Brian Christian DuSell

Contact Information	Email	brian.dusell@inf.ethz.ch
	Website	bdusell.com
	Google Scholar	https://scholar.google.com/citations?user=1AkLaFIAAAAJ
	GitHub	https://github.com/bdusell

Research Interests

Natural Language Processing, Artificial Intelligence, Formal Language Theory

Academic Appointments

ETH Zürich, Switzerland — Nov 2023 to present
Postdoc, Department of Computer Science
Supervisor: Ryan Cotterell

Education

University of Notre Dame, Notre Dame, IN — Aug 2016 to May 2023
M.S. and Ph.D., Computer Science
Dissertation: *Nondeterministic Stacks in Neural Networks*
Advisor: David Chiang

University of Notre Dame, Notre Dame, IN — Aug 2009 to May 2013
B.S., Computer Science, *magna cum laude*

Refereed Publications

Vsteinn Snbjarnarson, Anej Svete, Josef Valvoda, Reda Boumasmoud, **Brian DuSell**, Ryan Cotterell. Causally Evaluating the Learnability of Formal Language Tasks. In *Proc. ICML*. 2026.

Taiga Someya, Anej Svete, **Brian DuSell**, Timothy J. O'Donnell, Mario Giulianelli, and Ryan Cotterell. Information Locality as an Inductive Bias for Neural Language Models. In *Proc. ACL*. 2025.
Oral + Panel Discussion (awarded to 0.8% of accepted papers).

Tim Vieira, Ben LeBrun, Mario Giulianelli, Juan Luis Gastaldi, **Brian DuSell**, John Terilla, Timothy J. O'Donnell, Ryan Cotterell. From Language Models over Tokens to Language Models over Characters. In *Proc. ICML*. 2025.
Spotlight paper (awarded to 2.6% of submitted papers).

Tim Vieira, Tianyu Liu, Clemente Pasti, Yahya Emara, **Brian DuSell**, Benjamin LeBrun, Mario Giulianelli, Juan Luis Gastaldi, Timothy J. O'Donnell, and Ryan Cotterell. Language Models over Canonical Byte-Pair Encodings. In *Proc. ICML*. 2025.

Alexandra Butoi, Ghazal Khalighinejad, Anej Svete, Josef Valvoda, Ryan Cotterell, and **Brian DuSell**. Training Neural Networks as Recognizers of Formal Languages. In *Proc. ICLR*. 2025.

Juan Luis Gastaldi, John Terilla, Luca Malagutti, **Brian DuSell**, Tim Vieira, and Ryan Cotterell. The Foundations of Tokenization: Statistical and Computational Concerns. In *Proc. ICLR*. 2025.

Mario Giulianelli, Luca Malagutti, Juan Luis Gastaldi, **Brian DuSell**, Tim Vieira, and Ryan Cotterell. On the Proper Treatment of Tokenization in Psycholinguistics. In *Proc. EMNLP*. 2024.

Stephen Bothwell, **Brian DuSell**, David Chiang, and Brian Krostenko. PILA: A Historical-Linguistic Dataset of Proto-Italic and Latin. In *Proc. LREC–COLING*. 2024.

Brian DuSell and David Chiang. Stack Attention: Improving the Ability of Transformers to Model Hierarchical Patterns. In *Proc. ICLR*. 2024.
Spotlight paper (awarded to 5% of submitted papers).

Brian DuSell and David Chiang. The Surprising Computational Power of Nondeterministic Stack RNNs. In *Proc. ICLR*. 2023.

Alexandra Butoi, **Brian DuSell**, Tim Vieira, Ryan Cotterell, and David Chiang. Algorithms for Weighted Pushdown Automata. In *Proc. EMNLP*. 2022.

Brian DuSell and David Chiang. Learning Hierarchical Structures with Differentiable Nondeterministic Stacks. In *Proc. ICLR*. 2022.
Spotlight paper (awarded to 5% of submitted papers).

Brian DuSell and David Chiang. Learning Context-Free Languages with Nondeterministic Stack RNNs. In *Proc. CoNLL*. 2020.
Acceptance rate: 23%.

Kenton Murray, **Brian DuSell**, and David Chiang. Efficiency through Auto-Sizing: Notre Dame NLP's Submission to the WNGT 2019 Efficiency Task. In *Proc. Workshop on Neural Generation and Translation*. 2019.

Preprints

Brian DuSell and Ryan Cotterell. Bearing Syntactic Fruit with Stack-Augmented Neural Networks. 2025.

