Junyi Jessy Li

Department of Linguistics
The University of Texas at Austin

web: jessyli.com
email: jessy@austin.utexas.edu

Education

Ph.D.	Computer and Information Science, University of Pennsylvania, Philadelphia PA Dissertation: *From Discourse Structure to Text Specificity: Studies of Coherence Preferences*; Advisor: Ani Nenkova	2017
M.S.E.	Computer and Information Science, University of Pennsylvania, Philadelphia PA	2013
B.S.E.	Computer Science and Engineering, University of Michigan, Ann Arbor MI	2011
B.E.	Electrical and Computer Engineering, Shanghai Jiao Tong University, Shanghai, China	2011

Appointments

The University of Texas at Austin

Associate Professor, Department of Linguistics	2023 – present
Associate Professor by Courtesy, Department of Computer Science	2025 – present
Affiliated Faculty, Oden Institute for Computational Engineering and Sciences	2025 – present
Assistant Professor, Department of Linguistics	2017 – 2023

Publications

Referred Conference/Journal Articles

[1] Gauri Kambhatla, Sanjana Gautam, Angela Zhang, Alex Liu, Ravi Srinivasan, **Junyi Jessy Li**, and Matthew Lease. "Beyond Sociodemographic Prompting: Using Supervision to Align LLMs with Human Response Distributions." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2026, to appear.

[2] Kaijie Mo, Siddhartha Venkatayogi, Chantal Shaib, Ramez Kouzy, Wei Xu, Byron C. Wallace, and **Junyi Jessy Li**. "Decide less, communicate more: On the construct validity of end-to-end fact-checking in medicine." In: *Findings of the Association for Computational Linguistics: ACL 2026*. 2026, to appear.

[3] Sebastian Joseph, Lily Chen, Barry Wei, Michael Mackert, Iain J. Marshall, Paul Pu Liang, Ramez Kouzy, Byron C. Wallace, and **Junyi Jessy Li**. "Decide less, communicate more: On the construct validity of end-to-end fact-checking in medicine." In: *Findings of the Association for Computational Linguistics: ACL 2026*. 2026, to appear.

[4] Anshun Asher Zheng, **Junyi Jessy Li**, and David I Beaver. "Strategic Dialogue Assessment: The Crooked Path to Innocence." In: *Dialogue & Discourse* 17.1 (2026), pp. 1–53.

[5] Daniel Brubaker, William Berkeley Sheffield, **Junyi Jessy Li**, and Kanishka Misra. "WUGNECTIVES: Novel Entity Inferences of Language Models from Discourse Connectives." In: *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. 2026, pp. 6109–6127.

[6] Samer Alabed, Abigail Anderson, Ahmed Maiter, Anthony Hughes, Niamh McAnenly, Mahan Salehi, Michael Sharkey, Krit Dwivedi, Alireza Hokmabadi, Fares Alahdab, Mark Stevenson, Ning Ma, Robert Gaizauskas, Tim J Chico, Andy J Swift, **Junyi Jessy Li**, Jens Kleesiek, and Curtis Langlotz. "Large language models for simplifying radiology reports: a systematic review and meta-analysis of patient, public, and clinician evaluations." In: *The Lancet Digital Health* (2025). ISSN: 2589-7500.

[7] Sebastian Joseph, Syed Murtaza Husain, Stella S. R. Offner, Stéphanie Juneau, Paul Torrey, Adam S. Bolton, Juan P. Farias, Niall Gaffney, Greg Durrett, and **Junyi Jessy Li**. "AstroVisBench: A Code Benchmark for Scientific Computing and Visualization in Astronomy." In: *Proceedings of the Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS): Datasets and Benchmarks*. 2025.

[8] Ramya Namuduri, Yating Wu, Anshun Asher Zheng, Manya Wadhwa, Greg Durrett, and **Junyi Jessy Li**. "QUDsim: Quantifying Discourse Similarities in LLM-Generated Text." In: *Proceedings of the Conference on Language Modeling (COLM)*. 2025.

[9] Manya Wadhwa, Zayne Sprague, Chaitanya Malaviya, Philippe Laban, **Junyi Jessy Li**, and Greg Durrett. "EvalAgent: Discovering Implicit Evaluation Criteria from the Web." In: *Proceedings of the Conference on Language Modeling (COLM)*. 2025.

[10] Zichao Hu, **Junyi Jessy Li**, Arjun Guha, and Joydeep Biswas. "Robo-Instruct: Simulator-Augmented Instruction Alignment For Finetuning CodeLLMs." In: *Proceedings of the Conference on Language Modeling (COLM)*. 2025.

[11] Robin C. Hilsabeck, Jeffrey N. Keller, Maya L. Henry, **Junyi Jessy Li**, Lokesh Pugalenthi, Paul Toprac, Patrick Chang, Joshua Chang, Suzanne Schmitz, Avery Largent, Heather Foil, Robert Brouillette, Rosemary A. Lester-Smith, and Paul J. Rathouz. "Development and classification accuracy of an automated cognitive screening tool combining working memory and connected speech tasks for early detection of cognitive impairment in primary care." In: *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 11.3 (2025), e70145.

[12] Heather Dial, Lokesha S. Pugalenthi, G. Nike Gnanateja, **Junyi Jessy Li**, and Maya L. Henry. "Application of machine learning and temporal response function modeling of EEG data for differential diagnosis in primary progressive aphasia." In: *Scientific Reports* 15.1 (2025), p. 29539.

[13] William Berkeley Sheffield, Kanishka Misra, Valentina Pyatkin, Ashwini Deo, Kyle Mahowald, and **Junyi Jessy Li**. "Is It JUST Semantics? A Case Study of Discourse Particle Understanding in LLMs." In: *Findings of the Association for Computational Linguistics: ACL 2025*. 2025, pp. 21704–21715.

[14] Jan Trienes, Jörg Schlötterer, **Junyi Jessy Li**, and Christin Seifert. "Behavioral Analysis of Information Salience in Large Language Models." In: *Findings of the Association for Computational Linguistics: ACL 2025*. 2025, pp. 23428–23454.

[15] Hongli Zhan, Muneeza Azmat, Raya Horesh, **Junyi Jessy Li**, and Mikhail Yurochkin. "SPRI: Aligning Large Language Models with Context-Situated Principles." In: *Proceedings of the 42nd International Conference on Machine Learning (ICML)*. 2025, pp. 74370–74405.

[16] Roxanna Attar-Olyaee, Ramez Kouzy, Michael K. Rooney, Zakaria El-Kouzi, Karen E. Hoffman, Sherif M. Gadoue, Comron J. Hassanzadeh, **Junyi Jessy Li**, and Osama Mohamad. "Cancer Risk Concerns and Communication Gaps Regarding GLP-1 Medications." In: *JAMA Network Open* 8.7 (2025), e2521878–e2521878.

[17] Hye Sun Yun, Karen Y.C. Zhang, Ramez Kouzy, Iain J. Marshall, **Junyi Jessy Li**, and Byron C. Wallace. "Caught in the Web of Words: Do LLMs Fall for Spin in Medical Literature?" In: *Proceedings of the Sixth Conference on Health, Inference, and Learning (CHIL)*. Vol. 287. 2025, pp. 458–479.

[18] Linghan Zhong, Samuel Yuan, Jiyang Zhang, Yu Liu, Pengyu Nie, **Junyi Jessy Li**, and Milos Gligoric. "A Tool for Generating Exceptional Behavior Tests With Large Language Models." In: *Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering: formal tool demonstrations (FSE Demo)*. 2025, pp. 1193–1197.

[19] Jiyang Zhang, Yu Liu, Pengyu Nie, **Junyi Jessy Li**, and Milos Gligoric. "exLong: Generating Exceptional Behavior Tests with Large Language Models." In: *2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*. 2025, pp. 1462–1474.

[20] Shiyang Zhang, Zexi Zhou, Yee To Ng, Elizabeth Muñoz, **Junyi Jessy Li**, and Karen Fingerman. "Everyday Language and Cognitive Functioning in Late Life." In: *The Journals of Gerontology: Series B* 80.7 (May 2025), gbaf084. ISSN: 1758-5368.

