

Aaron M. Mueller

| | | |
|---------------------|---|---|
| CONTACT INFORMATION | Center for Language & Speech Processing Johns Hopkins University 3400 N. Charles St., Hackerman 319 Baltimore, MD 21218-2608 (U.S.A.) | <i>E-mail:</i> amueller@jhu.edu <i>Website:</i> aaronmueller.github.io <i>Code:</i> github.com/aaronmueller |
| RESEARCH INTERESTS | <ul style="list-style-type: none">• Multilingual natural language processing• Computational psycholinguistics• Machine translation | |
| EDUCATION | Johns Hopkins University , Baltimore, MD Ph.D. student, Computer Science. M.S.E., Computer Science. GPA: 3.9/4.0. <i>Advisors:</i> Mark Dredze, David Yarowsky. University of Kentucky , Lexington, KY B.S., Computer Science. Honors. B.S., Linguistics. Honors. GPA: 4.0/4.0. <i>Summa cum laude.</i> <i>Thesis:</i> Neural Machine Translation for Canadian French. | Present May 2020 May 2018 May 2018 |
| RESEARCH EMPLOYMENT | BBN Technologies , Cambridge, MA <i>Analytics & Machine Intelligence Research Intern</i> Supervisors: Roger Bock, Ilana Heintz – Research in improving the speed-accuracy trade-off in neural machine translation, word alignment, and entity linking. University of Kentucky , Lexington, KY Institutional Research & Advanced Analytics <i>Research Assistant (Machine Learning)</i> Supervisors: Nathan Jacobs, Craig Rudick – Implemented deep LSTMs to learn student and course profiles for grade prediction. – Designed various ordinal loss functions. University of Massachusetts Amherst , Amherst, MA Statistical Social Language Analysis Lab <i>Research Assistant (Natural Language Processing)</i> Supervisor: Brendan O'Connor – Research in entity-event extraction. – Integrated entity mention and relation systems into pre-existing sentence-level model. Also integrated document-level information (e.g., noun coreferences) as features. University of Kentucky , Lexington, KY Linguistics Department <i>Juried Project</i> Jurors: Hilaria Cruz, Raphael Finkel, Phil Harling – Preservation and revitalization project for a low-resource indigenous language—Chatino—spoken natively by approximately 40,000 individuals. – Trained an automatic speech recognition system with Sphinx. – Promoted and created Chatino language-learning resources, including open-source speech data, a speech corpus, and a website. | May - August 2019 May - August 2018 May - August 2017 August 2016 - May 2017 |

Johns Hopkins University, Baltimore, MD
Center for Language & Speech Processing
Research Assistant (Natural Language Processing)

May - August 2016

Supervisor: David Yarowsky

- Research in statistical machine translation for low-resource languages.
- Implemented a lemma-based English-Uyghur translation model, built a morphological generator for Crimean Tatar, and created parsers to extract translation tables from foreign editions of Wiktionary.

PAPERS

REFEREED CONFERENCE PROCEEDINGS

1. **Aaron Mueller**, Garrett Nicolai, Panayiota Petrou-Zeniou, Natalia Talmina, and Tal Linzen. *Cross-Linguistic Syntactic Evaluation of Word Prediction Models*. Association for Computational Linguistics (ACL), 2020.
2. **Aaron Mueller**, Garrett Nicolai, Arya D. McCarthy, Dylan Lewis, Winston Wu, and David Yarowsky. *An Analysis of Massively Multilingual Neural Machine Translation for Low-Resource Languages*. Language Resources and Evaluation Conference (LREC), 2020.
3. Arya D. McCarthy, Rachel Wicks, Dylan Lewis, **Aaron Mueller**, Winston Wu, Oliver Adams, Garrett Nicolai, Matt Post, and David Yarowsky. *The Johns Hopkins University Bible Corpus: 1600+ Tongues for Typological Exploration*. Language Resources and Evaluation Conference (LREC), 2020.
4. Garrett Nicolai, Dylan Lewis, Arya D. McCarthy, **Aaron Mueller**, Winston Wu, and David Yarowsky. *Fine-Grained Morphosyntactic Analysis and Generation Tools for More Than One Thousand Languages*. Language Resources and Evaluation Conference (LREC), 2020.
5. Marten van Schijndel, **Aaron Mueller**, and Tal Linzen. *Quantity Doesn't Buy Quality Syntax with Neural Language Models*. Empirical Methods in Natural Language Processing (EMNLP), 2019.
6. Arya D. McCarthy, Winston Wu, **Aaron Mueller**, Bill Watson, and David Yarowsky. *Modeling Color Terminology Across Thousands of Languages*. Empirical Methods in Natural Language Processing (EMNLP), 2019.
7. **Aaron Mueller*** and Yash Kumar Lal*. *Sentence-Level Adaptation for Low-Resource Neural Machine Translation*. Machine Translation Summit (MTSummit) Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT), 2019. [***Equal contribution**]

NON-REFEREED AND OTHERS

8. **Aaron Mueller**, Mark Richard Lauersdorf, Kevin McGowan, and Ramakanth Kavuluru. *Neural Machine Translation for Canadian French*. Undergraduate thesis, 2018.

