

ABTEEN EBRAHIMI

abteen.ebrahimi@colorado.edu

abteen.github.io

Boulder, CO

EDUCATION

University of Colorado Boulder

August 2020 - Present

PhD in Computer Science

University of California, Santa Cruz

Sept. 2016 - June 2020

B.A. in Computer Science with Highest Honors

B.A. in Computational Mathematics with Highest Honors

University Honors: Magna cum laude, GPA: 3.88

RESEARCH AND WORK EXPERIENCE

University of Colorado Boulder

August 2020 - Present

Graduate Program Teaching Instructor - Natural Language Processing (NLP)

Graduate Teaching Assistant - Algorithms, NLP, Data Structures

Google Research

January 2024 - April 2024

Student Researcher – GARCON Team

Jelinek Summer Workshop on Speech and Language Technology (JSALT)

Summer 2023

Junior Researcher

Natural Language and Dialogue Systems Lab

June 2018 - June 2020

Undergraduate Research Assistant

PUBLICATIONS

- [1] **Abteen Ebrahimi** and Katharina von der Wense. 2024. **Zero-Shot vs. Translation-Based Cross-Lingual Transfer: The Case of Lexical Gaps**. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*, Mexico City, Mexico. Association for Computational Linguistics
- [2] **Abteen Ebrahimi**, Arya D. McCarthy, Arturo Oncevay, John E. Ortega, Luis Chiruzzo, Gustavo Giménez-Lugo, Rolando Coto-Solano, and Katharina Kann. 2023b. **Meeting the Needs of Low-Resource Languages: The Value of Automatic Alignments via Pretrained Models**. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, Dubrovnik, Croatia. Association for Computational Linguistics
- [3] **Abteen Ebrahimi**, Manuel Mager, Arturo Oncevay, Vishrav Chaudhary, Luis Chiruzzo, Angela Fan, John Ortega, Ricardo Ramos, Annette Rios, Ivan Vladimir Meza Ruiz, Gustavo Giménez-Lugo, Elisabeth Mager, Graham Neubig, Alexis Palmer, Rolando Coto-Solano, Thang Vu, and Katharina Kann. 2022a. **AmericasNLI: Evaluating Zero-shot Natural Language Understanding of Pretrained Multilingual Models in Truly Low-resource Languages**. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics

- [4] **Abteen Ebrahimi** and Katharina Kann. 2021. **How to Adapt Your Pretrained Multilingual Model to 1600 Languages**. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. Association for Computational Linguistics
- [5] Katharina Kann, **Abteen Ebrahimi**, Kristine Stenzel, and Alexis Palmer. 2022c. **Machine Translation Between High-resource Languages in a Language Documentation Setting**. In *Proceedings of the First Workshop on NLP Applications to Field Linguistics*, Gyeongju, Republic of Korea. Association for Computational Linguistics
- [6] Katharina Kann, **Abteen Ebrahimi**, Joewie Koh, Shiran Dudy, and Alessandro Roncone. 2022b. **Open-domain Dialogue Generation: What We Can Do, Cannot Do, and Should Do Next**. In *Proceedings of the 4th Workshop on NLP for Conversational AI*, Dublin, Ireland. Association for Computational Linguistics
- [7] Katharina Kann, **Abteen Ebrahimi**, Manuel Mager, Arturo Oncevay, John E. Ortega, Annette Rios, Angela Fan, Ximena Gutierrez-Vasques, Luis Chiruzzo, Gustavo A. Giménez-Lugo, Ricardo Ramos, Ivan Vladimir Meza Ruiz, Elisabeth Mager, Vishrav Chaudhary, Graham Neubig, Alexis Palmer, Rolando Coto-Solano, and Ngoc Thang Vu. 2022a. **AmericasNLI: Machine Translation and Natural Language Inference Systems for Indigenous Languages of the Americas**. *Frontiers in Artificial Intelligence*, 5
- [8] Shereen Oraby, Vrindavan Harrison, **Abteen Ebrahimi**, and Marilyn Walker. **Curate and Generate: A Corpus and Method for Joint Control of Semantics and Style in Neural NLG**. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics

NON-REFEREED PUBLICATIONS

- [1] Ona De Gibert, Robert Pugh, Ali Marashian, Raul Vazquez, **Abteen Ebrahimi**, Pavel Denisov, Enora Rice, Edward Gow-Smith, Juan Prieto, Melissa Robles, Rubén Manrique, Oscar Moreno, Angel Lino, Rolando Coto-Solano, Aldo Alvarez, Marvin Agüero-Torales, John E. Ortega, Luis Chiruzzo, Arturo Oncevay, Shruti Rijhwani, Katharina Von Der Wense, and Manuel Mager. 2025. **Findings of the AmericasNLP 2025 Shared Tasks on Machine Translation, Creation of Educational Material, and Translation Metrics for Indigenous Languages of the Americas**. In *Proceedings of the Fifth Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP)*, pages 134–152, Albuquerque, New Mexico. Association for Computational Linguistics
- [2] **Abteen Ebrahimi** and Kenneth Church. 2024. **Since the Scientific Literature Is Multilingual, Our Models Should Be Too**. *Preprint*.
- [3] Kenneth Church, Omar Alonso, Peter Vickers, Jiameng Sun, **Abteen Ebrahimi**, and Raman Chandrasekar. 2024. **Academic Article Recommendation Using Multiple Perspectives**. *Preprint*.
- [4] **Abteen Ebrahimi**, Ona de Gibert, Raul Vazquez, Rolando Coto-Solano, Pavel Denisov, Robert Pugh, Manuel Mager, Arturo Oncevay, Luis Chiruzzo, Katharina von der Wense, and Shruti Rijhwani. 2024. **Findings of the AmericasNLP 2024 Shared Task on Machine Translation into Indigenous Languages**. 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