[21] Yating Wu, Ritika Mangla, Alexandros G. Dimakis, Greg Durrett, and **Junyi Jessy Li**. "Which questions should I answer? Salience Prediction of Inquisitive Questions." In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. **Outstanding Paper Award**. 2024, pp. 19969–19987.

[22] Chantal Shaib, Yanai Elazar, **Junyi Jessy Li**, and Byron Wallace. "Detection and Measurement of Syntactic Templates in Generated Text." In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2024, pp. 6416–6431.

[23] Venkata Subrahmanyan Govindarajan, Matianyu Zang, David Beaver, Kyle Mahowald, and **Junyi Jessy Li**. "Do they mean 'us'? Interpreting Referring Expressions in Intergroup Bias." In: *Findings of the Association for Computational Linguistics: EMNLP 2024 (EMNLP Findings)*. 2024, pp. 9772–9785.

[24] Manya Wadhwa, Xinyu Zhao, **Junyi Jessy Li**, and Greg Durrett. "Learning to Refine with Fine-Grained Natural Language Feedback." In: *Findings of the Association for Computational Linguistics: EMNLP 2024 (EMNLP Findings)*. 2024, pp. 12281–12308.

[25] Manya Wadhwa, Jifan Chen, **Junyi Jessy Li**, and Greg Durrett. "Using Natural Language Explanations to Rescale Human Judgments." In: *Proceedings of the Conference on Language Modeling (COLM)*. 2024.

[26] Hongli Zhan, Allen Zheng, Yoon Kyung Lee, Jina Suh, **Junyi Jessy Li**, and Desmond C. Ong. "Large Language Models are Capable of Offering Cognitive Reappraisal, if Guided." In: *Proceedings of the Conference on Language Modeling (COLM)*. 2024.

[27] Sebastian Joseph, Lily Chen, Jan Trienes, Hannah Louisa Göke, Monika Coers, Wei Xu, Byron C Wallace, and **Junyi Jessy Li**. "FactPICO: Factuality Evaluation for Plain Language Summarization of Medical Evidence." In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*. 2024, pp. 8437–8464.

[28] Jan Trienes, Sebastian Joseph, Jörg Schlötterer, Christin Seifert, Kyle Lo, Wei Xu, Byron C. Wallace, and **Junyi Jessy Li**. "InfoLossQA: Characterizing and Recovering Information Loss in Text Simplification." In: *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*. 2024, pp. 4263–4294.

[29] Yoon Kyung Lee, Jina Suh, Hongli Zhan, **Junyi Jessy Li**, and Desmond C. Ong. "Large Language Models Produce Responses Perceived to be Empathic." In: *Proceedings of the 12th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 2024, To Appear.

[30] Smriti Singh, Cornelia Caragea, and **Junyi Jessy Li**. "Language Models (Mostly) Do Not Consider Emotion Triggers When Predicting Emotion." In: *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. 2024, pp. 603–614.

[31] Tiberiu Sosea, **Junyi Jessy Li**, and Cornelia Caragea. "Sarcasm Detection in a Disaster Context." In: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*. 2024, pp. 14313–14324.

[32] Yating Wu, Ritika Mangla, Greg Durrett, and **Junyi Jessy Li**. "QUDeval: The Evaluation of Questions Under Discussion Discourse Parsing." In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2023, pp. 5344–5363.

[33] Yating Wu, William Sheffield, Kyle Mahowald, and **Junyi Jessy Li**. "Elaborative Simplification as Implicit Questions Under Discussion." In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2023, pp. 5525–5537.

[34] Sebastian Joseph, Kathryn Kazanas, Keziah Reina, Vishnesh J. Ramanathan, Wei Xu, Byron Wallace, and **Junyi Jessy Li**. "Multilingual Simplification of Medical Texts." In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2023, pp. 16662–16692.

[35] Hongli Zhan, Desmond C. Ong, and **Junyi Jessy Li**. "Evaluating Subjective Cognitive Appraisals of Emotions from Large Language Models." In: *Findings of the Association for Computational Linguistics: EMNLP 2023 (EMNLP Findings)*. 2023, pp. 14418–14446.

[36] Jiyang Zhang, Pengyu Nie, **Junyi Jessy Li**, and Milos Gligoric. "Multilingual Code Co-Evolution Using Large Language Models." In: *Proceedings of the ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*. 2023, pp. 695–707.

[37] Tiberiu Sosea, Hongli Zhan, **Junyi Jessy Li**, and Cornelia Caragea. "Unsupervised Extractive Summarization of Emotion Triggers." In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*. 2023, pp. 9550–9569.

[38] Chantal Shaib, Millicent Li, Sebastian Joseph, Iain Marshall, **Junyi Jessy Li**, and Byron Wallace. "Summarizing, Simplifying, and Synthesizing Medical Evidence using GPT-3 (with Varying Success)." In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*. 2023, pp. 1387–1407.

[39] Wei-Jen Ko, Yating Wu, Cutter Dalton, Dananjay Srinivas, Greg Durrett, and **Junyi Jessy Li**. "Discourse Analysis via Questions and Answers: Parsing Dependency Structures of Questions Under Discussion." In: *Findings of the Association for Computational Linguistics: ACL 2023 (ACL Findings)*. 2023, pp. 11181–11195.

[40] Venkata Subrahmanyan Govindarajan, David Beaver, Kyle Mahowald, and **Junyi Jessy Li**. "Counterfactual Probing for the Influence of Affect and Specificity on Intergroup Bias." In: *Findings of the Association for Computational Linguistics: ACL 2023 (ACL Findings)*. 2023, pp. 12853–12862.

[41] Venkata Subrahmanyan Govindarajan, Katherine Atwell, Barea Sinno, Malihe Alikhani, David I. Beaver, and **Junyi Jessy Li**. "How people talk about each other: Modeling Generalized Intergroup Bias and Emotion." In: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. 2023, pp. 2496–2506.

[42] Pengyu Nie, Rahul Banerjee, **Junyi Jessy Li**, Raymond J. Mooney, and Milos Gligoric. "Learning Deep Semantics for Test Completion." In: *Proceedings of the 45th International Conference on Software Engineering (ICSE)*. 2023, pp. 2111–2123.

[43] Yaman Kumar, Swapnil Parekh, Somesh Singh, **Junyi Jessy Li**, Rajiv Ratn Shah, and Changyou Chen. "Automatic Essay Scoring Systems Are Both Overstable And Oversensitive: Explaining Why And Proposing Defenses." In: *Dialogue & Discourse* 14.1 (2023), pp. 1–33.

[44] Wei-Jen Ko, Cutter Dalton, Mark Simmons, Eliza Fisher, Greg Durrett, and **Junyi Jessy Li**. "Discourse Comprehension: A Question Answering Framework to Represent Sentence Connections." In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2022, pp. 11752–11764.

[45] Hongli Zhan, Tiberiu Sosea, Cornelia Caragea, and **Junyi Jessy Li**. "Why Do You Feel This Way? Summarizing Triggers of Emotions in Social Media Posts." In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2022, pp. 9436–9453.

[46] Tanya Goyal, **Junyi Jessy Li**, and Greg Durrett. "SNaC: Coherence Error Detection for Narrative Summarization." In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2022, pp. 444–463.

[47] Sheena Panthaplackel, Milos Gligoric, **Junyi Jessy Li**, and Raymond Mooney. "Using Developer Discussions to Guide Fixing Bugs in Software." In: *Findings of the Association for Computational Linguistics: EMNLP 2022 (EMNLP Findings)*. 2022, pp. 2292–2301.

[48] Tanya Goyal, **Junyi Jessy Li**, and Greg Durrett. "FALTE: A Toolkit for Fine-grained Annotation for Long Text Evaluation." In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (EMNLP Demo)*. 2022, pp. 351–358.

[49] Zachary W. Taylor, Maximus H. Chu, and **Junyi Jessy Li**. "Text Simplification of College Admissions Instructions: A Professionally Simplified and Verified Corpus." In: *Proceedings of the 29th International Conference on Computational Linguistics (COLING)*. 2022, pp. 6505–6515.

[50] Katherine Atwell, Remi Choi, **Junyi Jessy Li**, and Malihe Alikhani. "The Role of Context and Uncertainty in Shallow Discourse Parsing." In: *Proceedings of the International Conference on Computational Linguistics (COLING)*. 2022, pp. 797–811.