Invited Talks

- "Training Neural Networks to Recognize Formal Languages: Bridging Theory and Experiments" — Sep 2025
  New York University
- "Canonical Tokenizations" — Feb 2025
  GenLM
- "Training Neural Networks as Recognizers of Formal Languages" — Jan 2025
  Saarland University
- "Stack Attention: Improving the Ability of Transformers to Model Hierarchical Patterns" — Apr 2024
  Seminars on Formal Languages and Neural Networks
  Recording: `https://www.youtube.com/watch?v=NrKLnGfEeeg`
- "Stack Attention: Improving the Ability of Transformers to Model Hierarchical Patterns" — Feb 2024
  ZurichNLP Meetup
  URL: `https://zurich-nlp.ch/event/zurichnlp-meetup-8/`
- "Nondeterministic Stacks in Neural Networks" — Oct 2022
  Seminars on Formal Languages and Neural Networks
  Recording: `https://www.youtube.com/watch?v=tkj6E9_n82U`
- "Stack Nondeterminism in Neural Networks" — Nov 2021
  Notre Dame NL+
  URL: `https://nlp.nd.edu/nlplus/2021/11/10/dusell.html`
- "How to Install Literally Anything: A Practical Guide to Singularity." — May 2019
  XSEDE Campus Champions Tech Talk
  Recording: `https://www.youtube.com/watch?v=D5pe4ewtDe8`

Research Positions

Postdoctoral Research Associate — Jun 2023
University of Notre Dame
Department of Computer Science and Engineering
Natural Language Processing Group
Supervisor: David Chiang

Applied Scientist Intern Jun to Sep 2021
Amazon Web Services
Team: Amazon Translate
Manager: Georgiana Dinu
Mentors: Xing Niu and Anna Currey

Applied Scientist Intern Jun to Sep 2020
Amazon Web Services
Team: Amazon Translate
Manager: Georgiana Dinu
Mentors: Xing Niu and Greg Hanneman

Research Assistant Sep 2011 to May 2013
University of Notre Dame
Department of Computer Science and Engineering
Cooperative Computing Lab
Supervisor: Douglas Thain
Summary: Contributed to BioCompute, a distributed computing environment for bioinformatics accessible via a web interface.

Teaching Experience

Instructor of Record Spring 2022
CSE 30151: Theory of Computing
University of Notre Dame
Department of Computer Science and Engineering
Designed and led an in-person course in formal language and complexity theory for a class of over 80 undergraduate students. I received a median course instructor feedback score of 4, on a scale from 0 to 5, which is the unofficial threshold for excellent teaching used in making tenure decisions at Notre Dame.

Teaching Assistant Fall 2018
CSE 40657/60657: Natural Language Processing
University of Notre Dame
Department of Computer Science and Engineering
Instructor: David Chiang

Teaching Assistant Spring 2017
CSE 30151: Theory of Computing
University of Notre Dame
Department of Computer Science and Engineering
Instructor: David Chiang
Received the department's Outstanding Graduate TA award.

Teaching Assistant Fall 2016
CSE 30151: Theory of Computing
University of Notre Dame
Department of Computer Science and Engineering
Instructor: Peter Kogge
Received an Honorable Mention for the department's Outstanding Graduate TA award.

Teaching Assistant Fall 2012
CSE 30331: Data Structures
University of Notre Dame
Department of Computer Science and Engineering
Instructors: Paul Brenner and Raul Santelices

Tutor Spring and Fall 2012
University of Notre Dame
Academic Services for Student-Athletes
Tutored a student-athlete for the courses CSE 30151: Theory of Computing and CSE 40113: Design and Analysis of Algorithms.

Mentoring

- Melisa Sude Gonca, undergraduate student at ETH Zürich (2026). Supervising bachelor's thesis on NP-hardness proofs for transformer language models.
- Yahya Emara, master's student at ETH Zürich (2025 to present). Supervising master's thesis project on character-level generation from token-level language models.
- Taiga Someya, PhD student at University of Tokyo (2024 to 2025). Informally co-supervised project on inductive biases of neural networks and hosted as visitor at ETH.
- Pascal Dominic Müller, master's student at ETH Zürich (2024 to present). Co-supervised term project on semiring-general backpropagation; now supervising master's thesis.
- Octave Arevian, master's student at ETH Zürich (2024 to 2025). Supervised master's thesis project on sample efficiency of stack-augmented neural networks.

