

ACADEMIC POSITIONS

2020-present	Assistant Professor of Linguistics and of Computer Science (by courtesy)	Northwestern University
2024-present	Faculty Fellow, Institute for Policy Research	
2020-present	Affiliated Faculty, Cognitive Science Program	
2020-2024	Faculty Affiliate, Institute for Policy Research	

EDUCATION

2019-2020	Postdoctoral Scholar in Computer Science	Stanford University Advisor: Dan Jurafsky
2014-2019	Ph.D. in Linguistics	Stanford University Dissertation: <i>Computational Linguistic Models of Police-Community Interaction</i> Committee: Dan Jurafsky (chair), John R. Rickford, Penelope Eckert, Robert J. Podesva, Jennifer L. Eberhardt
2011-2013	M.A. in East Asian Studies	Stanford University Thesis: <i>Computational Linguistics for Literature: Case Studies in Chinese</i> Advisor: Dan Jurafsky
2010-2011	Fulbright Scholar in Translation Theory	National Taiwan Normal University
2009	Graduate Certificate in China Studies	The Hopkins-Nanjing Center
2008	B.A. in Chinese, with Departmental and General Honors.	Vassar College

AWARDS

2018	Cialdini Prize for Voigt et al. (2017), "an outstanding contribution that uses field methods and demonstrates relevance to outside groups," Society for Personality and Social Psychology
2017	Cozzarelli Prize for Voigt et al. (2017), Best Paper in the Behavioral and Social Sciences, Proceedings of the National Academy of Sciences
2016-2019	Michelle and Kevin Douglas Stanford Interdisciplinary Graduate Fellowship, Stanford University
2014	2nd Place Best Paper for Voigt et al. (2014), Speech Prosody 7
2011-2012	Foreign Languages and Area Studies Fellowships, Chinese (2011) and Japanese (2012), Stanford University

GRANT ACTIVITY

As Co-Principal Investigator

- 2025-2029
in preparation Co-Principal Investigator, with Harold Pollack. [911, What's Your Emergency?: Multimodal Computational Modeling to Improve 911 Call Coding and Response](#). In preparation.
- 2023 Co-Principal Investigator, with Andrew V. Papachristos. [The Media Accountability Project: Understanding Racial Disparity in Media Depictions of Gun Violence](#). Community Justice Action Fund. Total award amount \$104k.
- 2021-2024 Co-Principal Investigator, with Nusret Sahin. [Enhancing Procedural Justness of Encounters Through Substantiation \(EPJETS\): The Atlantic County Randomized Controlled Trial](#). National Institute of Justice. Total award amount \$703k.

As Co-Investigator, Advisor, or Personnel

- 2024-2028
under review Co-Investigator, with Vijay Mittal. [Precursors and Trajectory of Early Stage Biobehavioral Indicators of Psychosis](#). Under review at the National Institute of Mental Health, NIH.
- 2024-2028 Co-Investigator, with Molly Losh. [A cross-linguistic, family-genetic study of language in autism and the broad autism phenotype](#). National Institute on Deafness and Other Communication Disorders, NIH. Total award amount \$4M.
- 2023-2025 Co-Investigator, with V.S. Subramanian and Brayden King. [AI and Social Movements Working Group](#). Northwestern Buffett Institute for Global Affairs. Total award amount \$300k.
- 2023-2027 Senior Advisor. PI Rebecca Neusteter. [911, What's Your Emergency?: Factors Impacting Call Coding](#). National Institute of Justice. Total award amount \$500k.
- 2023-2028 Co-Investigator, with Molly Losh. [A Family-Genetic Study of Language in Autism](#). National Institute on Deafness and Other Communication Disorders, NIH. Total award amount \$4M.
- 2021-2024 Senior Personnel. PI Suzan van der Lee. [HDR DSC: Collaborative Research: The Metropolitan Chicago Data Science Corps \(MCDC\): Learning from Data to Support Communities](#). National Science Foundation. Total award amount \$904k.

PUBLICATIONS

LEGEND: * equal co-authorship (first or senior) † student authors with me as primary advisor

under review

Jacob Katz†, Nusret Sahin, and **Rob Voigt**. *Quantifying Script Adherence in Behavioral RCTs*.

Luz Maria Allende, **Rob Voigt**, Jason Schiffman, and Vijay Mittal. *Psychosis and Schizophrenia Stigma in Media: a Comprehensive Natural Language Based Approach*.