[51] Jiyang Zhang, Sheena Panthaplackel, Pengyu Nie, **Junyi Jessy Li**, and Milos Gligoric. "CoditT5: Pretraining for Source Code and Natural Language Editing." In: *Proceedings of the IEEE/ACM International Conference on Automated Software Engineering (ASE)*. 2022, pp. 1–12.

[52] Barea Sinno, Bernardo Oviedo, Katherine Atwell, Malihe Alikhani, and **Junyi Jessy Li**. "Political Ideology and Polarization: A Multi-dimensional Approach." In: *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. 2022, pp. 231–243.

[53] Tiberiu Sosea, Chau Pham, Alexander Tekle, Cornelia Caragea, and **Junyi Jessy Li**. "Emotion analysis and detection during COVID-19." In: *Proceedings of the Language Resources and Evaluation Conference (LREC)*. 2022, pp. 6938–6947.

[54] Ashwin Devaraj, William Sheffield, Byron C. Wallace, and **Junyi Jessy Li**. "Evaluating Factuality in Text Simplification." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. **Outstanding Paper Award**. 2022, pp. 7331–7345.

[55] Anubrata Das, Chitrank Gupta, Venelin Kovatchev, Matthew Lease, and **Junyi Jessy Li**. "ProtoTEx: Explaining Model Decisions with Prototype Tensors." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2022, pp. 2986–2997.

[56] Fangyuan Xu, **Junyi Jessy Li**, and Eunsol Choi. "How do we answer complex questions: Discourse structure of long form answers." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2022, pp. 3556–3572.

[57] Pengyu Nie, Jiyang Zhang, **Junyi Jessy Li**, Raymond J. Mooney, and Milos Gligoric. "Impact of Evaluation Methodologies on Code Summarization." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2022, pp. 4936–4960.

[58] Sheena Panthaplackel, **Junyi Jessy Li**, Milos Gligoric, and Raymond J. Mooney. "Learning to Describe Solutions for Bug Reports Based on Developer Discussions." In: *Findings of the Association for Computational Linguistics: ACL 2022 (ACL Findings)*. 2022, pp. 2935–2952.

[59] Tanya Goyal, Jiacheng Xu, **Junyi Jessy Li**, and Greg Durrett. "Training Dynamics for Text Summarization Models." In: *Findings of the Association for Computational Linguistics: ACL 2022 (ACL Findings)*. 2022, pp. 2061–2073.

[60] Anubha Kabra, Mehar Bhatia, Yaman Kumar Singla, **Junyi Jessy Li**, and Rajiv Ratn Shah. "Evaluation Toolkit For Robustness Testing Of Automatic Essay Scoring Systems." In: *Joint International Conference on Data Science & Management of Data (CODS-COMAD)*. 2022, pp. 90–99.

[61] Neha Srikanth and **Junyi Jessy Li**. "Elaborative Simplification: Content Addition and Explanation Generation in Text Simplification." In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021 (ACL Findings)*. 2021, pp. 5123–5137.

[62] Katherine Atwell, **Junyi Jessy Li**, and Malihe Alikhani. "Where Are We in Discourse Relation Recognition?" In: *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*. 2021, pp. 314–325.

[63] Ashwin Devaraj, Iain J. Marshall, Byron C. Wallace, and **Junyi Jessy Li**. "Paragraph-level Simplification of Medical Texts." In: *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. 2021, pp. 4972–4984.

[64] Elisa Ferracane, Greg Durrett, **Junyi Jessy Li**, and Katrin Erk. "Did they answer? Subjective acts and intents in conversational discourse." In: *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. 2021, pp. 1626–1644.

[65] Sheena Panthaplackel, **Junyi Jessy Li**, Milos Gligoric, and Raymond J. Mooney. "Deep Just-In-Time Inconsistency Detection Between Comments and Source Code." In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)* 35.1 (2021), pp. 427–435.

[66] Laura Manor, Yiran Su, Zachary W. Taylor, and **Junyi Jessy Li**. "How to apply for financial aid: Exploring perplexity and jargon in texts for non-expert audiences." In: *Proceedings of the Society for Computation in Linguistics (SCiL)* 4.1 (2021), pp. 380–383.

[67] Pengyu Nie, Karl Palmskog, **Junyi Jessy Li**, and Milos Gligoric. "Roosterize: Suggesting Lemma Names for Coq Verification Projects Using Deep Learning." In: *Proceedings of the 43rd International Conference on Software Engineering: Demonstrations Track (ICSE Demo)*. 2021, pp. 21–24.

[68] Wei-Jen Ko, Te-Yuan Chen, Yiyan Huang, Greg Durrett, and **Junyi Jessy Li**. "Inquisitive Question Generation for High Level Text Comprehension." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020, pp. 6544–6555.

[69] Venkata Subrahmanyan Govindarajan, Benjamin Chen, Rebecca Warholic, Katrin Erk, and **Junyi Jessy Li**. "Help! Need Advice on Identifying Advice." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020, pp. 5295–5306.

[70] Wei-Jen Ko and **Junyi Jessy Li**. "Assessing Discourse Relations in Language Generation from GPT-2." In: *Proceedings of the International Conference on Natural Language Generation (INLG)*. 2020, pp. 52–59.

[71] Jaeseong Lee, Pengyu Nie, **Junyi Jessy Li**, and Milos Gligoric. "On the Naturalness of Hardware Descriptions." In: *Proceedings of the ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*. 2020, pp. 530–542.

[72] Shrey Desai, Cornelia Caragea, and **Junyi Jessy Li**. "Detecting Perceived Emotions in Hurricane Disasters." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2020, pp. 5290–5305.

[73] Sheena Panthaplackel, Pengyu Nie, Milos Gligoric, **Junyi Jessy Li**, and Raymond Mooney. "Learning to Update Natural Language Comments Based on Code Changes." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2020, pp. 1853–1868.

[74] Pengyu Nie, Karl Palmskog, **Junyi Jessy Li**, and Milos Gligoric. "Deep Generation of Coq Lemma Names Using Elaborated Terms." In: *Proceedings of the International Joint Conference on Automated Reasoning (IJCAR)*. 2020, pp. 97–118.

[75] Swapnil Dhanwal, Hritwik Dutta, Hitesh Nankani, Nilay Shrivastava, Yaman Kumar, **Junyi Jessy Li**, Debanjan Mahata, Rakesh Gosangi, Haimin Zhang, Rajiv Ratn Shah, and Amanda Stent. "An Annotated Dataset of Discourse Modes in Hindi Stories." In: *Proceedings of the Conference on Language Resources and Evaluation (LREC)*. 2020, pp. 1191–1196.

[76] Yang Zhong, Chao Jiang, Wei Xu, and **Junyi Jessy Li**. "Discourse Level Factors for Sentence Deletion in Text Simplification." In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. 2020, pp. 9709–9716.

[77] Sheena Panthaplackel, Milos Gligoric, Raymond Mooney, and **Junyi Jessy Li**. "Associating Natural Language Comment and Source Code Entities." In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. 2020, pp. 8592–8599.

[78] Shrey Desai, Barea Sinno, Alex Rosenfeld, and **Junyi Jessy Li**. "Adaptive Ensembling: Unsupervised Domain Adaptation for Political Document Analysis." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing and the International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 2019, pp. 4712–4724.

[79] Hsin-Ping Huang and **Junyi Jessy Li**. "Unsupervised Adversarial Domain Adaptation for Implicit Discourse Relation Classification." In: *Proceedings of the Conference on Computational Natural Language Learning (CoNLL)*. 2019, pp. 686–695.

[80] Pengyu Nie, Rishabh Rai, **Junyi Jessy Li**, Sarfraz Khurshid, Raymond J. Mooney, and Milos Gligoric. "A Framework for Writing Trigger-Action Todo Comments in Executable Format." In: *Proceedings of the ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*. **ACM SIGSOFT Distinguished Paper Award**. 2019, pp. 385–396.

[81] Elisa Ferracane, Greg Durrett, **Junyi Jessy Li**, and Katrin Erk. "Evaluating Discourse in Structured Text Representations." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2019, pp. 646–653.

[82] Wei-Jen Ko, Greg Durrett, and **Junyi Jessy Li**. "Linguistically-Informed Specificity and Semantic Plausibility for Dialog Generation." In: *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. 2019, pp. 3456–3466.

