Dr. Jonathan Dunn

jedunn@illinois.edu

Associate Professor
Department of Linguistics
University of Illinois, Urbana-Champaign

Papers: jdunn.name
Data: earthLings.io
Code: github.com/jonathandunn

Specialization

Computational Linguistics; Cognitive Linguistics; Computational Syntax; Computational Sociolinguistics

Overview

I am a computational linguist working to model both (i) the emergence of grammar within individuals and (ii) variation in grammar across individuals, populations, and registers. To support this research, I've developed large multi-lingual geographic corpora. My work has also examined the impact of linguistic variation on LLMs and the impact of LLMs on linguistic variation.

Books and Monographs

Dunn, J. (Under Contract). *The Computational Renaissance: Language Models in the History of Linguistics*. MIT Press.

Dunn, J. (2026). *Syntactic Variation from Individuals to Populations: Language as a Complex System*. Elements in Construction Grammar. Cambridge University Press.

Dunn, J. (2024). *Computational Construction Grammar: A Usage-Based Approach*. Elements in Cognitive Linguistics. Cambridge University Press. Interactive Labs.

Dunn, J. (2022). *Natural Language Processing for Corpus Linguistics*. Elements in Corpus Linguistics. Cambridge University Press. Interactive Labs
Reviews: *Natural Language Engineering*, *Corpus Pragmatics*, *Int'l Journal of Corpus Linguistics*

Appointments held

2024–	Associate Professor in Linguistics, University of Illinois Urbana-Champaign
2023-2023	Senior Lecturer Above the Bar in Linguistics, University of Canterbury (NZ)
2021-2022	Senior Lecturer in Linguistics, University of Canterbury (NZ)
2018-2020	Lecturer in Linguistics, University of Canterbury (NZ)
2015-2018	Research Assistant Professor in Computer Science, Illinois Institute of Technology
2015-2018	Visiting Scientist, National Geospatial-Intelligence Agency
2014-2015	Post-Doc in Computer Science, Illinois Institute of Technology

Education

2013	PhD, Purdue University Dissertation: *Automatic Identification of Metaphoric Utterances* Advisor: Victor Raskin
2010	MA, Purdue University Thesis: *Towards a Computational Model of Metaphor* Advisor: Victor Raskin
2008	BA, Classics, Hillsdale College

Leadership, Fellowships, and Visiting Positions

2024- Director of Undergraduate Studies, Department of Linguistics, University of Illinois
2026-2027 OpEd Project Public Voices Fellow
2025 Faculty, LSA Linguistic Institute, University of Oregon
2021-2023 Theme Leader at the New Zealand Institute for Language, Brain and Behaviour
2022 Visiting Scholar at the Center for Spatial Data Science, University of Chicago
2015 Faculty, LSA Linguistic Institute, University of Chicago

Major Funding

2024-2027 Marsden Fund Standard Grant (NZL)
Does Machine-Assisted Writing Erase Linguistic Diversity?
Co-PI: $660,000 (NZD)

2020-2022 Science for Technological Innovation, National Science Challenges, MBIE (NZL)
Domain Adaptation to Support Polynesian Language Technology
PI: $200,000 (NZD)

2015-2018 Visiting Scientist Research Fellowship, Oak Ridge Institute for Science and Education
Geolinguistics
PI: Approx. $500,000 (USD)

Awards and Other Funding

2026-2029 Conrad Professorial Scholar
College of Liberal Arts and Sciences,
University of Illinois Urbana-Champaign

2025 BRIDGE Seed Fund
University of Illinois and University of Birmingham
Co-PI, Prof. Jack Grieve, Birmingham
Award: $6,500 USD and £6,500 GBP

2023 Early Career Research Award in Humanities, Social Sciences or Creative Arts
Faculty of Arts, University of Canterbury
Award: $2,000 (NZD)

2021 Early Career Researcher Development Fund, University of Canterbury
Towards Equitable Language Models
Grant: $5,000 (NZD)

2020 Teaching Development Grant, University of Canterbury
Syntactic Annotations for Māori Corpora
Grant with Jeanette King: $5,000 (NZD)

2019 Support Grant, University of Canterbury
Automating Tourist Profiles By Integrating Spatial and Textual Artificial Intelligence
Grant with Ben Adams and Girish Prayag: $5,000

