings of the 3rd Workshop on the Use of Computational Methods in the Study of Endangered Languages Volume 1 (Papers)*. Honolulu: Association for Computational Linguistics. 2019. 6–13. <https://aclanthology.org/W19-6002>.

**Luke Gessler**, Yang Janet Liu and Amir Zeldes. 2019. A Discourse Signal Annotation System for RST Trees. In *Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019*. Minneapolis, MN: Association for Computational Linguistics. 2019. 56–61. <https://aclanthology.org/W19-2708>.

## Invited Presentations

- Low-resource Transformer Language Models. 2024. Guest lecture given in CSCI 3832 at University of Colorado Boulder.
- MicroBERT and PrOnto: Training and Evaluating Language Models for Low-resource Languages. 2023. Guest lecture given in CSC401 at University of Ontario Institute of Technology.
- Discourse Unit Segmentation, Connective Detection, and Relation Classification and DisCoDisCo at the DISRPT 2021 Shared Task. 2023. Guest lecture given in LING 7800 at University of Colorado Boulder.
- Sociolinguistic Variation in Japanese Particles. 2022. Guest lecture given in LING 424 at Georgetown University.
- Cases and Adpositions in Latin. 2022. Guest lecture given in LING 424 at Georgetown University.
- Neural Sequence-to-Sequence Models. 2022. Guest lecture given in COSC 572 at Georgetown

- University.
- Optimizing Developer Productivity in Endangered Language Documentation Apps. 2019. Talk given at deRSE19.

## Teaching

### Indiana University (Instructor)

- Foundations of Computational Linguistics (LING-L 435/LING-L 555)  
Fall 2024  
Graduate-level introduction to computer programming with a focus on applications in CL.

### Georgetown University (TA)

- All About Prepositions (LING 424)  
Fall 2022  
Cross-lingual study of adpositions and case marking, with emphasis on semantics.
- Empirical Methods in Natural Language Processing (COSC 572)  
Spring 2022  
Graduate-level introduction to natural language processing for computer science students.
- Computational Corpus Linguistics (LING 372)  
Fall 2019, Fall 2020  
Introduction to corpus linguistics with a focus on multilayer annotation and practical skills in corpus construction and querying.

## Institutional Service

### Indiana University

- Computational Linguistics Outreach Coordinator. Fall 2024. Promote CL degree programs at IU.
- Social Organizer, ClingDing. Fall 2024. Organize social activities for the IU computational linguistics community.

### Georgetown University

Service activities as a Ph.D. student, 2018–2023:

- Organizer, Georgetown University Computational Linguistics Reading Group, 2022. Organized and facilitated a semiweekly reading group for contemporary papers in NLP and CL, driven by student interest.
- Social Media Chair, Georgetown University Computational Linguistics. 2021–2023. Managed Twitter and other accounts for the GUCL community, promoting papers published by members and engaging with the computational linguistics and natural language processing community on Twitter.
- Instructor, AllenNLP Workshop. 2021. Designed and ran a half-day workshop where attendees, who know Python, were introduced to AllenNLP, a deep learning framework for NLP. <https://github.com/lgessler/allennlp-workshop>
- Social Chair, Georgetown University Computational Linguistics. 2020–2023. Organized social activities such as happy hours, movie nights, and game nights to build community in the Georgetown computational linguistics community.

- Organizer, History of Linguistics Reading Group. 2020. Organized and moderated a 3-month weekly reading group covering select primary literature in linguistics’ historical development from Panini to American Structuralism. [https://docs.google.com/document/d/10RjmOTCNwsBP8ErLqga\\_sptpOfCFN8Won5hle6FMlJpA/edit?usp=sharing](https://docs.google.com/document/d/10RjmOTCNwsBP8ErLqga_sptpOfCFN8Won5hle6FMlJpA/edit?usp=sharing)
- Instructor, Academic CV Latex Workshop. 2019. Designed and ran a day-long workshop where attendees learned the basics of LaTeX and created their own CVs. <https://github.com/lgessler/glsa-latex-workshop>
- Instructor, Python for Linguists. 2018. Designed and ran a day-long introduction Python for linguistics graduate students, with no programming experience assumed. <https://github.com/lgessler/python-for-linguists>
- Organizer, Georgetown University Fieldwork Forum. 2018–2019. Co-founded a working/reading group for students and faculty who do linguistic fieldwork.

## University of Virginia

- President, Virginia Linguistics Club. 2014–2016. Founded and ran UVA’s first linguistics club, organizing social events like documentary screenings and hosting guest speakers.

## Professional Service

### Conference and Workshop Roles

- Handbook Chair, NAACL 2024.
- Publication and Publicity Co-chair, Linguistic Annotation Workshop XIV. 2020. Managed LAW 2020’s website and social media presence and generated conference proceedings.

### Reviewing

- Action Editor, ACL Rolling Review (ARR). ARR is a centralized reviewing platform that serves all main conferences administered by the Association for Computational Linguistics. Action Editors assign and supervise reviews and write meta-reviews.  
December 2023, February 2024, June 2024, August 2024
- Area Chair for Corpora and Annotation, Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING). Area Chairs oversee the reviewing process and give write meta-reviews.
- Reviewer, Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL)  
2022, 2024
- Reviewer, Linguistic Annotation Workshop (LAW)  
2020, 2023, 2024
- Reviewer, Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP)  
2021, 2023, 2024
- Reviewer, Conference on Natural Language Learning (CoNLL)  
2024
- Reviewer, Association for Computational Linguistics  
2023

- Reviewer, Annual Meeting of the Cognitive Science Society (CogSci)  
2023
- Reviewer, Workshop on Natural Language Processing for Digital Humanities (NLP4DH)  
2021
- Reviewer, International Conference on Language Resources and Evaluation (LREC)  
2020, 2021

## **Professional Memberships**

- Association for Computational Linguistics