[83] Wei-Jen Ko, Greg Durrett, and **Junyi Jessy Li**. "Domain Agnostic Real-Valued Specificity Prediction." In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. 2019, pp. 6610–6617.

[84] Yifan Gao, Yang Zhong, Daniel Preoţiuc-Pietro, and **Junyi Jessy Li**. "Predicting and Analyzing Language Specificity in Social Media Posts." In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. 2019, pp. 6415–6422.

[85] Eric Holgate, Isabel Cachola, Daniel Preoţiuc-Pietro, and **Junyi Jessy Li**. "Why Swear? Analyzing and Inferring the Intentions of Vulgar Expressions." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2018, pp. 4405–4414.

[86] Isabel Cachola, Eric Holgate, Daniel Preoţiuc-Pietro, and **Junyi Jessy Li**. "Expressively vulgar: The socio-dynamics of vulgarity and its effects on sentiment analysis in social media." In: *Proceedings of the International Conference on Computational Linguistics (COLING)*. **Area Chair Favorite**. 2018, pp. 2927–2938.

[87] Benjamin Nye, **Junyi Jessy Li**, Roma Patel, Yinfei Yang, Iain Marshall, Ani Nenkova, and Byron Wallace. "A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2018, pp. 197–207.

[88] An Thanh Nguyen, Byron Wallace, **Junyi Jessy Li**, Ani Nenkova, and Matthew Lease. "Aggregating and Predicting Sequence Labels from Crowd Annotations." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2017, pp. 299–309.

[89] **Junyi Jessy Li**, Kapil Thadani, and Amanda Stent. "The Role of Discourse Units in Near-Extractive Summarization." In: *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*. **Best Paper Award Nomination**. 2016, pp. 137–147.

[90] **Junyi Jessy Li** and Ani Nenkova. "The Instantiation Discourse Relation: A Corpus Analysis of Its Properties and Improved Detection." In: *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*. 2016, pp. 1181–1186.

[91] **Junyi Jessy Li**, Bridget O'Daniel, Yi Wu, Wenli Zhao, and Ani Nenkova. "Improving the Annotation of Sentence Specificity." In: *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*. 2016, pp. 3921–3927.

[92] **Junyi Jessy Li** and Ani Nenkova. "Detecting Content-Heavy Sentences: A Cross-Language Case Study." In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2015, pp. 1271–1281.

[93] **Junyi Jessy Li** and Ani Nenkova. "Fast and Accurate Prediction of Sentence Specificity." In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. 2015, pp. 2281–2287.

[94] **Junyi Jessy Li**, Marine Carpuat, and Ani Nenkova. "Cross-lingual Discourse Relation Analysis: A corpus study and a semi-supervised classification system." In: *Proceedings of the International Conference on Computational Linguistics (COLING)*. 2014, pp. 577–587.

[95] **Junyi Jessy Li**, Marine Carpuat, and Ani Nenkova. "Assessing the Discourse Factors that Influence the Quality of Machine Translation." In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*. 2014, pp. 283–288.

[96] **Junyi Jessy Li** and Ani Nenkova. "Addressing Class Imbalance for Improved Recognition of Implicit Discourse Relations." In: *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*. 2014, pp. 142–150.

[97] **Junyi Jessy Li** and Ani Nenkova. "Reducing Sparsity Improves the Recognition of Implicit Discourse Relations." In: *Proceedings of the Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*. 2014, pp. 199–207.

Articles in Peer-Reviewed Workshop Proceedings

[98] **Junyi Jessy Li**, Yang Janet Liu, Kanishka Misra, Valentina Pyatkin, and William Sheffield. "Which course? Discourse! Teaching Discourse and Generation in the Era of LLMs." In: *Workshop on Teaching Natural Language Processing (TeachingNLP)*. 2026.

[99] Yiheng Su, **Junyi Jessy Li**, and Matthew Lease. "Wrapper Boxes for Faithful Attribution of Model Predictions to Training Data." In: *Proceedings of the BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP*. 2024, pp. 551–576.

[100] Negin Raoof, Yating Wu, Carlos Bonilla, **Junyi Jessy Li**, Stephanie M. Grasso, Alex Dimakis, and Zoi Gkalitsiou. "Modeling Bilingual Disfluencies with Large Language Models." In: *Proceedings of the ICML Workshop on Large Language Models and Cognition*. 2024.

[101] Venelin Kovatchev, Trina Chatterjee, Venkata S Govindarajan, Jifan Chen, Eunsol Choi, Gabriella Chronis, Anubrata Das, Katrin Erk, Matthew Lease, **Junyi Jessy Li**, Yating Wu, and Kyle Mahowald. "longhorns at DADC 2022: How many linguists does it take to fool a Question Answering model? A systematic approach to adversarial attacks." In: *Proceedings of the First Workshop on Dynamic Adversarial Data Collection (DADC)*. 2022, pp. 41–52.

[102] Pengyu Nie, Karl Palmskog, **Junyi Jessy Li**, and Milos Gligoric. "Learning to Format Coq Code Using Language Models." In: *Proceedings of the Coq Workshop*. 2020, 2pp.

[103] Laura Manor and **Junyi Jessy Li**. "Plain English Summarization of Contracts." In: *Proceedings of the NAACL Workshop on Natural Legal Language Processing (NLLP)*. 2019, pp. 1–11.

[104] Elisa Ferracane, Titan Page, **Junyi Jessy Li**, and Katrin Erk. "From News to Medical: Cross-domain Discourse Segmentation." In: *Proceedings of the NAACL Workshop on Rhetorical Structure Theory and Related Formalisms (DISRPT)*. 2019, pp. 22–29.

[105] Pengyu Nie, **Junyi Jessy Li**, Sarfraz Khurshid, Raymond J. Mooney, and Milos Gligoric. "Natural Language Processing and Program Analysis for Supporting Todo Comments as Software Evolves." In: *Proceedings of the AAAI Workshop on NLP for Software Engineering (NL4SE)*. 2018, pp. 775–778.

[106] **Junyi Jessy Li**. "Estimating text intelligibility via information packaging analysis." In: *Proceedings of the AAAI Conference on Artificial Intelligence: Doctoral Consortium (AAAI-DC)*. 2016, pp. 4305–4306.

Patents

- Lucretia H. Vanderwende, Sumit Basu, Charles E. Jacobs and **Junyi Li**. "Generating high-level questions from sentences". *US Patent US10769958B2*. July 2019.

Abstracts and Conference Posters

- Heather Dial, Lokesh Pugalenthi, **Junyi Jessy Li**, and Maya Henry. "Application of machine learning, natural language processing, and temporal response function modeling for differential diagnosis of primary progressive aphasia." In: *The 52nd Clinical Aphasiology Conference*. 2024.
- Robin C. Hilsabeck, Jeffrey N. Keller, Maya L. Henry, **Junyi Jessy Li**, Lokesh Pugalenthi, Paul Rathouz, Paul Toprac, Patrick Chang, Joshua Chang, Suzanne Schmitz, Avery Largent, Heather Foil, Robert Brouillette, Rosemary A. Lester-Smith. "Combining working memory and speech data in a digital app for early detection of cognitive impairment in primary care." In: *Alzheimer's Association International Conference (Alzheimer's & Dementia, 19: e082210)*. 2023.
- Heather Dial, Lokesh Pugalenthi, Nike Gnanateja Gurindapalli, **Junyi Jessy Li**, and Maya Henry. "Cortical tracking of acoustic and linguistic features in primary progressive aphasia: Emerging findings and methodological considerations." In: *The 61st Annual Meeting of the Academy of Aphasia*. 2023.
- Robin C. Hilsabeck, Jeff Keller, Maya L. Henry, **Junyi Jessy Li**, Lokesh Pugalenthi, Paul Rathouz, Paul Toprac, Joshua Chang, Suzanne Schmitz, Avery Largent, Heather Foil, Robert Brouillette, and Rosemary Lester-Smith. "Combining working memory and speech data in a digital app for early

detection of cognitive impairment in primary care." In: *Alzheimer's Association International Conference (AAIC)*. 2023.