2017 Certificate of Merit, Oak Ridge Institute for Science and Education
Award: $3,000 (USD)

2014-2015 IC Postdoctoral Research Fellowship, Office of the Director of National Intelligence
Computational Cognitive Stylistics For Multi-Modal Identity Analytics

2013 Bilsland Dissertation Fellowship, Purdue University Graduate School
Automated Identification of Metaphoric Utterances

2011-2012 Bilsland Strategic Initiatives Grant, Purdue University Graduate School
Evaluating Cross-Listed Graduate Courses

MOOCs on edX

Now over 14,000 students

Text Analytics 1: Introducing Natural Language Processing
Text Analytics 2: Visualizing Natural Language Processing

Papers

My research is situated at intersection between linguistics and natural language processing. Thus, I have published widely in both fields. Here I am distinguishing between NLP proceedings and linguistics journals for convenience. Both represent peer-reviewed publications.

In Press

1. **Dunn, J.** (in press). "Cognition and Computation." In the *Handbook of Cognitive Linguistics*. Bloomsbury. Link.
2. **Dunn, J.** & Edia, M. (in press). "Data-Driven Dialectology and Regional Variation." In the *Research Handbook on Sociolinguistics*. Edward Elgar Press. Link.

2026

3. Morin, C.; Coats, S.; & **Dunn, J.** (2026). "How register and region shape the language network: Evidence from Computational Construction Grammar." *Constructions*, 18(1). DOI: 10.24338/cons-755.
4. Gan, Q.; **Dunn, J.**; Nini, A.; & Adams, B. (2026). "A Multi-Dialectal, Longitudinal Corpus of Human-AI Hybrid Language Production." In *Proceedings of the International Conference on Language Resources and Evaluation*. ERLA. Link.
5. Eida, M.; Dolan, R.; de Nijs, P.; & **Dunn, J.** (2026). "Some Dialects Are More Equal Than Others: Non-Prestigious Arabic Dialectal Bias in LLMs." In *Proceedings of The Fourth Arabic Natural Language Processing Conference*. Association for Computational Linguistics. Link.
6. Martins, M. & **Dunn, J.** (2026). "Where Are Language-Specific Features Represented in Multilingual LLMs? Everywhere." In *Proceedings of 6th Multilingual Representation Learning Workshop*. Association for Computational Linguistics. Link.
7. **Dunn, J.** (2026). "Association Measures." In the *International Encyclopedia of Language and Linguistics*, 3rd Edition. Elsevier. DOI: 10.1016/B978-0-323-95504-1.01533-7
8. **Dunn, J.** (2026). "Diffusion Across the Grammar: Complexity in Areal Interactions Between Dialects of English" in Enrique-Arias, A., C. De Benito Moreno and F. Del Barrio De La Rosa (eds.) *The Spatial Diffusion of Linguistic Changes: New Methods and Theoretical Perspectives.* Berlin: De Gruyter. Studies in Language Change 26. Link.

2025

9. **Dunn, J.** & Wong, S. (2025). "Language Contact and Population Contact As Sources of Dialect Similarity." *Languages*, 10(8), 188. DOI: 10.3390/languages10080188.
10. Anand, G., & **Dunn, J.** (2025). "Mapping Semantic Domains Across India's Social Media: Networks, Geography, and Social Factors." In *Proceedings of the International Conference on Computational Semantics*. Association for Computational Linguistics. 307-320. Link.
11. **Dunn, J.** & Eida, M. (2025). "LLMs Learn Constructions That Humans Do Not Know." In *Proceedings of the Second International Workshop on Construction Grammars and NLP*. Association for Computational Linguistics. 13-23. Link.

2024

12. **Dunn, J.**; Adams, B.; & Tayyar Madabushi, H. (2024). "Pre-Trained Language Models Represent Some Geographic Populations Better Than Others." In *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, LREC/COLING. 12966–12976. Link.
13. **Dunn, J.** & Edwards-Brown, L. (2024). "Geographically-Informed Language Identification." In *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, LREC/COLING. 7672–7682. Link.

