

# Aaron M. Mueller

---

**COORDONNÉES** Center for Language & Speech Processing  
Johns Hopkins University  
3400 N. Charles St., Hackerman 319  
Baltimore, MD 21218-2608 (É-U)

*Courriel:* amueller@jhu.edu  
*Site web:* aaronmueller.github.io

**INTÉRÊTS DE RECHERCHE**

- TALN multilingue
- Psycholinguistique informatique
- Traduction automatique
- Langues menacées et à faibles ressources

**FORMATION** **Johns Hopkins University**, Baltimore, MD, É-U

Doctorant en informatique.

Depuis août 2018

Maîtrise en informatique.

Mai 2020

*Affiliation:* Center for Language and Speech Processing.

*Directeurs de thèse:* Mark Dredze, David Yarowsky.

**Université du Kentucky**, Lexington, KY, É-U

Baccalauréat en informatique (avec mention).

Mai 2018

Baccalauréat en linguistique (avec mention).

Mai 2018

GPA: 4.0/4.0. *Summa cum laude*.

*Thèse:* Neural Machine Translation for Canadian French.

*Directeurs de thèse:* Ramakanth Kavuluru, Mark Richard Lauersdorf.

**EXPÉRIENCE** **BBN Technologies**, Cambridge, MA, É-U

**EN RECHERCHE** *Stagiaire en recherche linguistique et multimédia*

Mai - août 2019

Superviseurs: Roger Bock, Ilana Heintz

– Recherche en traduction automatique neuronale, alignement des mots et annotation sémantique.

– Financé par le programme AIDA (Active Interpretation of Disparate Alternatives) de la DARPA.

**Université du Kentucky**, Lexington, KY, É-U

**Institutional Research & Advanced Analytics**

*Assistant de recherche en apprentissage automatique*

Mai - août 2018

Superviseurs: Nathan Jacobs, Craig Rudick

– Dessiner un système qui prédit les notes futures d'un étudiant donné sa performance précédente.

– Mettre en œuvre les LSTMs profonds et employer les techniques TALN pour apprendre les profils étudiants et cours.

– Dessiner une fonction objectif ordinaire pour améliorer la performance du système.

**University du Massachusetts à Amherst**, Amherst, MA, É-U

**Statistical Social Language Analysis Lab**

*Assistant de recherche en TALN*

Mai - août 2017

Superviseur: Brendan O'Connor

– Collaborer avec des doctorants et professeurs pour améliorer un système d'extraction des entités et événements.

– Donner des articles de nouvelles, utiliser la supervision distante pour soutenir les noms des citoyens américains qui ont été tués par la police.

– Intégrer les systèmes d'extraction des relations des entités dans le modèle préexistant au niveau des phrases complètes.

– Intégrer l'information au niveau du document (p. ex., les coréférences de noms) comme caractéristiques pour améliorer notre système.

**Université du Kentucky**, Lexington, KY, É-U

**Département de linguistique**

*Projet de jury*

Août 2016 - mai 2017

Jurés: Hilaria Cruz, Raphael Finkel, Phil Harling

- Projet de préservation et revitalisation pour une langue autochtone à faible ressources—Chatino—parlée par 40 000 personnes.
- Entraîner un système de reconnaissance automatique de la parole avec Sphinx.
- Promouvoir et créer des ressources pour l'apprentissage des langues chatino, dont les données source-ouvert de parole, un corpus de parole et un site web.

**Johns Hopkins University**, Baltimore, MD, É-U

**Center for Language & Speech Processing**

*Assistant de recherche en TALN*

Mai - août 2016

Superviseur: David Yarowsky

- Collaborer avec des doctorants et professeurs pour créer un système de traduction automatique statistique pour les langues à faibles ressources.
- Mettre en œuvre un modèle de traduction anglais-ouïghour basé sur les lemmes.
- Construire un générateur morphologique pour le tatar de Crimée.
- Créer des analyseurs syntaxiques pour soutirer les tables de traduction des éditions étrangères de Wiktionnaire.
- Financé par le programme LORELEI (Low-Resource Languages for Emergent Incidents) de la DARPA.

BOURSES ET  
PRIX

**Boursier de recherche de la National Science Foundation**

2018 - 2023

**Boursier Gaines**

2016 - 2018

Bourse de recherche de deux ans décernée aux étudiants de premier cycle pour leur rendement scolaire, capacité de mener des recherches, intérêt pour les affaires publiques et désir de faire avancer les connaissances de la condition humaine à travers les sciences humaines. Elle nécessite l'achèvement d'un projet de jury, un thèse et un cours séminaire en sciences humaines. (\$5 000 USD)

**Boursier Patterson**

2014 - 2018

Décernée aux étudiants à l'Université du Kentucky qui ont gagné le Bourse national du mérite. (\$80 000+ USD)

**Boursier Raymond F. Betts**

2017

Décernée à un étudiant par an pour aider ses travaux de thèse. Utilisé les fonds pour étudier le français québécois à Montréal et à Québec pendant l'hiver de 2017–2018. (\$2 500 USD)

**Bourse Goldwater (Mention honorable)**

2017

**Phi Beta Kappa**

2017

**Prix de recherche en linguistique**

2016

Attribué à un étudiant à l'Université du Kentucky pour faciliter un projet de recherche d'une année en linguistique. Utilisé les fonds pour un projet de sociolinguistique quantitative examinant les glissements sémantiques des mots sexospécifiques dans les médias américains sur 200 ans. (\$500 USD)

**Bourse nationale du mérite**

2014

PUBLICATIONS

1. **Aaron Mueller**, Garrett Nicolai, Panayiota Petrou-Zeniou, Natalia Talmina, et 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, et 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, et 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, et 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**, et 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, et David Yarowsky. *Modeling Color Terminology Across Thousands of Languages*. Empirical Methods in Natural Language Processing (EMNLP), 2019.
7. **Aaron Mueller\*** et 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. [**\*Contribution égale**]

**PRÉSENTATIONS** **Aaron Mueller**. Québec French and Language Technology. *Cérémonie de Boursiers Betts & Rowland* 2018, Lexington, KY, É-U, 4 avril 2018.

