Luke Gessler

Ph.D. Student, Georgetown University
Last updated June 10, 2023

Contact
lukegessler@gmail.com
https://lgessler.com

Education

2018 – 2023	**Georgetown University** – Washington, DC Ph.D. student in Computational Linguistics Dissertation Title: *Low-resource Monolingual Transformer Language Models* Committee: Amir Zeldes, Nathan Schneider, Antonios Anastasopoulos GPA: 3.97/4.00 Selected coursework: Syntax I & II; Semantics I & II; Phonology I; Computational Grammar Formalisms; Sociolinguistic Variation; Dialogue Systems; Social Factors in AI; Second Language Acquisition; Philosophy of Language
2018 – 2021	**Georgetown University** – Washington, DC M.S. in Computational Linguistics
2012 – 2016	**University of Virginia** – Charlottesville, VA B.A. Computer Science, B.A. Linguistics, with distinction GPA: 3.83/4.00

Awards and Honors

2018–2023	**Georgetown University Graduate Student Assistantship** Merit-based award of tuition and a stipend for the duration of the Ph.D. program.
2016	**Phi Beta Kappa**
2013, 2014	**Critical Language Scholarship** Full funding to spend two summers in India to study Hindi. Tested at CEFR level C1.
2012	**Maureen O'Donnell Oxford Classical Dictionary Award** Awarded for winning five gold medals in the National Latin Exam.

Conference and Workshop Publications

2022	**Luke Gessler**, Amir Zeldes. 2022. MicroBERT: Effective Training of Low-resource Monolingual BERTs through Parameter Reduction and Multitask Learning. In *Proc. of MRL 2022* at EMNLP, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
2022	**Luke Gessler**. 2022. Closing the NLP Gap: Documentary Linguistics and NLP Need a Shared Software Infrastructure. In *Proc. of ComputEL-5* at ACL, Dublin, Ireland. Association for Computational Linguistics.

2022   **Luke Gessler**, Lauren Levine, Amir Zeldes. 2022. Midas Loop: A Prioritized Human-in-the-Loop Annotation for Large Scale Multilayer Data. In: *Proc. of LAW XVI* at LREC, Marseille, France. European Language Resources Association.

2022   **Luke Gessler**, Austin Blodgett, Joseph Ledford, Nathan Schneider. 2022. Xposition: An Online Multilingual Database of Adpositional Semantics. In *Proc. of LREC*, Marseille, France. European Language Resources Association.

2021   **Luke Gessler**, Nathan Schneider. 2021. BERT has uncommon sense: similarity ranking for word sense BERTology. In *Proc. of BlackboxNLP* at EMNLP, Punta Cana, Dominican Republic. Association for Computational Linguistics.

2021   **Luke Gessler**, Shabnam Behzad, Yang Janet Liu, Siyao Peng, Yilun Zhu, Amir Zeldes. 2021. DisCoDisCo at the DISRPT2021 shared task: a system for discourse segmentation, classification, and connective detection. In *Proc. of CODI-DISRPT* at EMNLP, Punta Cana, Dominican Republic. Association for Computational Linguistics. **Scored 1st place in all subtasks.**

2020   **Luke Gessler**, Siyao Peng, Yang Janet Liu, Yilun Zhu, Shabnam Behzad, and Amir Zeldes. 2020. AMALGUM – a free, balanced, multilayer English web corpus. In *Proc. of LREC*, pages 5267-5275, online. European Language Resources Association.

2020   Graham Neubig, Shruti Rijhwani, Alexi Palmers, Jordan MacKenzie, Hilaria Cruz, Xinjian Li, Matthew Lee, Aditi Chaudhary, **Luke Gessler**, Steven Abney, Shirley Anugrah Hayati, Anastasopoulos, Antonios, Olga Zamaraeva, Emily Prud'hommeaux, Jennette Child, Sara Child, Knowles, Rebecca, Sarah Moeller, Jeffrey Micher, Yiyuan Li, Sydney Zink, Mengzhou Xia, Roshan S. Sharma, and Patrick Littell. 2020. A summary of the first workshop on language technology for language documentation and revitalization. In *Proc. of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)* at LREC, pages 342-351, online. European Language Resources Association.

2020   Aryaman Arora, **Luke Gessler**, and Nathan Schneider. 2020. Supervised grapheme-to-phoneme conversion of orthographic schwas in Hindi and Punjabi. In *Proc. of ACL*, pages 7791-7795, online. Association for Computational Linguistics.

2019   **Luke Gessler**. 2019. Developing without developers: choosing labor-saving tools for language documentation apps. In *Proc. of ComputEL-3* at ICLDC, pages 6-13, Honolulu, Hawaii, United States. Association for Computational Linguistics.

2019   **Luke Gessler**, Yang Janet Liu, and Amir Zeldes. 2019. A discourse signal annotation system for RST trees. In *Proc. of the Workshop on Discourse Relation Parsing and Treebanking* at NAACL-HLT, pages 56-61, Minneapolis, Minnesota, United States. Association for Computational Linguistics..