14. **Dunn, J.** (2024). "Validating and Exploring Large Geographic Corpora." In *Proceedings of the Joint International Conference on Computational Linguistics, Language Resources and Evaluation*, LREC/COLING. 17348–17358. Link.

15. Eida, M.; Nassar, M.; & **Dunn, J.** (2024). "How Well Do Tweets Represent Sub-Dialects of Egyptian Arabic?" In *Proceedings of the Eleventh Workshop on NLP for Similar Languages, Varieties, and Dialects*. Association for Computational Linguistics. 41-55. Link.

2023

16. Li, H.; **Dunn, J.***; & Nini, A. (2023). "Register Variation Remains Stable Across 60 Languages." *Corpus Linguistics and Linguistic Theory*, 19(3): 397-426. Link. **Corresponding author*.

17. **Dunn, J.** (2023). "Syntactic variation across the grammar: Modelling a complex adaptive system." *Frontiers in Complex Systems*. DOI: 10.3389/fcpxs.2023.1273741.

18. **Dunn, J.** (2023). "Exploring the Constructicon: Linguistic Analysis of a Computational CxG." In *Proceedings of the First International Workshop on Construction Grammars and NLP*. Association for Computational Linguistics. Link.

19. **Dunn, J.** (2023). "Variation and Instability in Dialect-Based Embedding Spaces." In *Proceedings of the Workshop on NLP for Similar Languages, Varieties and Dialects*. Association for Computational Linguistics. Link.

20. Wong, S. G.-J., Durward, M., Adams, B., & **Dunn, J.** (2023). "Cantnlp@ LT-EDI-2023: Homophobia/transphobia detection in social media comments using spatio-temporally retrained language models." In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*. Link.

2022

21. **Dunn, J.** (2022). "Exposure and Emergence in Usage-Based Grammar: Computational Experiments in 35 Languages." *Cognitive Linguistics*, 33(4): 659-699. Data. Link.

22. **Dunn, J.** & Wong, S. G.-J. (2022). "Stability of Syntactic Dialect Classification Over Space and Time." In *Proceedings of the International Conference on Computational Linguistics*, COLING. 26-36. Link.

23. **Dunn, J.**; Li, H.; & Sastre, D. (2022). "Predicting Embedding Reliability in Low-Resource Settings Using Corpus Similarity Measures." In *Proceedings of the International Conference on Language Resources and Evaluation*, LREC. European Language Resources Association. 6461-6470. Link.

24. **Dunn, J.** & Nijhof, W. (2022). "Language Identification for Austronesian Languages." In *Proceedings of the International Conference on Language Resources and Evaluation*, LREC. European Language Resources Association. 6530-6539. Link.

25. Li, H. & **Dunn, J.*** (2022). "Corpus Similarity Measures Remain Robust Across Diverse Languages." *Lingua*, 275: 103377. Link. * *Corresponding author*.

26. **Dunn, J.** (2022). "Cognitive Linguistics Meets Computational Linguistics: Construction Grammar, Dialectology, and Linguistic Diversity." In Tay, D. & Xie Pan, M. (eds.), *Data Analytics in Cognitive Linguistics: Methods and Insights*. 273-308. Berlin: De Gruyter. Link

2021

27. **Dunn, J.** & Tayyar Madabushi, H. (2021). "Learned Construction Grammars Converge Across Registers Given Increased Exposure." In *Proceedings of the Conference on Computational Natural Language Learning*, CoNLL. Association for Computational Linguistics. 268-278. Link.

28. **Dunn, J.** (2021). "Representations of Language Varieties Are Reliable Given Corpus Similarity Measures." In *Proceedings of the Workshop on NLP for Similar Languages, Varieties and Dialects* @ EACL. Association for Computational Linguistics. 28-38. Link.

29. **Dunn, J.** & Nini, A. (2021). "Production vs Perception: The Role of Individuality in Usage-Based Grammar Induction." In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics* @ NAACL. Association for Computational Linguistics. 149-159. Link.

2020

30. **Dunn, J.** (2020). "Mapping Languages: The Corpus of Global Language Use." *Language Resources and Evaluation*, 54(4): 999-1,018. Link.

31. **Dunn, J.** & Adams, B. (2020). "Geographically-Balanced Gigaword Corpora for 50 Language Varieties." In *Proceedings of the Conference on Language Resources and Evaluation*, LREC. European Language Resources Association. 2,521-2,529. Link.

