

Brian Christian DuSell

CONTACT INFORMATION	Email	bdusell@ttic.edu
	Website	bdusell.com
	Google Scholar	https://scholar.google.com/citations?user=1AkLaFIAAAAJ
	GitHub	https://github.com/bdusell
RESEARCH INTERESTS	Artificial Intelligence, Natural Language Processing, Formal Language Theory	
ACADEMIC APPOINTMENTS	Toyota Technological Institute at Chicago , Chicago, IL	Oct 2026 to present
	Research Assistant Professor	
	ETH Zürich , Switzerland	Nov 2023 to Aug 2026
	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: <i>Nondeterministic Stacks in Neural Networks</i> Advisor: David Chiang	
	University of Notre Dame , Notre Dame, IN	Aug 2009 to May 2013
	B.S., Computer Science, <i>magna cum laude</i>	
REFEREED PUBLICATIONS	Vsteinn Snbjarnarson, Anej Svete, Josef Valvoda, Reda Boumasmoud, Brian DuSell , Ryan Cotterell. Causally Evaluating the Learnability of Formal Language Tasks . In <i>Proc. ICML</i> . 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 <i>Proc. ACL</i> . 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 <i>Proc. ICML</i> . 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 <i>Proc. ICML</i> . 2025.	
	Alexandra Butoi, Ghazal Khalighinejad, Anej Svete, Josef Valvoda, Ryan Cotterell, and Brian DuSell . Training Neural Networks as Recognizers of Formal Languages . In <i>Proc. ICLR</i> . 2025.	
	Juan Luis Gastaldi, John Terilla, Luca Malagutti, Brian DuSell , Tim Vieira, and Ryan Cotterell. The Foundations of Tokenization: Statistical and Computational Concerns . In <i>Proc. ICLR</i> . 2025.	
	Mario Giulianelli, Luca Malagutti, Juan Luis Gastaldi, Brian DuSell , Tim Vieira, and Ryan Cotterell. On the Proper Treatment of Tokenization in Psycholinguistics . In <i>Proc. EMNLP</i> . 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”
New York University Sep 2025
- “Canonical Tokenizations”
[GenLM](#) Feb 2025
- “Training Neural Networks as Recognizers of Formal Languages”
Saarland University Jan 2025
- “Stack Attention: Improving the Ability of Transformers to Model Hierarchical Patterns”
[Seminars on Formal Languages and Neural Networks](#)
Recording: <https://www.youtube.com/watch?v=NrKLnGfEeeg> Apr 2024
- “Stack Attention: Improving the Ability of Transformers to Model Hierarchical Patterns”
[ZurichNLP Meetup](#)
URL: <https://zurich-nlp.ch/event/zurichnlp-meetup-8/> Feb 2024
- “Nondeterministic Stacks in Neural Networks”
[Seminars on Formal Languages and Neural Networks](#)
Recording: https://www.youtube.com/watch?v=tkj6E9_n82U Oct 2022
- “Stack Nondeterminism in Neural Networks”
[Notre Dame NL+](#)
URL: <https://nlp.nd.edu/nlplus/2021/11/10/dusell.html> Nov 2021
- “How to Install Literally Anything: A Practical Guide to Singularity.”
[XSEDE Campus Champions Tech Talk](#)
Recording: <https://www.youtube.com/watch?v=D5pe4ewtDe8> May 2019

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 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.	Spring and Fall 2012
MENTORING	• Melisa Sude Gonca, undergraduate student at ETH Zürich (2026). Supervised bachelor's thesis on NP-hardness proofs for transformer language models. • Pascal Dominic Müller, master's student at ETH Zürich (2024 to 2026). Co-supervised term project and master's thesis on semiring-general backpropagation. • Yahya Emara, master's student at ETH Zürich (2025). Supervised 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. • 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: ARR August 2026, EMNLP 2026, 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 • First Place, Chinese Speech Contest (2nd Year Chinese) • Outstanding Graduate Teaching Assistant • Honorable Mention, Outstanding Graduate Teaching Assistant • Notebaert Premier Fellowship • B.S. <i>magna cum laude</i>, University of Notre Dame • Member, Tau Beta Pi Engineering Honor Society • Member, Upsilon Pi Epsilon Computing Honor Society • College of Engineering Dean's List	2023 Apr 2019 May 2018 May 2017 2016 2013 2012 2012 Fall 2010 to Spring 2013
CAMPUS SERVICE	• Mentor, Graduate Resilience Alliance at Notre Dame • Graduate Orientation Ambassador, University of Notre Dame • Graduate Representative, University of Notre Dame • CSE Peer Mentor, University of Notre Dame	Jan to Apr 2023 2017 to 2019 2019 2018 to 2019

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.