FELLOWSHIPS AND AWARDS

National Science Foundation Graduate Research Fellow 2018 - 2023

Gaines Fellow 2016 - 2018

Two-year fellowship awarded to undergraduates based on academic performance, ability to conduct research, an interest in public issues, and a desire to enhance understanding of the human condition through the humanities. Requires the completion of a juried project, a thesis project, and a seminar in the humanities. (\$5,000)

Patterson Scholar 2014 - 2018

Awarded to University of Kentucky students who have earned National Merit semifinalist standing or better. (\$80,000+)

Raymond F. Betts Scholar 2017

Awarded to rising seniors conducting thesis research. Used funds to study Québec French in Montréal and Québec City during winter of 2017–2018. (\$2,500)

Goldwater Scholarship (Honorable Mention) 2017

Phi Beta Kappa 2017

Linguistics Research Award 2016

Awarded to a University of Kentucky undergraduate to facilitate a year-long research project in linguistics. Used funds for a quantitative sociolinguistics project examining semantic shifts in gendered words in U.S. media over 200 years. (\$500)

National Merit Semifinalist

2014

SELECTED **Aaron Mueller.** *Québec French and Language Technology.* 2018 Betts & Rowland Scholar Awards
PRESENTATIONS Ceremony, Lexington, KY, April 4, 2018.

Aaron Mueller, Katherine Keith, Abram Handler, Su Lin Blodgett, and Brendan O'Connor. *The Identification of Civilians Killed by Police with Supervised Entity-Event Extraction.* 2017 UMass Amherst Research Experience for Undergraduates (REU) Showcase, Amherst, MA, August 9, 2017.

Aaron Mueller, Raphael Finkel, and Hilaria Cruz. *Documenting and Promoting the Chatino Language and Orthography.* Juried Presentation in Satisfaction of the Requirements of the Gaines Fellowship, Lexington, KY, February 21, 2017.

Aaron Mueller, Huda Khayrallah, Winston Wu, and David Yarowsky. *A Lemma-Based Approach for English-Uyghur Statistical Machine Translation.* 9th Annual Conference of the Illinois Language and Linguistics Society (ILLS9), Urbana, IL, March 31–April 1, 2017.

Aaron Mueller and Mark Richard Lauersdorf. *Lexical and Semantic Shifts in the Linguistic Construction of Social Gender: A Corpus-Based Analysis of Written U.S. English.* Poster. 9th Annual Toronto Undergraduate Linguistics Conference (TULCON9), Toronto, ON, March 4–6, 2016.

TEACHING **Johns Hopkins University,** Baltimore, MD

EXPERIENCE *Teaching Assistant*

- Machine Learning: AI System Design & Development, Mathias Unberath. Spring 2020.

ADDITIONAL **Chellgren Center for Undergraduate Excellence,** Lexington, KY

EXPERIENCE *Student Outreach Team*

August 2017 - May 2018

- Organized panel discussions, workshops, and lectures on abstract writing, research poster design, and academic presentations.
- Led the review committee for our Showcase of Undergraduate Scholars, an annual research conference for University of Kentucky undergraduates.

UK Special Collections Research Center, Lexington, KY

Archiving Intern

September 2015 - May 2016

- Created a diachronic corpus of Lexingtonian architectural documentation which includes zoning ordinances and Lexington Planning Committee meeting minutes.
- Cooperated on a multifformat project that layered GIS technology, city government data, archival photos, and digital humanities tools.

COMPUTING **Programming:**

SKILLS

- Languages: Python, C++, C, HTML, CSS, Javascript, Java, UNIX shell scripting.
- Machine Learning Toolkits: PyTorch (incl. huggingface, fairseq, sockeye), Keras, TensorFlow.
- Version Control: DVCS (Git, Bitbucket).
- Development Environments: Vim, Emacs, VSCode.

Linguistic Tools:

Praat, AntConc, QGIS, Audacity.

LANGUAGES English (native language), French (B2, Canadian). Experience with Chatino and Uyghur through research.

INTERESTS AND **– Music:** guitar (classical, electric), bass, banjo, synthesizer, MaxMSP, Ableton Live.

HOBBIES

- Athletics: cold-weather/snow hiking, swing dance, contra.
- Hobbies: natural language learning, reading (philosophy, politics), city exploration.
- Cooking: Thai, French, German, Southern.