32. **Dunn, J.**; Coupe, T.; & Adams, B. (2020). "Measuring Linguistic Diversity During COVID-19." In *Proceedings of the Workshop on Natural Language Processing and Computational Social Science* @ EMNLP. Association for Computational Linguistics. 1-10. Link.

33. **Dunn, J.** (2020). "Ontological and Grammatical Constraints on Metaphor Productivity." In Attardo, S. (ed.), *Script-Based Semantics: Foundations and Applications. Essays in Honor of Victor Raskin*. Berlin: De Gruyter. 55-76. Link

34. Mohammadhassan N., Mitrovic A., Neshatian K. & **Dunn J.** (2020). "Automatic Quality Assessment of Comments in Active Video Watching Using Machine Learning Techniques." In So H-J; Rodrigo M; Mason J; Mitrovic A (Eds). *Proceedings of the 28th International Conference on Computers in Education*. I: 1-10. Taiwan: Asia-Pacific Society for Computers in Education. Link

35. Mohammadhassan N., Mitrovic A., Neshatian K. & **Dunn J.** (2020) "Developing Personalized Nudges to Improve Quality of Comments in Active Video Watching." In *Proceedings of 28th International Conference on Computers in Education* 2: 766-769. Taiwan: Asia-Pacific Society for Computers in Education. Link

2019

36. **Dunn, J.** (2019). "Global Syntactic Variation in Seven Languages: Towards a Computational Dialectology." *Frontiers in Artificial Intelligence: Language and Computation.* Section on Computational Sociolinguistics. 10.3389/frai.2019.00015.

37. **Dunn, J.** (2019). "Modeling Global Syntactic Variation in English Using Dialect Classification." In *Proceedings of the Workshop on NLP for Similar Languages, Varieties and Dialects* @ NAACL. Association for Computational Linguistics. 42-53. Link.

38. **Dunn, J.** (2019). "Frequency vs. Association for Constraint Selection in Usage-Based Construction Grammar." In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics* @ NAACL. Association for Computational Linguistics. 117-128. Link.

39. **Dunn, J.** & Adams, B. (2019). "Mapping Languages and Demographics with Georeferenced Corpora." In *Proceedings of GeoComputation 19*. Link

2018

40. **Dunn, J.** (2018). "Multi-Unit Association Measures: Moving Beyond Pairs of Words." *International Journal of Corpus Linguistics*, 23(2): 183-215. Link

41. **Dunn, J.** (2018). "Finding Variants for Construction-Based Dialectometry: A Corpus-Based Approach to Regional CxGs." *Cognitive Linguistics*, 29(2): 275-311. Link.

42. **Dunn, J.** (2018). "Modeling the Complexity and Descriptive Adequacy of Construction Grammars." In *Proceedings of the Society for Computation in Linguistics*. Association for Computational Linguistics. 81-90. Link.

2017

43. **Dunn, J.** (2017). "Computational Learning of Construction Grammars." *Language and Cognition*, 9(2): 254-292. Link.

2016

44. **Dunn, J.**; Argamon, S.; Rasooli, A.; & Kumar, G. (2016) "Profile-Based Authorship Analysis." *Digital Scholarship in the Humanities*, 31(4): 689-710. Link.

2015

45. **Dunn, J.** (2015) "Modeling Abstractness and Metaphoricity." *Metaphor & Symbol*, 30(4): 259-289. Link.

46. **Dunn, J.** (2015). "Three Types of Metaphoric Utterances That Can Synthesize Theories of Metaphor." *Metaphor & Symbol*, 30(1): 1-23. Link.

2014

47. **Dunn, J.** (2014). "Measuring Metaphoricity." In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, ACL. Association for Computational Linguistics. 745-751. Link.

48. **Dunn, J.** (2014). "Multi-Dimensional Abstractness in Cross-Domain Mappings." In *Proceedings of the Workshop on Metaphor in NLP*. Association for Computational Linguistics. 27-32. Link.