Joseph C.Y. Lau, Emily Landau, Qingcheng Zeng†, Ruichun Zhang, Stephanie Crawford, **Rob Voigt***, Molly Losh*. *Short Report: Pre-trained AI language model represents pragmatic language variability central to autism and genetically related conditions*. Revise-and-resubmit at *Autism*.

forthcoming

Qingcheng Zeng†, Mingyu Jin, and **Rob Voigt**. *Adaptive Axes: A Pipeline for In-domain Social Stereotype Analysis*. Forthcoming at Empirical Methods for Natural Language Processing (EMNLP 2024).

Mourad Heddaya, Qingcheng Zeng†, Chenhao Tan*, **Rob Voigt***, and Alexander Zentefis*. *Causal Micro-Narratives*. Forthcoming at Empirical Methods for Natural Language Processing, Workshop on Narrative Understanding (WNU 2024).

Nicholas P. Camp and **Rob Voigt**, *Body Camera Footage as Data: Using Natural Language Processing to Capture Policing At Scale and In Depth*. Forthcoming at Behavioral Science & Policy.

C. Michael Senko† and **Rob Voigt**. *Transgressing the binary: A computational approach to gendered reference*. Forthcoming in *Battlefield Linguistics: Contemporary contestations of language, gender, and sexuality*, De Gruyter Mouton.

2024 Nicholas P. Camp*, **Rob Voigt***, MarYam G. Hamedani, Dan Jurafsky, and Jennifer L. Eberhardt. *Leveraging body-worn camera footage to assess the effects of training on officer communication during traffic stops*. PNAS Nexus.

Lucas Zurbuchent† and **Rob Voigt**. *A Computational Analysis and Exploration of Linguistic Borrowings in French Rap Lyrics*. Proceedings of the Association for Computational Linguistics Student Research Workshop (ACL SRW 2024).

2023 Ran Abramitzky*, Leah Platt Boustan*, Peter Catron*, Dylan Connor* and **Rob Voigt***. *The Refugee Advantage: English-Language Attainment in the Early Twentieth Century*. Sociological Science.

Qingcheng Zeng†, Lucas Garay, Peilin Zhou, DaDing Chong, Yining Hua, Jiageng Wu, Yikang Pan, Han Zhou, **Rob Voigt**, and Jie Yang. *GreenPLM: Cross-Lingual Transfer of Monolingual Pre-Trained Language Models at Almost No Cost*. Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2023).

Mourad Heddaya, Solomon Dworkin, Chenhao Tan*, **Rob Voigt***, and Alexander Zentefis*. *The Language of Bargaining*. Proceedings of the Association for Computational Linguistics (ACL 2023).

Suet-Ying Lam*, Qingcheng Zeng†*, Kexun Zhang*, Chenyu You and **Rob Voigt**. *Large Language Models Are Partially Primed in Pronoun Interpretation*. Findings of the Association for Computational Linguistics (ACL 2023 Findings).

2022 Dallas Card, Serina Chang, Chris Becker, Julia Mendelsohn, **Rob Voigt**, Leah Boustan, Ran Abramitzky, Dan Jurafsky. *Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration*. Proceedings of the National Academy of Sciences (PNAS).