- [5] Luis Chiruzzo, Pavel Denisov, Alejandro Molina-Villegas, Silvia Fernandez-Sabido, Rolando Coto-Solano, Marvin Agüero-Torales, Aldo Alvarez, Samuel Canul-Yah, Lorena Hau-Ucán, **Abteen Ebrahimi**, Robert Pugh, Arturo Oncevay, Shruti Rijhwani, Katharina von der Wense, and Manuel Mager. 2024. **Findings of the AmericasNLP 2024 Shared Task on the Creation of Educational Materials for Indigenous Languages**. 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
- [6] **Abteen Ebrahimi**, Manuel Mager, Shruti Rijhwani, Enora Rice, Arturo Oncevay, Claudia Baltazar, María Cortés, Cynthia Montaña, John E. Ortega, Rolando Coto-solano, Hilaria Cruz, Alexis Palmer, and Katharina Kann. 2023a. **Findings of the AmericasNLP 2023 Shared Task on Machine Translation into Indigenous Languages**. In *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP)*, Toronto, Canada. Association for Computational Linguistics
- [7] **Abteen Ebrahimi**, Manuel Mager, Adam Wiemerslage, Pavel Denisov, Arturo Oncevay, Danni Liu, Sai Koneru, Enes Yavuz Ugan, Zhaolin Li, Jan Niehues, Monica Romero, Ivan G Torre, Tanel Alumäe, Jiaming Kong, Sergey Polezhaev, Yury Belousov, Wei-Rui Chen, Peter Sullivan, Ife Adenigba, Bashar Talafha, Alcides Alcoba Inciarte, Muhammad Abdul-Mageed, Luis Chiruzzo, Rolando Coto-Solano, Hilaria Cruz, Sofía Flores-Solórzano, Aldo Andrés Alvarez López, Ivan Meza-Ruiz, John E. Ortega, Alexis Palmer, Rodolfo Joel Zevallos Salazar, Kristine Stenzel, Thang Vu, and Katharina Kann. 2022b. **Findings of the Second AmericasNLP Competition on Speech-to-Text Translation**. In *Proceedings of the NeurIPS 2022 Competitions Track*, Volume 220 of *Proceedings of Machine Learning Research*. PMLR
- [8] Manuel Mager, Arturo Oncevay, **Abteen Ebrahimi**, John Ortega, Annette Rios, Angela Fan, Ximena Gutierrez-Vasques, Luis Chiruzzo, Gustavo Giménez-Lugo, Ricardo Ramos, Ivan Vladimir Meza Ruiz, Rolando Coto-Solano, Alexis Palmer, Elisabeth Mager-Hois, Vishrav Chaudhary, Graham Neubig, and Ngoc Thang Vu. **Findings of the AmericasNLP 2021 Shared Task on Open Machine Translation for Indigenous Languages of the Americas**. In *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas*. Association for Computational Linguistics
- [9] Vrindavan Harrison, Juraj Juraska, Wen Cui, Lena Reed, Kevin K. Bowden, Jiaqi Wu, Brian Schwarzmann, **Abteen Ebrahimi**, Rishi Rajasekaran, Nikhil Varghese, Max Wechsler-Azen, Steve Whittaker, Jeffrey Flanigan, and Marilyn Walker. 2020. **Athena: Constructing Dialogues Dynamically with Discourse Constraints**. *Preprint*.

SERVICE

NAACL Student Research Workshop Chair	2025
<i>Student Research Workshop</i>	
Workshop and Shared Task Co-Organizer	2023-2025
<i>Third Workshop on NLP for Indigenous Languages of the Americas</i> , co-located with ACL	
<i>Fourth Workshop on NLP for Indigenous Languages of the Americas</i> , co-located with NAACL	
<i>Fifth Workshop on NLP for Indigenous Languages of the Americas</i> , co-located with NAACL	
Co-Organizer of the Second AmericasNLP Competition Speech-to-Text Translation for Indigenous Languages of the Americas	2022
<i>Co-located with NeurIPS 2022</i>	
Co-Organizer of the AmericasNLP Shared Task on Open Machine Translation	2021
<i>First Workshop on NLP for Indigenous Languages of the Americas</i> , co-located with NAACL	

AWARDS

Summer Research Fellowship - \$6000	<i>Summer 2025</i>
Outstanding Research Award - \$500	<i>AY 2022-2023</i>
Ralph J. Slutz Student Excellence Award - \$1k	<i>Fall 2021</i>
Departmental Publication Recognition - \$200	<i>Fall 2021</i>
Dean's Graduate Assistantship - \$11.6k	<i>AY 2020-2021</i>

SKILLS

Languages and Frameworks

- **Fluent:** Python, PyTorch, Huggingface, Farsi
- **Proficient:** Pandas, NumPy, Bash, L^AT_EX
- **Familiar:** Java, C, C++