49. **Dunn, J.**; Beltran de Heredia, J.; Burke, M.; Gandy, L.; Kanareykin, S.; Kapah, O.; Taylor, M.; Hines, D.; Frieder, O.; Grossman, D.; Howard, N.; Koppel, M.; Morris, S.; Ortony, A.; & Argamon, S. (2014). "Language-Independent Ensemble Approaches to Metaphor Identification." In *Proceedings of the 28th Conference on Artificial Intelligence: Workshop on Cognitive Computing for Augmented Human Intelligence*. 6-12. Link.

2013

50. **Dunn, J.** (2013). "How Linguistic Structure Influences and Helps To Predict Metaphoric Meaning." *Cognitive Linguistics*, 24(1): 33-66. Link.

51. **Dunn, J.** (2013). "Evaluating the Premises and Results of Four Metaphor Identification Systems." In *Proceedings of the Conference on Intelligent Text Processing and Computational Linguistics, Vol. 1*. Heidelberg: Springer. 471-486. Link

52. **Dunn, J.** (2013). "What Metaphor Identification Systems Can Tell Us About Metaphor-in-Language." In *Proceedings of the Workshop on Metaphor in NLP* @ NAACL. Association for Computational Linguistics. 1-10. Link.

2011

53. **Dunn, J.** (2011). "Gradient Semantic Intuitions of Metaphoric Expressions." *Metaphor & Symbol*, 26(1): 53-67. Link.

Teaching

Computational Linguistics

Computational Corpus Linguistics (LING 413: Mixed-level)
University of Illinois Urbana-Champaign
Recent Teaching Evaluation: 4.89 (out of 5)

Computational Sociolinguistics (LING 442: Mixed-level)
University of Illinois Urbana-Champaign
Recent Teaching Evaluation: 4.83 (out of 5)

Computational Syntax (LING 444: Mixed-level)
University of Illinois Urbana-Champaign
Recent Teaching Evaluation: 4.75 (out of 5)

Capstone in Computational Linguistics (LING 453: Undergraduate)
University of Illinois Urbana-Champaign
Recent Teaching Evaluation: 4.75 (out of 5)

Natural Language Processing (COSC 441: Graduate-level)
University of Canterbury
Teaching Evaluation: 4.75 (out of 5)

Text Analytics (LING 223: Undergraduate-level)
University of Canterbury
Teaching Evaluation: 4.90 (out of 5)

Linguistics

Topics in Syntactic Theory (LING 306: Undergraduate-level)
University of Canterbury
Teaching Evaluation: 4.50 (out of 5)

Grammatical Structure (LING 217: Undergraduate-level)
University of Canterbury
Teaching Evaluation: 4.64 (out of 5)

Forensic Linguistics (LING 225: Undergraduate-level)
University of Canterbury
Teaching Evaluation: 4.70 (out of 5)

General Linguistics

English Structures (LING 400: Graduate-level Intro)
University of Canterbury
Teaching Evaluation: 4.90 (out of 5)

The English Language (LING 101: Undergraduate-level Intro)
University of Canterbury
Teaching Evaluation: 4.10 (out of 5)

Other Teaching

Computational Construction Grammar: LSA Linguistic Institute @ University of Oregon
Data-Driven Computational Pragmatics: LSA Linguistic Institute @ University of Chicago
Introduction to Linguistics (LING 227) @ Purdue University
Classroom Communication for International Graduate Students @ Purdue University
First Year Composition for International Students (ENGL 1011) @ Purdue University
First-Year Composition (ENGL 101) @ Purdue University
Residential Learning Community (ENGL 101) @ Purdue University

Teaching Awards and Recognition

2025 List of Teachers Ranked as Excellent, University of Illinois Urbana-Champaign
FLEX Award, University of Illinois Urbana-Champaign
2024 List of Teachers Ranked as Excellent, University of Illinois Urbana-Champaign
2012 Quintilian Award (Top 10% of Instructor Evaluations), Purdue University, Department of English
2012 Graduate Teacher Certificate, Center for Instructional Excellence, Purdue University

Supervisions, Post-Doctoral

2021-2022 Haipeng Li (Post-doc, New Zealand Institute for Language Brain and Behaviour)
2024-2026 Qiao Gan (Post-doc, New Zealand Institute for Language Brain and Behaviour)

Supervisions, Ph.D.