**REPRÉSENTATIVES**

**Aaron Mueller**, Katherine Keith, Abe Handler, Su Lin Blodgett, et 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, É-U, 9 août 2017.

**Aaron Mueller**, Raphael Finkel, et Hilaria Cruz. Documenting and Promoting the Chatino Language and Orthography. *Présentation de jury dans la satisfaction des besoins de la bourse Gaines*, Lexington, KY, É-U, 21 février 2017.

**Aaron Mueller**, Huda Khayrallah, Winston Wu, et 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, É-U, 31 mars–1 avril 2017.

**Aaron Mueller** et 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, Canada, 4–6 mars 2016.

**EXPÉRIENCE EN** **Johns Hopkins University**, Baltimore, MD, É-U

**ENSEIGNEMENT** *Assistant*

Janvier - mai 2020

– Machine Learning: AI System Design and Development, Mathias Unberath. Printemps 2020.

**EXPÉRIENCE** **Chellgren Center for Undergraduate Excellence**, Lexington, KY, É-U

**SUPPLÉMENTAIRE** *Équipe de sensibilisation des étudiants*

Août 2017 - mai 2018

- Promouvoir le Bureau des recherches de 1<sup>er</sup> cycle et la Bourse Chellgren.
- Organiser les tables rondes, ateliers et lectures sur l'écriture des résumés, le dessin des posters de recherches et les présentations académiques.
- Diriger le comité d'examen pour la Showcase of Undergraduate Scholars, une conférence universitaire annuelle pour les étudiants de l'Université du Kentucky de présenter leurs recherches aux autres étudiants et professeurs.

**UK Special Collections Research Center**, Lexington, KY, É-U

*Stagiaire en archivage*

Septembre 2015 - mai 2016

- Créer un corpus diachronique de documentation architecturale de Lexington, dont les ordonnances de zonage et les procès-verbaux du comité organisateur de Lexington.
- Rendre accessible une collection d'archives sur l'histoire architecturale de Lexington.
- Coopérer sur un projet multimédia, dont les logiciels SIG, données de municipalité, photos d'archives, et outils d'humanités numériques.

**VOLONTARIAT** **Amnesty International**

Depuis 2012

- Organisation internationale dévouée à la protection des droits de l'homme.
- Membre d'Amnesty International à Johns Hopkins et à Washington, D.C.
- *Fondateur* et ancien *coprésident* d'Amnesty International à l'Université du Kentucky.
- Organiser les collectes de fonds et manifestations d'information : les concerts avec des groupes locaux, tables rondes sur les questions contemporaines, débats de faculté, pétitions, protestations et plus.
- Participant à la Conférence nationale d'Amnesty International sur les droits de l'homme.

**Oxfam International**

2014 - 2018

- Organisation internationale dévouée à l'élimination de la pauvreté, les secours aux sinistrés, and l'action politique pour les victimes de l'injustice dans le monde entier.
- *Fondateur* and ancien *co-coordonnateur* d'Oxfam à l'Université du Kentucky.
- Planifier les collectes de fonds, manifestations de sensibilisation et campagnes de signature des pétitions pour la campagne GROW d'Oxfam.

**Société pour la promotion de la recherche de premier cycle**

2014 - 2017

- Organisation basée à l'Université du Kentucky fondée pour aider les étudiants de 1<sup>er</sup> cycle à trouver les opportunités de recherche et à améliorer leurs compétences en recherche.
- *Coprésident* et ancien *Directeur des événements*.
- Planifier and exécuter la Showcase of Undergraduate Scholars.
- Planifier les ateliers de recherche pour aider les étudiants à améliorer leurs résumés, présentations et documents académiques.
- Organiser les tables rondes de faculté, événements promotionnels et réunions d'information sur s'impliquant dans la recherche, les bourse d'études et les opportunités de recherche d'été.
- Fusionné avec l'Équipe de sensibilisation des étudiants en mai 2017 (rejoint dès la fusion).

**COMPÉTENCES** Programmation:

- INFORMATIQUES**
- Languages: Python, C, C++, Java, HTML, CSS, Javascript, MATLAB, R, script shell UNIX.
  - Toolkits d'apprentissage automatique: TensorFlow, Sockeye, Keras, PyTorch, OpenNMT.
  - Contrôle de version: DVCS (Git, Bitbucket).
  - Environnements de développement: Vim, Emacs, Eclipse.

## Outils linguistiques:

Praat, AntConc, QGIS, Audacity.

## Logiciel bureautique de montage et productivité:

- $\text{\LaTeX}$  ( $\text{\L\TeX}$ ,  $\text{\BIB\TeX}$ ).
- GIMP, Inkscape.
- Microsoft Office, G Suite, OpenOffice.org, LibreOffice

**LANGUES** Français (C1), anglais (langue maternelle), inuktitut (rudimentaire). Expérience en chatino et ouïghour par la recherche.

**INTÉRÊTS ET LOISIRS**

- Musique: guitare (classique, électrique), basse, banjo, synthétiseur, MaxMSP, Ableton Live.
- Athlétisme: randonnée d'hiver, danse disco et swing.
- Loisirs: apprentissage du langage naturel, lecture (philosophie, politique), exploration urbaine.
- Compétences culinaires: thaïlandais, français, allemand, sud des États-Unis.