Professional Service

- Area Chair: EMNLP 2025, ACL 2025
- Reviewer: ICML 2026 (**silver reviewer**), TACL 2026, ICLR 2026, ICML 2025, ICLR 2025, ACL 2024, ICML 2024, EMNLP 2023, NeurIPS 2023 (**top reviewer**), ACL 2023, EMNLP 2022, EMNLP 2021
- Organizing Committee for Midwest Speech and Language Days May 2018

Awards

- Top Reviewer for NeurIPS 2023
- First Place, Chinese Speech Contest (2nd Year Chinese) Apr 2019
  University of Notre Dame Department of East Asian Languages
- Outstanding Graduate Teaching Assistant May 2018
  Department of Computer Science and Engineering
  University of Notre Dame
- Honorable Mention, Outstanding Graduate Teaching Assistant May 2017
  Department of Computer Science and Engineering
  University of Notre Dame
- Notebaert Premier Fellowship 2016
  University of Notre Dame Graduate School
  The most prestigious fellowship for graduate students at Notre Dame.
- B.S. *magna cum laude*, University of Notre Dame 2013
- Member, Tau Beta Pi Engineering Honor Society 2012
- Member, Upsilon Pi Epsilon Computing Honor Society 2012
- College of Engineering Dean's List Fall 2010 to Spring 2013
  University of Notre Dame

Campus Service

- Mentor, Graduate Resilience Alliance at Notre Dame Jan to Apr 2023
  Mentored a group of first-year Notre Dame graduate students.
- Graduate Orientation Ambassador, University of Notre Dame 2017 to 2019
  Organized the fall orientation program for incoming graduate students.

- Graduate Representative, University of Notre Dame 2019
  Recruited fellowship awardees for the Graduate School.
- CSE Peer Mentor, University of Notre Dame 2018 to 2019
  Mentored first-year graduate students in the Computer Science and Engineering Department.

Industry Experience

Software Developer May 2014 to Aug 2016
Oak Financial Software Corp
Developed Chapulín, a hybrid mobile and web application for executing international money transfers to Latin America. Implemented frontend and contributed to backend functionality, tools for analytics, and test automation. Technologies used: JavaScript, Cordova, Python, Node.js.

Member of Technical Staff Jul 2013 to May 2014
NetApp, Inc.
Performed quality assurance for data replication software included in the Data ONTAP storage OS. Technologies used: Perl, Jenkins.

Software Engineer, Intern May to Aug 2012
Wolverine Trading, LLC
Developed a high-performance Syslog daemon with a configurable message handling system and real-time GUI client. Achieved 300-fold improvement in message processing rate over previous tool. My code was deployed to 80 production servers within the next two months. Technologies used: C++, C#, WPF, XAML.

Programming Skills

Proficient in Python, PyTorch, Bash scripting, Docker, JavaScript/Node.js, frontend/backend web development.
Very familiar with C, C++, Java, PHP, SQL.

Software

Rau (https://github.com/bdusell/rau)
Python package and command-line tool for training and evaluating language models and sequence-to-sequence models using PyTorch.

Semiring Einsum (https://bdusell.github.io/semiring-einsum/)
Python package for performing efficient einsum operations (a generalization of matrix multiplication) in various semirings in PyTorch.

Neural Network Recognizers (https://github.com/rycolab/neural-network-recognizers)
PyTorch code for training RNNs, LSTMs, and transformers as recognizers of formal languages. Supports multi-task learning and efficient dataset generation.

Stack Attention (https://github.com/bdusell/stack-attention)
PyTorch implementation of transformers with stack attention, including a full machine translation pipeline.

Nondeterministic Stack RNN (https://github.com/bdusell/nondeterministic-stack-rnn)
PyTorch implementation of my Nondeterministic Stack RNN model, as well as other Stack RNN models.

QFunnel (https://github.com/bdusell/qfunnel)
Command-line tool for efficiently queueing large numbers of experiments on Notre Dame's research computing cluster.

dockerdev (https://github.com/bdusell/dockerdev)
Shell scripts for easily managing development environments in Docker containers.

rougescore (https://github.com/bdusell/rougescore)
Python implementation of the ROUGE metric.

Jishosen (jishosen.com)
A Japanese-English dictionary website based on freely available data.

pycfg (https://github.com/bdusell/pycfg)
: Implementation of several context-free grammar algorithms, including Tomita's GLR parsing algorithm.

romaji-cpp (https://github.com/bdusell/romaji-cpp)
: C++ library for transliterating Japanese phonetic characters to Latin letters.

jgreek (https://github.com/bdusell/jgreek)
: Java library for dealing with orthography in ancient Greek.

xlator (https://github.com/bdusell/xlator)
: Programmable machine translation system based on syntactic transfer.