2024-2025 Brennan Dell (Ph.D., Linguistics, Illinois)
2021-2026 Sidney Wong (Ph.D., Linguistics, Canterbury; With Ben Adams)
2020-2022 Negar Mohammadhassan (Ph.D., Computer Science, Canterbury; Primary: Tanja Mitrovic)
2024-Present Mai Mohamed Eida (Ph.D. Student, Linguistics, Illinois)
2024-Present Sofya Styrina (Ph.D. Student, Linguistics, Illinois)
2024-Present Mikaela Martins (Ph.D. Student, Linguistics, Illinois)
2024-Present Gunjan Anand (Ph.D. Student, Linguistics, Illinois)
2021-Present Matthew Durward (Ph.D. Student, Linguistics, Canterbury; With Chris Thomson)

Supervisions, Other

2024 Zofia Graham (Summer Mentoring Program, Undergraduate-level, Illinois)
2023 Fernando Miguez (M.A. Student, Linguistics, Canterbury; With Dineke Schokkin)

Software and Data

earthLings: Computational Linguistic Atlas
A 440 billion word geo-referenced corpus from the web and social media
https://www.earthLings.io
https://www.github.com/jonathandunn/earthLings

C2xG: Computational Construction Grammar
Python package for learning, evaluating, and annotating CxGs
https://www.github.com/jonathandunn/c2xg

GeoLID: Identifying minority languages more accurately using geographic priors
Python package for state-of-the-art LID
https://www.github.com/jonathandunn/geoLid

pacificLID: Identifying languages and code-switching with a focus on Austronesian languages
Python Version: https://github.com/jonathandunn/pacific_CodeSwitch
fastText Version: https://jdunn.name/corpora/

text_analytics: Package for teaching computational linguistics in Python
https://www.github.com/jonathandunn/text_analytics

corpus_similarity: Package for measuring corpus similarity and homogeneity
https://www.github.com/jonathandunn/corpus_similarity

Common Crawl Corpus: Corpus building from the common crawl data
https://www.github.com/jonathandunn/common_crawl_corpus

Service to the Discipline

Editorial Work

2026 Area Chair, Conference on Computational Natural Language Learning (CoNLL)
2018-2024 Associate Editor, *Frontiers in Artificial Intelligence* (Language and Computation)

Reviewing, Books

Textbooks in Linguistics, Cambridge University Press
Linguistics and Language Education, Bloomsbury Academic
Synthesis Lectures in Human Language Technologies, Morgan Claypool

Reviewing, Proposals

Austrian Science Fund
Croatian Science Foundation

Endeavour Fund, MBIE, New Zealand
Swiss National Science Foundation
European Science Foundation, EU
Fund for Scientific Research (F.R.S.-FNRS), Belgium
Social Sciences and Humanities Research Council, Canada

Reviewing, Journals

Advances in Methods and Practices in Psychological Science
Cognitive Linguistics
Cognitive Science
Constructions
Corpus Linguistics and Linguistic Theory
IEEE Transactions on Audio, Speech and Language Processing
Journal of Artificial Intelligence Research
Journal of Cognitive Science
Language & Cognition
Language, Cognition, and Neuroscience
Language Resources and Evaluation
Language Variation and Change
Lingua
Linguistics Vanguard
Nature Communication
npj Heritage Science
Metaphor & Symbol
Psychological Review
Royal Society Open Science
Scientific Reports
The Computer Journal

Reviewing, Proceedings Articles

Asia-Pacific Association for Computational Linguistics (AACL)
Association for Computational Linguistics (ACL)
Association for Computational Linguistics Rolling Review (ARR)
Conference on Computational Natural Language Learning (CoNLL)
Empirical Methods in Natural Language Processing (EMNLP)
European Association for Computational Linguistics (EACL)
International Conference on Computational Linguistics (COLING)
International Conference on Language Resources and Evaluation (LREC)
International Joint Conference on Natural Language Processing (ACL-IJCNLP)
North American Association for Computational Linguistics (NAACL)
Joint Linguistic Annotation Workshop and Designing Meaning Representations
Workshop on Linguistic Annotation
Workshop on Figurative Language Processing
Workshop on Metaphor in NLP
Workshop on Natural Language Processing and Computational Social Science
Workshop on NLP for Similar Languages, Varieties and Dialects
International Workshop on Construction Grammars and NLP