- Heather Dial, Lokesh Pugalenthi, Maya Henry, and **Junyi Jessy Li**. "Application of machine learning, natural language processing, and temporal response function modeling for differential diagnosis of primary progressive aphasia." In: *Clinical Aphasiology Conference*. 2023.
- Heather Dial, G. Nike Gnanateja, Lokesh Pugalenthi, Rachel Tessmer, Bharath Chandrasekaran, Maya Henry, and **Junyi Jessy Li**. "Cortical Tracking of Semantic Dissimilarity for Features Derived Using Static and Contextualized Embeddings in Primary Progressive Aphasia." In: *45th Association for Research in Otolaryngology (ARO) Mid-Winter Meeting*. 2022.
- Avery Largent, Sonal Alla, Jeff Keller, Maya Henry, Paul Toprac, Joshua Chang, Heather Cuevas, Paul Rathouz, **Junyi Jessy Li**, Thomas Parsons, Anna Norman, Varshinee Sreekanth Robert Brouillette, Heather Foil, and Robin Hilsabeck. "Leveraging Technology to Make Cognitive Screening Easy for Older Adults and Primary Care Providers." In: *Aging and Health Informatics Conference (AHIC)*. 2021.
- Lokesh Pugalenthi, Heather Dial, Nike Gnanateja, Rachel Tessmer, Maya Henry, and **Junyi Jessy Li**. "Contrasting Static and Contextualized Embeddings in the Use of Semantic Feature Vectors in Neurophysiological Prediction." In: *TACC Symposium for Texas Researchers (TACCSTER)*. 2021.
- **Junyi Jessy Li**, Julia Parish-Morris, Leila Bateman, and Ani Nenkova. "Autism quotient scores modulate the perception and production of text specificity in adult females." In: *International Meeting for Autism Research (IMFAR)*. 2017.

Invited Talks

- "Discourse Centric Models and Evaluation." Machine Learning Seminar, Georgia Institute of Technology, Atlanta, GA. April 2026.
- "Benchmarking and Evaluating NLP models." Language AI in the Space Sciences Workshop. Space Telescope Science Institute, Baltimore MA. March 2026.
- "Faithfulness vs. Safety in AI for Medicine." Texas Symposium on Machine Learning, Responsible AI & Robotics, Austin TX. March 2026.
- "Discourse Centric Generative Models." Computer Science Colloquium, Rice University, Houston, TX. February 2026.
- "Engaging experts and LLMs in corpora development." The 19th Linguistic Annotation Workshop (LAW-XIX) Keynote. Vienna, Austria. July 2025.
- "Discourse Models with Questions Under Discussion." Society for Computation in Linguistics (SCiL) Keynote. Eugene, OR. July 2025.
- "Discourse Models with Language Models." The Centre for Linguistic Theory and Studies in Probability (CLASP) Seminar, University of Gothenburg, Gothenburg, Sweden. June 2025.
- "Exploring the LLM universe for astronomy research." CosmicAI Hybrid Seminar Series, UT Austin, Austin, TX. April 2025.
- "Discourse Models with Language Models." Computer Science Colloquium, University of Massachusetts Lowell, Lowell, MA. April 2025.
- "Discourse Models with Language Models." Khoury College of Computer Sciences, Northeastern University, Boston, MA. April 2025.
- "Discourse Models with Language Models." Language Technology Lab Seminars, University of Cambridge, Cambridge, UK. March 2025.
- "Discourse Models with Language Models." Computer Science Seminar Series, University of Birmingham, Birmingham, UK. March 2025.
- "Helpful and Safe(r) LLMs? Deciphering emotions and faithfully communicating medical evidence." Workshop on Mind, Language and Ethics in AI, UT Austin, Austin TX. October 2024.
- "Rethinking Discourse Processing and Generation in the Era of Large Language Models." Center for Mind/Brain Sciences (CIMeC), University of Trento, Trento, Italy. October 2024.
- "What can we do with massive generative abilities? Rethinking Discourse Processing and Generation." Marburg AI Talks — Seminar Series, University of Marburg, Marburg, Germany. July 2024.

- "Large Language Models for Software Evolution." University of Michigan–Shanghai Jiao Tong University Joint Institute, Shanghai, China. July 2024.
- "Which questions should I answer? Salience Prediction of Inquisitive Questions." Computational Linguistics Group, Georgetown University, Washington DC. May 2024.
- "Modeling Discourse as Questions and Answers." Data Science Seminar Series, Texas A&M University, College Station, TX. March 2024.
- "Deciphering Emotions with Language Models." Seminar on language generation and summarization, Columbia University, New York, NY. November 2023.
- "A Computational Approach to Questions Under Discussion in Discourse Modeling." South by Semantics Workshop, UT Austin, Austin, TX. November 2023.
- "Working with Language Models: Promises and Woes." St. David's CHPR Colloquium Series, School of Nursing, UT Austin, Austin, TX. November 2023.
- "Tell us plainly! The promises and woes of using LLMs to make medical information more accessible." *Part of panel: "Simplifying Medical Texts with Large Language Models: Challenges and Opportunities"*. Health Literacy Annual Research Conference. Online. October 2023.
- "Modeling Discourse as Questions and Answers." The University of Texas at Dallas, Dallas, TX. October 2023.
- "Deciphering Emotional Dynamics with Language Models." Smart Cities and Generative AI Symposium, UT Austin, Austin, TX. August 2023.
- "Modeling Discourse as Questions and Answers." The 12th Joint Conference on Lexical and Computational Semantics, Toronto, Canada. July 2023.
- "Deciphering Emotional Dynamics on Social Media." Department of Linguistics, Northwestern University, Evanston, IL. May 2023.
- "Text Simplification and Discourse Comprehension." Department of Computer Science, University of Illinois at Urbana-Champaign, Champaign, IL. August 2022.
- "Designing Responsible AI Technologies to Curb Disinformation." Good Systems Annual Symposium, UT Austin, Austin, TX. April 2022.
- "New Challenges in Text Simplification." Institute for Foundations of Machine Learning, UT Austin, Austin, TX. March 2022.
- "Artificial Intelligence: The Language Chapter." Plática Series, Mentoring to Achieve Latino Educational Success (Project MALES), College of Education, UT Austin, Austin, TX. October 2021.
- "New Challenges in Text Simplification." Center for Language and Speech Processing, Johns Hopkins University, Baltimore, MD. September 2021.
- "New Challenges in Text Simplification." Media Data Science Research Lab, Adobe Systems, Noida, India. June 2021.
- "New Challenges in Text Simplification." Amazon AWS Research, Seattle, WA. June 2021.
- "Comment Maintenance as Software Evolves." Department of Computer Science, University of Houston, Houston, TX. April 2021.
- "Help! Need Advice on Discourse Comprehension." College of Computing, Georgia Institute of Technology, Atlanta, GA. March 2021.
- "Help! Need Advice on Discourse Comprehension." Department of Computer Science, University of Pittsburgh, Pittsburgh, PA. March 2021.
- "Help! Need Advice on Discourse Comprehension." Computational Linguistics and Information Processing, University of Maryland College Park, College Park, MD. March 2021.
- "High level inferences on social media language: advice and emotions." Valkyrie Intelligence, Austin, TX. January 2021.
- "Help! Need Advice on Identifying Advice." Salesforce Research, Palo Alto, CA. December 2020.
- "Forensic Linguistics." Texas Forensics Guest Lecture Series at UT Austin, Austin, TX. November 2019.
- "Expressively Vulgar: The socio-dynamics of vulgarity." Undergraduate Linguistics Society, UT Austin, Austin, TX. April 2019.
- "How Computers Handle Language." Osher Lifelong Learning Institute, UT Austin, Austin, TX. November 2018.
- "Assessing Text Intelligibility with Discourse Structure." Department of Computer Science, Texas

State University, San Marcos, TX. March 2017.
- "Assessing Text Intelligibility with Discourse Structure." Department of Computer Science, University of California Davis, Davis, CA. March 2017.
- "Assessing Text Intelligibility with Discourse Structure." College of Information Science and Technology, Pennsylvania State University, State College, PA. March 2017.
- "Assessing Text Intelligibility with Discourse Structure." Department of Linguistics, UT Austin, Austin, TX. February 2017.
- "Assessing Text Intelligibility with Discourse Structure." Department of Computer Science, University of Illinois at Chicago, Chicago, IL. February 2017.
- "Assessing Text Intelligibility with Discourse Structure." Department of Linguistics, University of Illinois at Urbana-Champaign, Champaign, IL. February 2017.
- "Assessing Text Intelligibility with Discourse Structure." Department of Computer Science, Washington University in St. Louis, St. Louis, MO. February 2017.
- "Assessing Text Intelligibility with Discourse Structure." Department of Computer Science, University of Iowa, Iowa City, IA. January 2017.
- "Assessing Text Intelligibility with Discourse Structure." School of Information, UT Austin, Austin, TX. January 2017.
- "Text Specificity: How and Why." Bloomberg LP, New York, NY. November 2016.
- "Text Specificity: How and Why." Department of Computer Science, UT Austin, Austin, TX. October 2016.
- "Understanding Text Specificity." Interactions Corporation, New Providence, NJ. June 2016.
- "Analyzing General and Specific Content in Text." Computational Linguistics and Information Processing, University of Maryland College Park, College Park, MD. September 2015.
- "Content-heavy Sentences in Cross-Lingual Context." Department of Computer Science, Columbia University, New York, NY. September 2015.
- "Predicting Sentence Specificity." Center for Undergraduate Research and Fellowships, University of Pennsylvania, Philadelphia, PA. September 2015.