2019   Mitchell Abrams, **Luke Gessler**, and Matthew Marge. 2019. B. Rex: a dialogue agent for book recommendations. In Proc. of SIGDIAL, pages 418-421, Stockholm, Sweden. Association for Computational Linguistics.

Presentations

2022	**Luke Gessler**, *Sociolinguistic Variation in Japanese Particles*. Guest lecture given in LING 424 at Georgetown University.
2022	**Luke Gessler**, *Cases and Adpositions in Latin*. Guest lecture given in LING 424 at Georgetown University.
2022	**Luke Gessler**, *Neural Sequence-to-Sequence Models*. Guest lecture given in COSC 572 at Georgetown University.
2019	**Luke Gessler**, *Optimizing Developer Productivity in Endangered Language Documentation Apps*. Talk given at deRSE19.

Teaching: Courses

Fall 2022	**TA, All About Prepositions (LING 424) – Georgetown University** Assisted in developing and teaching a new upper-level undergraduate course focused on cross-lingual study of all aspects of adpositions and case marking, with emphasis on semantics. Duties included developing assignments, grading, and holding office hours.
Spring 2022	**TA, Empirical Methods in Natural Language Processing (COSC 572) – Georgetown University** Assisted the graduate-level introduction to NLP in the computer science department, covering fundamental NLP algorithms and an introduction to deep learning for NLP. Duties included grading assignments, lecturing, and holding office hours.
Fall 2020	**TA, Computational Corpus Linguistics (LING 372) – Georgetown University** Assisted an intensive semester-long course focused on sophisticated linguistic annotation of text documents, including part of speech tags, Universal Dependencies syntactic parses, named and non-named entities, and more. Duties included grading assignments, advising students' final projects, and holding office hours.
Fall 2019	**TA, Computational Corpus Linguistics (LING 372) – Georgetown University** *See above.*

Teaching: Workshops

Nov. 2021	**Instructor, AllenNLP Workshop** 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`
Nov. 2019	**Instructor, Academic CV Latex Workshop** 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`

Nov. 2018	**Instructor, Python for Linguists** Taught a day-long course with original materials (github.com/lgessler/python-for-linguists) introduction Python for linguistics graduate students, with no programming experience assumed. `https://github.com/lgessler/python-for-linguists`

Department Service

2022	**Reading Group Organizer, Georgetown University Computational Linguistics** Organized and facilitated a semiweekly reading group for contemporary papers in NLP and CL, driven by student interest.
2021 –	**Social Media Chair, Georgetown University Computational Linguistics** Manage 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.
2020 –	**Social Chair, Georgetown University Computational Linguistics** Organize social activities such as happy hours, movie nights, and game nights to build community in the Georgetown computational linguistics community.
2020	**Organizer, History of Linguistics Reading Group** Organized and hosted a 3-month weekly reading group covering select primary literature in linguistics' historical development from Panini to American Structuralism. Syllabus: `https://docs.google.com/document/d/10RjmOTCNwsBP8ErLqgasptpOfCFN8Won5hle6FMlJpA/edit?usp=sharing`
2018 – 2019	**Founder and Organizer, Georgetown University Fieldwork Forum** Created a working/reading group for students and faculty who do linguistic fieldwork.
2014 – 2016	**Founder, Virginia Linguistics Club** Founded and ran the University of Virginia's first linguistics club, organizing social events like screenings of linguistics-related documentaries and hosting guest speakers.

Academic Service

2019	**Co-publicity chair, Linguistic Annotation Workshop XIV 2020** Managing LAW 2020's website and social media presence.

Reviewer Service

2023	Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP)
2023	Linguistic Annotation Workshop (LAW-XVII)
2023	Association for Computational Linguistics (ACL 2023)
2023	Annual Meeting of the Cognitive Science Society (CogSci 2023)
2022	Workshop On The Use of Computational Methods in the Study of Endangered Languages (ComputEL-5)
2022	International Conference on Language Resources and Evaluation (LREC)

2021	Workshop on Natural Language Processing for Digital Humanities (NLP4DH)
2021	Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP)
2020	International Conference on Language Resources and Evaluation (LREC)

Other Service

| 2022 – | **Mentor, Spark The Journey**
Provide one-to-one academic mentoring to a high schooler in DC public schools. |
| 2018 – 2022 | **Webmaster, Veg Society of DC**
Manage a small 501(c)(3) nonprofit's web operations (G Suite, DNS, etc.). |

Professional Experience

2021	**Software Development Consultant, Luminare Inc.** – Houston, TX Rapid prototyping on a web application for health analytics leveraging my expertise in tools including Clojure, ClojureScript, and Fulcro.
2016 – 2018	**Software Developer, Epic Systems Corporation** – Madison, WI Made software for the United States' most widely-used electronic medical record software, including a headliner feature for the 2018 update.
2015	**Software Development Intern, Epic Systems Corporation** – Madison, WI Developed an IVR (interactive voice response) system integrated into Epic's software.

Programming Languages

| Proficient | Clojure, Python, JavaScript, Java, C#, MUMPS |
| Familiar | Scheme, Common Lisp, OCaml, C, Go, bash |

Languages

Native	English
Proficient	Hindi
Reading	Latin, Sahidic Coptic