Reviewing, Conferences

Linguistic Society of America Annual Meeting
Symposium on Corpus Approaches to Lexicogrammar
International Conference on Construction Grammar

University Service

Committees

2025-Present Graduate College Fellowship Board (Humanities and Arts), University of Illinois
2024-2025 Committee on Courses and Curricula, College of Liberal Arts and Sciences, University of Illinois
2024 Ad Hoc Hiring Committee, Department of Linguistics, University of Illinois
2023 Academic Programmes Committee, College of Arts, University of Canterbury
2023 Hiring Committee, Department of Linguistics, University of Canterbury
2023 Hiring Committee, NZILBB, University of Canterbury
2020-2021 Academic Programmes Committee, College of Arts, University of Canterbury
2018-2023 Academic Programmes Committee, College of Science, University of Canterbury
2020-2022 Steering Group for Bachelors of Data Science Degree, University of Canterbury
2019-2020 Research Committee, School of Language, Social & Political Sciences, University of Canterbury
2019 Hiring Committee, Department of Linguistics, University of Canterbury

Administrative Achievements

2026 Developed Linguistics + DS program at the University of Illinois
2025 Revised and updated CS+Linguistics program at the University of Illinois
2020 Developed and delivered one of the flagship MOOCs for the University of Canterbury
2020 Developed the *Computational Linguistics* undergraduate major, University of Canterbury
2020 Helped develop the undergraduate degree, *Bachelors of Data Science*, University of Canterbury

Partially Peer-Reviewed Publications

2018 1. Dunn, J. (2018). "Recursively Emerging Structure: A Discovery-Device CxG." In *Proceedings of the Chicago Linguistics Society*, 53: 71-86.

2017 2. Dunn, J. (2017). "Learnability and Falsifiability of Construction Grammars: A Learning-Based Approach." *Proceedings of the Linguistic Society of America Annual Meeting*, 2(1):1-15.

3. Dunn, J. (2017). *Automating Human Geography With Dialectology*. Technical Report for National Geospatial-Intelligence Agency. Washington, D.C.

2015 4. Dunn, J. (2015). "Review of Frames and Constructions in Metaphoric Language (Constructional Approaches to Language, 14)." *Cognitive Linguistics*, 26(2): 371-375.

5. Dunn, J. (2015). "Review of The Semantic Representation of Natural Language." *Studies in Language*, 39(2): 492-500.

2013 6. Dunn, J. (2013). "Review of Converging Evidence: Methodological and Theoretical Issues for Linguistic Research. (Human Cognitive Processing, 33)." *Cognitive Linguistics*, 24(4): 711-717.

Invited Talks (Since 2022)

2027
1. Illinois Language and Linguistics Society: "Can Language Models Change Language?"
2. Rochester Institute of Technology, AWARE-AI Symposium: "Next Word Prediction as a Sociolinguistic Task"

2026
3. Penn State University: "Linguistic Variation and Language Models: Causes and Consequences of Homogeneity."
4. University of Manchester: “Computational Linguistics Will Not Solve All Your Problems.”
5. University of Birmingham: "Towards a Geographic Corpora Consortium."

2025
6. University of Colorado, Boulder: Institute for Cognitive Science: "Computational Sociolinguistics: Population-Level Variation and Its Impact on Language Technology."
7. University of Colorado, Boulder: Department of Linguistics: "Language as a Complex System: Syntactic Variation from Individuals to Populations."

2024
8. University of Galway: "Linguistic Diversity in the Digital World: From Perception to Production."
9. University of Illinois, Chicago: "Impacts of Linguistic Diversity on Language Technology."

2023
10. "Emerging Structure in Computational Construction Grammar." Plenary at the Construction Grammars and NLP Workshop, Georgetown University (CxGs+NLP)
11. "Learnability and Variability in Computational Linguistics: Implications for Second Language Acquisition." Purdue University

2022
12. "Improving Corpus Resources for Low-Resource Languages." University of Galway
13. "Computational Linguistics for Socially-Responsible Technology." University of Wisconsin-Madison
14. "Linguistic Variation in NLP." Pacific Northwest National Lab

Last updated: November 18, 2026