Invited Talks/Panels (Community and Public Interest)

- Lecturer, AI in Health Communication, Texas Epidemic Public Health Institute & UT Center for Health Communication, Austin, TX. September 2025.
- Guest panelist, Linguistics Society of America (LSA) Annual Meeting, Philadelphia, PA. January 2025.
- Guest speaker, Undergraduate Linguistic Society. College of Liberal Arts, UT Austin, Austin, TX. November 2024.
- 2nd Women and Gender Minorities Research Day Panel Discussion. Department of Computer Science, UT Austin, Austin, TX. March 2024.
- "Artificial Intelligence: The Language Chapter." Westlake High School Senior Career Day, Austin, TX. February 2024.
- Liberal Arts Week Research Showcase. College of Liberal Arts, UT Austin, Austin, TX. February 2024.
- Biannual Advisory Council AI Panel Discussion. College of Liberal Arts, UT Austin, Austin, TX. November 2023.
- NSF CAREER "Ask Us Anything" Panel Discussion. The Office of the Vice President for Research, Scholarship and Creative Endeavors, UT Austin, Austin, TX. February 2023.
- "Artificial Intelligence: The Language Chapter." Westlake High School Senior Career Day, Austin, TX. February 2023.
- "Defining a Research Agenda." EECS Rising Stars Workshop, UT Austin, Austin, TX. October 2022.
- "Artificial Intelligence: The Language Chapter." Westlake High School Senior Career Day, Austin, TX. February 2022.
- "Speciteller – an NLP tool developed at UPenn." Philadelphia Python Users Group, Philadelphia, PA. June 2016.
- "How does the computer understand language?" Women in Computer Science High School Open Day, University of Pennsylvania, Philadelphia, PA. April 2016.
- "How does the computer understand language?" Lower Merion High School, Philadelphia, PA.

March 2016.

Selected Research Funding

- **Automated Reasoning Amazon Research Awards**, "Documenting and Recommending Tactics in HOL Light" 2026
  PI with Gligoric, Milos. Total funding: $75,000 plus $40,000 in AWS credits
- **Cisco Research**, "Developing Code Verifier Agents" 2025
  PI with Gligoric, Milos. Total funding: $75,000
- **NSF and Simons Foundation**, "NSF-Simons AI Institute for Cosmic Origins" 2024 – 2029
  Research Group Lead (AI). Director: Offner, Stella; co-director: Lease, Matthew. Total funding: $20,000,000
- **NIH R01**, "Building Safety Guards into LLMs for Trustworthy Automatic Simplification of Medical Documents" 2024 – 2028
  PI. M-PIs: Wallace, Byron; Xu, Wei. Total funding: $1,374,201
- **Good Ventures Foundation via Open Philanthropy**, "AI Safety Research & Field-Building Project" 2024 – 2027
  Co-I. PIs: Lederman, Harvey; Mahowald, Kyle; Knox, Bradley. Total funding: $775,000
- **NSF Core**, "Collaborative Research: SHF: Medium: Natural Language Models with Execution Data for Software Testing" 2023 – 2027
  Co-PI. PI: Gligoric, Milos; co-PI: Mahowald, Kyle. Total funding: $900,000
  NSF collaborative project; collaboration institution: University of Illinois at Urbana Champaign, PI: Misailovic, Sasa. Total project award (both UT and UIUC): $1,200,000.
- **NSF CAREER Award**, "CAREER: Discourse Processing and Content Generation for Document Simplification" 2022 – 2027
  PI. Total funding: $540,543
- **Cisco Systems**, "Classifying Text with Intuitive and Faithful Model Explanations" 2022 – 2024
  PI with Lease, Matthew A. Total funding: $199,458
- **NSF Core**, "Collaborative Research: HCC: Medium: Fine-grained Emotion Analysis in Crises" 2021 – 2026
  PI. Co-PI: Pennebaker, James W. Total funding: $452,312
  NSF collaborative project; collaboration institution: University of Illinois at Chicago, PI: Caragea, Cornelia. Total project award (both UT and UIC): $799,926.
- **Good Systems, UT Austin OVPR**, "Designing Responsible AI Technologies to Protect Information Integrity" 2021 – 2027
  Co-I. PI: Lease, Matthew A. Total funding: $725,000
- **NSF Computer Research Initiation Initiative (CRII) Award**, "CRII:RI: A Multi-level Framework for Text Specificity" 2019 – 2022
  PI. Total funding: $190,496
- **Google Faculty Research Award**, "Co-evolution of Code and Comments" 2019
  Co-PI. PI: Mooney, Raymond J.; co-PI: Gligoric, Milos. Total funding: $55,314
- **Salesforce Research Deep Learning Grant**, "Hierarchical Graph-based Advice Summarization from Online Forums" 2018
  PI. Co-PI: Erk, Katrin. Total funding: $50,000
- **Amazon Alexa Fund**, "Generating Informative Conversation Responses with Text Specificity and Opportunistic Learning" 2018

PI. Co-PIs: Durrett, Greg and Mooney, Raymond J. Total funding: $63,000 plus $13,200 credits for Amazon Web Services.

Industry Experience

- **Intelligent Automation Inc., Rockville, MD**
  Technical Consultant — 2020 – 2022
- **ConverseNow, Inc., Austin, TX**
  Technical Advisor — 2020 – 2021
- **Yahoo! Labs, New York, NY**
  Research Intern — Summer 2015
- **Microsoft Research, Redmond, WA**
  Research Intern — Summer 2013

Awards and Honors

2024	**Outstanding Paper Award**, Conference on Empirical Methods in Natural Language Processing (EMNLP)
2023	**Best Area Chair**, Conference on Empirical Methods in Natural Language Processing (EMNLP)
2022	**Outstanding Paper Award**, Annual Meeting of the Association for Computational Linguistics (ACL)
2020	**Outstanding Area Chair**, Conference on Empirical Methods in Natural Language Processing (EMNLP)
2020	**Outstanding Senior Program Committee Member**, AAAI Conference on Artificial Intelligence (AAAI)
2019	**ACM SIGSOFT Distinguished Paper Award**, the ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)
2018	**Area Chair Favorite Paper**, International Conference on Computational Linguistics (COLING)
2016	**Best Paper Award Nomination**, Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)
2016	**Best Reviewer**, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

Teaching

Instructor

University of Texas at Austin | Linguistics (LIN), Computer Science (CS)

LIN 313	*Language and Computers*, Spring 2018, Fall semesters 2018 – 2020
LIN 350	*Analyzing Linguistic Data*, Fall 2023
LIN 353D/CS 378	*Computational Discourse and Natural Language Generation*, Fall 2025
LIN 371/373N	*Machine Learning for Text Analysis*, Spring semesters 2019 – 2024, Fall 2024
CS/AI/DES 388	*Natural Language Processing (For CS Online MS Programs)*, Fall 2025
LIN 389C	*Research in Computational Linguistics*, Spring semesters 2018 – 2022, Fall 2024, Spring 2026
LIN 393 Topics	*Social Applications and Impact of NLP*, Spring 2023, 2024 *Natural Language Generation*, Fall 2022 *Discourse Processing and Natural Language Generation*, Fall 2020 *Computational Linguistics with Weak Supervision*, Fall 2019 *Computational Discourse*, Fall 2018

Advising and Student-Related Service

Ph.D. Students Supervised

- Sebastian Joseph, Computer Science.
- Kaijie Mo, Linguistics.
- William Sheffield, Computational Linguistics.
- Manya Wadhwa, Computer Science, co-advised with Greg Durrett.
- Yating Wu, Electrical and Computer Engineering, co-advised with Alex Dimakis.
- Asher Zheng, Computational Linguistics, co-advised with David Beaver.
- Hongli Zhan, Ph.D. in Computational Linguistics, 2026. (Dissertation chair).
- Anubrata Das, Ph.D. in Information Studies, 2025. (Dissertation co-chair; chair: Matthew Lease).
- Jiyang Zhang, Ph.D. in Electrical and Computer Engineering, 2025. (Dissertation co-chair; chair: Milos Gligoric).
- Venkata Subrahmanyan Govindarajan, Ph.D. in Computational Linguistics, 2024. (Dissertation chair; co-chair: David Beaver).
- Wei-Jen Ko, Ph.D. in Computer Science, 2022. (Dissertation chair; co-chair: Greg Durrett).
- Sheena Panthaplackel, Ph.D. in Computer Science, 2022. (Dissertation co-chair; chair: Raymond J. Mooney).
- Elisa Ferracane, Ph.D. in Computational Linguistics, 2020 (Dissertation co-chair; chair: Katrin Erk).

Dissertation Committees (as Member)

- Yejin Cho, Computational Linguistics
- Rajarshi Haldar, Computer Science, University of Illinois Urbana-Champaign
- Jierui Li, Computer Science
- Tiffany M. Phan, Computational Astronautical Sciences and Technologies (Oden Institute)
- Yidan Sun, College of Computing and Data Science, Nanyang Technological University, Singapore
- Yang Zhong, Computer Science, University of Pittsburgh
- Zexi Zhou, Human Development and Family Sciences
- Barea Sinno, Ph.D. in Political Science, Rutgers University, 2026
- Giuliano Martinelli, Ph.D. in Artificial Intelligence, Sapienza University of Rome, Italy, 2026
- Yinhong Liu, Ph.D. in Natural Language Processing, Cambridge University, UK, 2025
- Eric Holgate, Ph.D. in Computational Linguistics, 2025
- Gabriela O'Connor, Ph.D. in Spanish and Portuguese, 2025
- Jan Trienes, Ph.D. in Computer Science, Marburg University, Germany, 2025
- René Knaebel, Ph.D. in Computational Linguistics, University of Potsdam, Germany, 2025

- Chuyuan Li, Ph.D. in Computer Science, Université de Lorraine, Nancy, France, 2023
- Tanya Goyal, Ph.D. in Computer Science, 2023
- Pengyu Nie, Ph.D. in Electrical and Computer Engineering, 2023
- Tiberiu Sosea, Ph.D. in Computer Science, University of Illinois at Chicago, 2023
- Xuewen Yao, Ph.D. in Electrical and Computer Engineering, 2023
- Jiacheng Xu, Ph.D. in Computer Science, 2022
- Rozen Neupane, Ph.D. in French Studies, 2020
- Zachary Taylor, Ph.D. in Education Leadership, 2020
- Su Wang, Ph.D. in Computational Linguistics, 2020

Master's Theses

- Emma Gueorguieva, M.A. in Psychology, Fall 2024 (Reader).
- Sebastian Joseph, M.S. in Computer Science, Spring 2024 (Chair).
- Ritika Mangla, M.S. in Computer Science, Spring 2024 (Chair; Co-chair: Greg Durrett).
- Smriti Singh Bondili, M.S. in Computer Science, Spring 2024 (Chair).
- Anisha Gunjal, M.S. in Computer Science, 2024. (2nd Reader; Chair: Greg Durrett).
- Yian Wong, M.S. in Computer Science, 2024. (2nd Reader; Chair: Matt Lease).
- Yiheng Su, M.S. in Computer Science, 2023. (Co-chair with Matt Lease).
- Michail Mersinias, M.S. in Computer Science, 2023. (2nd Reader; Chair: Kyle Mahowald).
- Ashwin Devaraj, M.S. in Computer Science, 2022. (Chair).
- Laura Manor, M.A. in Computational Linguistics, 2020. (Chair).
- Neha Srikanth, M.S. in Computer Science, 2020. (Chair).
- Elisa Ferracane, M.A. in Computational Linguistics, 2018. (Co-chair with Katrin Erk).
- Eric Holgate, M.A. in Computational Linguistics, 2018. (Co-chair with Katrin Erk).

Master's Research Supervision

- Anurag R. Patil, Computational Data Science, Spring – Summer 2021
- Prakhar Singh, Computer Science, Spring 2020
- Yiran (Erin) Su, Electrical and Computer Engineering, Spring 2020
- Te-Yuan Chen, School of Information, Spring 2020
- Yiyan Huang, School of Information, Spring 2020
- Hsin-Ping Huang, Computer Science, Spring – Summer 2018

Undergraduate Honors Theses

- Victor Wang, Computer Science, 2025 (2nd Reader)
- Ramya Namuduri, Computer Science, 2025 (Co-superviser with Greg Durrett)
- Allen Zheng, Computer Science, 2024 (Supervisor)
- Lokesh Pugalenthi, Computer Science, 2023 (Supervisor)
- Ajith Kemisetti, Computer Science, 2023 (Supervisor)
- Ojas Ahuja, Computer Science, 2022 (2nd Reader)
- Shrey Desai, Computer Science, 2020 (2nd Reader)
- Akshay Kumar Gupta, Computer Science, 2020 (2nd Reader)
- Jaydeep Singh, Computer Science, 2019 (2nd Reader)

Undergraduate Research Supervision

- Siddhartha Venkatayogi, Computer Science, Fall 2025 – present
- Daniel Brubaker, Linguistics, Fall 2024 – present
- Marco Cheung, Computer Science, Spring 2025 – Fall 2025
- Syed Murtaza Husain, Computational Physics and Astronomy, Summer 2024 – Spring 2025
- Katie Yan, Computer Science and Mathematics, Fall 2024 – Spring 2025
- Lily Chen, Electrical Engineering and Computer Science, MIT, Summer 2023 – Spring 2025

CRA-W Distributed Research Experiences for Undergraduates (DREU) mentee

- Karthik Nemmani, Comupter Science, Fall 2024.
- Jasmine Xu, Statistics and Data Science, Fall 2024.
- Allen Zheng, Computer Science, Fall 2023 – Spring 2024
- Yuki Zang, Computer Science, Brown University, Summer 2023 – Spring 2024
  CRA-W Distributed Research Experiences for Undergraduates (DREU) mentee
- Karim Villaescusa Feuchter, Linguistics, Spring – Fall 2023
- Ritvik Renikunta, Computer Science, Fall 2023
- Yian Wong, Computer Science, Spring 2022 – Spring 2023
- Sebastian Joseph, Computer Science, Spring 2022 – Summer 2023
- Ajith Kemisetti, Computer Science, Fall 2022
- Lokesh Pugalenthi, Computer Science & French, Spring 2021 – Spring 2023
- William Sheffield, Linguistics & Mathematics, Summer 2021 – Summer 2022
- Anna Alvis, Linguistics, Spring 2022
- Rahul Banerjee, Electrical and Computer Engineering, Spring 2022
- Maximus Chu, Computer Science, Spring 2022
- Kathryn Kazanas, Linguistics, Spring 2022 – Summer 2023
- Keziah Kaylyn Reina, Linguistics, Spring 2022 – Fall 2023
- Apollo Reyes, Linguistics, Spring 2022
- John Henry Cruz, Computer Science, Fall 2021
- Chitrank Gupta, Computer Science, India Institute of Technology Bombay, Summer 2021 – Fall 2021
- Ian Shin, Computer Science & Mathematics, Fall 2021
- Cutter Dalton, Linguistics, Spring 2021 – Summer 2021
- Lisa DiSalvo, Computer Science, Arcadia University, Summer 2021
  CRA-W Distributed Research Experiences for Undergraduates (DREU) mentee
- Arjun Kunjilwar, Computer Science, Spring 2021 – Summer 2021
- Chau Pham, Computer Science w/ minor in Applied Mathematics and Linguistics, Colgate University, Summer 2020 – Summer 2021
  CRA-W Distributed Research Experiences for Undergraduates (DREU) mentee
- Ashwin Devaraj, Computer Science, Spring 2020 – Spring 2021
- Gauraang Dhamankar, Computer Science, Summer 2020 – Spring 2021
- Eliza J Fisher, Linguistics, Spring 2021
- Mark Simmons, Linguistics, Spring 2021
- Neha Srikanth, Computer Science, Spring 2019 – Summer 2021
- Bernardo Oviedo, Computer Science & Mathematics, Spring 2020 – Fall 2020
- Alexander Tekle, Electrical and Computer Engineering, Summer 2020
  Honors: Undergraduate Research Fellowship
- Benjamin Chen, Computer Science, Summer 2019 – Spring 2020
- Shrey Desai, Computer Science, Spring 2019 – Fall 2019
  Honors: Computing Research Association (CRA) Outstanding Undergraduate Researcher 2020 Honorable Mention, Lala Fay & Claude DeVan Watts Endowed Presidential Scholarship, College of Natural Sciences Dean's Honored Graduate, Undergraduate Research Fellowship
- Rebecca Warholic, Computer Science & Italian Studies, Smith College, Summer 2019
  CRA-W Distributed Research Experiences for Undergraduates (DREU) mentee
- Yang Zhong, Mathematics, Summer 2018 – Spring 2019
- Isabel Cachola, Mathematics, Spring 2018 – Summer 2018
  Honors: College of Natural Sciences Dean's Honored Graduate, Aspire Research Excellence Award, Undergraduate Research Fellowship
- Yifan Gao, Mathematics, Summer 2018
- Victor Mena, Linguistics, Summer 2018

Professional Service

Executive Office

- **Secretary**, Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL), 2024 – present.
- **Board member**, Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2019 – 2021, 2023 – 2025.

Organizer, Editor

- **Program co-chair**, Conference on Empirical Methods in Natural Language Processing (EMNLP) 2026.
- **Action editor**, Computational Linguistics Journal (CL), 2025 – present.
- **Action editor**, Transactions of the Association for Computational Linguistics (TACL), 2024 – present.
- **Co-organizer**, Workshop on Computational Approaches to Discourse (CODI), 2020 – present.
- **Associate editor**, Dialogue and Discourse Journal, 2021 – 2025.
- **Tutorial co-chair**, Conference on Empirical Methods in Natural Language Processing (EMNLP) 2024.
- **Program co-chair**, Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2022.
- **Publication chair**, Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2021.
- **Co-organizer**, Workshop on Natural Language Processing for Programming (NLP4Prog), 2021.
- **Local organizer**, Young Researchers' Roundtable on Spoken Dialog Systems (YRRSDS), 2014.

Senior Program Committee Member

- **Best paper award committee**, International Joint Conference on Natural Language Processing & Asia-Pacific Chapter of the Association for Computational Linguistics (IJCAI-AACL), 2025.
- **Senior area chair**, *Summarization*, International Joint Conference on Natural Language Processing & Asia-Pacific Chapter of the Association for Computational Linguistics (IJCNLP-AACL) 2025.
- **Area chair**, Conference on Language Modeling (COLM) 2025.
- **Senior area chair**, *Discourse and Pragmatics*, Conference of the Association for Computational Linguistics (ACL) 2025.
- **Senior area chair**, *Discourse and Pragmatics*, Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL) 2025.
- **Senior area chair**, *Discourse and Pragmatics*, Conference of the Association for Computational Linguistics (ACL) 2024.
- **Senior area chair**, *Discourse and Pragmatics; Resources and Evaluation*, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2024.
- **Senior area chair**, *Discourse and Pragmatics*, European Chapter of the Association for Computational Linguistics (EACL) 2024.
- **Senior area chair**, *Natural Language Generation, Summarization and Simplification*, Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING) 2024.
- **Area chair**, AAAI Conference on Artificial Intelligence (AAAI) 2024.
- **Area chair**, *Discourse and Pragmatics*, Conference on Empirical Methods in Natural Language Processing (EMNLP) 2023.
- **Area chair**, *Discourse and Pragmatics*, Annual Meeting of the Association for Computational Linguistics (ACL) 2023.
- **Area chair**, *Resources and Evaluation*, Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL), 2023.
- **Area chair**, *Discourse, dialogue, and generation*, Joint Conference on Lexical and Computational Semantics (*SEM), 2023.
- **Senior area chair**, *Discourse and Pragmatics*, Annual Meeting of the Association for Computational Linguistics (ACL) 2022.
- **Area chair**, AAAI Conference on Artificial Intelligence (AAAI) 2022.
- **Senior area chair**, *Discourse and Pragmatics*, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2021.

- **Best paper selection committee**, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2021.
- **Area chair**, *Machine Learning*, Annual Meeting of the Association for Computational Linguistics (ACL) 2021.
- **Area chair**, *Dialog Systems*, International Conference on Natural Language Processing and Chinese Computing (NLPCC) 2021.
- **Senior PC member**, AAAI Conference on Artificial Intelligence (AAAI), 2021.
- **Area chair**, *Discourse and Pragmatics*, Annual Meeting of the Association for Computational Linguistics (ACL) 2020.
- **Area chair**, *Discourse and Pragmatics*, Discourse and Pragmatics, Conference on Empirical Methods in Natural Language Processing (EMNLP) 2020.
- **Senior PC member**, International Joint Conference on Artificial Intelligence (IJCAI) 2020.
- **Area chair**, *Discourse, dialogue and generation*, Joint Conference on Lexical and Computational Semantics (*SEM) 2020.
- **Area chair**, *Discourse and Pragmatics*, Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2018.

Program Committee Member

- Conference on Language Modeling (COLM), 2024
- Conference on Empirical Methods in Natural Language Processing (EMNLP), 2016 – 2022
- International Conference on Computational Linguistics (COLING), 2014 – 2022
- European Chapter of the Association for Computational Linguistics (EACL), 2017, 2021
- Conference on Computational Natural Language Learning (CoNLL), 2015, 2019, 2021
- Workshop on Understanding Implicit and Underspecified Language (UnImplicit), 2021
- Workshop on Natural Legal Language Processing (NLLP), 2020, 2021
- Widening Natural Language Processing Workshop (WiNLP), 2018, 2019, 2021, 2023
- Workshop on Noisy User-generated Text (W-NUT), 2016 – 2021
- Asian Chapter of the Association for Computational Linguistics (AACL), 2020
- Joint Conference on Lexical and Computational Semantics (*SEM), 2019
- AAAI Conference on Artificial Intelligence (AAAI), 2018, 2019
- International Joint Conference on Artificial Intelligence (IJCAI), 2018, 2019
- Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2018, 2019
- Annual Meeting of the Association for Computational Linguistics (ACL), 2017 – 2019
- North American Chapter of the Association for Computational Linguistics (NAACL), 2015 – 2019
- Student Research Workshops at *ACL (SRW), 2013, 2015 – 2017
- Recent Advances in Natural Language Processing (RANLP), 2013, 2015

Reviewer

- **Standing reviewer**, Computational Linguistics Journal, 2021 – 2025.
- **Standing reviewer**, Transactions of the Association for Computational Linguistics (TACL), 2020 – 2024.
- **Standing reviewer**, ACL Rolling Review, 2021 – 2022.
- **Reviewer**, Journal of Artificial Intelligence Research, 2023.
- **Reviewer**, Nature Human Behavior, 2023.
- **Reviewer**, Engineering Applications of Artificial Intelligence, 2023.
- **Reviewer**, Big Data & Society, 2021.
- **Reviewer**, Dialogue and Discourse, 2017 – 2019.

Panelist and Proposal Reviewer

- **Panelist**, National Science Foundation, 2019 – present.
- **External reviewer**, Natural Sciences and Engineering Research Council of Canada, 2023.
- **Panelist and board member**, Science Fund, Republic of Serbia, 2020, 2023.

MENTOR

- **Research mentor**, CRA-W Distributed Research Experiences for Undergraduates, 2019 – 2021, 2023.
- **Mentor**, Student Research Workshop at ACL, 2016, 2023.
- **Mentor**, Student Research Workshop at NAACL, 2019.