- 2021 Nicholas P. Camp, **Rob Voigt**, Dan Jurafsky, and Jennifer L. Eberhardt. **The Thin Blue Waveform: Racial Disparities in Officer Prosody Undermine Institutional Trust in the Police**. *Journal of Personality and Social Psychology (JPSP)*.
- 2019 Dorottya Demszky, Nikhil Garg, **Rob Voigt**, James Zou, Matthew Gentzkow, Jesse Shapiro, and Dan Jurafsky. **Analyzing Polarization in Social Media: Method and Application to Tweets on 21 Mass Shootings**. *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2019)*.
- 2018 **Rob Voigt**, David Jurgens, Vinodkumar Prabhakaran, Dan Jurafsky, and Yulia Tsvetkov. 2018. **RtGender: A Corpus of Responses to Gender for Studying Gender Bias**. *Proceedings of the Language Resources and Evaluation Conference (LREC 2018)*.
- 2017 **Rob Voigt**, Nicholas P. Camp, Vinodkumar Prabhakaran, William L. Hamilton, Rebecca C. Hetey, Camilla M. Griffiths, David Jurgens, Dan Jurafsky, and Jennifer L. Eberhardt. 2017. **Language from police body camera footage shows racial disparities in officer respect**. *Proceedings of the National Academy of Sciences (PNAS)*. *Cozzarelli Prize*. *Cialdini Prize*.
- 2016 **Rob Voigt**, Penelope Eckert, Dan Jurafsky, and Robert J. Podesva. **Cans and cants: Computational potentials for multimodality with a case study in head position**. *Journal of Sociolinguistics* 20 (5), 677-711.
- Rob Voigt**, Dan Jurafsky, and Meghan Sumner. **Between- and Within-Speaker Effects of Bilingualism on F0 Variation**. *Proceedings of INTERSPEECH 2016*, 1122-1126.
- 2015 **Rob Voigt** and Dan Jurafsky, 2015. **The Users Who Say "Ni": Audience Identification in Chinese-language Restaurant Reviews**. *Proceedings of the Association of Computational Linguistics (ACL 2015)*.
- Robert J. Podesva, Patrick Callier, **Rob Voigt**, and Dan Jurafsky. **The Connection Between Smiling and GOAT Fronting: Embodied Affect in Sociophonetic Variation**. *Proceedings of the International Congress of Phonetic Sciences 18 (ICPHS)*.
- 2014 Mengqiu Wang, **Rob Voigt**, and Christopher D. Manning. **Two Knives Cut Better Than One: Chinese Word Segmentation with Dual Decomposition**. *Proceedings of the Association of Computational Linguistics (ACL 2014)*, Short Papers.
- Rob Voigt**, Robert J. Podesva and Dan Jurafsky. **Speaker Movement Correlates with Prosodic Indicators of Engagement**. *Proceedings of Speech Prosody 7*, 70-74. *2nd Place Best Paper*.
- 2013 Spence Green, Daniel Cer, Kevin Reschke, **Rob Voigt**, John Bauer, Sida Wang, Natalia Silveira, Julia Neidert, and Christopher D. Manning, 2013. **Feature-Rich Phrase-based Translation: Stanford University's Submission to the WMT 2013 Translation Task**. *Proceedings of the Eighth Workshop on Statistical Machine Translation (WMT 2013)*.
- Rob Voigt** and Dan Jurafsky. **Tradition and Modernity in 20th Century Chinese Poetry**. *Proceedings of the North American Association for Computational Linguistics, Second Workshop on Computational Linguistics for Literature (CLfL 2013)*.
- 2012 **Rob Voigt** and Dan Jurafsky. **Towards a Literary Machine Translation: The Role of Referential Cohesion**. *Proceedings of the North American Association for Computational Linguistics, Workshop on Computational Linguistics for Literature (CLfL 2012)*.

SELECTED TALKS AND PRESENTATIONS

- 2024 **Rob Voigt**, "Multimodality and 'Slow Big Data'." Invited talk, **Summer Institute in Computational Social Science, Chicago**. August 2024.
- 2022 **Rob Voigt**, "Computational Linguistics and the Linguistic Mechanisms of Procedural Justice." Colloquium, University of Texas at Austin, Linguistics Department. April 2022.
- 2021 **Rob Voigt**, "Language and Immigration: Historical Views from Personal and Political Speech." Colloquium, University of Chicago Language Evolution, Acquisition & Processing Workshop, November 2021.
- Rob Voigt**, "Developing a Linguistically-Informed Procedural Justice Training and Quantifying its Impact." Panel Presentation, Association of Applied Linguistics, February 2021.
- 2018 Yulia Tsvetkov, Vinodkumar Prabhakaran, and **Rob Voigt**, "**Socially Responsible Natural Language Processing**." Tutorial Presentation, North American Association for Computational Linguistics. June 2018, New Orleans, Louisiana.

ADVISING

Current

- 2021-present [Grace LeFevre](#), PhD in Linguistics. Pre-candidacy.
 2021-present [Adrian Ray Avalani](#), PhD in Linguistics. Pre-candidacy.
 2022-present [Qingcheng Zeng](#), PhD in Linguistics. Pre-candidacy.
 2022-present [Lynna Neri](#), PhD in Linguistics. Pre-candidacy.
 2023-present [Nneoma Udeze](#), PhD in Linguistics. Pre-candidacy.

Alumni

- 2023-2024 Lucas Zurbuchen, MS in Computer Science.
 Undergraduate Honors Thesis: *A Computational Analysis and Exploration of Linguistic Borrowings in French Rap Lyrics*
- 2023-2024 Jacob Katz, BA in Linguistics.
 Undergraduate Honors Thesis: *Quantifying Script Adherence in Behavioral RCTs with Large Language Models: A Study on Procedural Justice in Policing*
- 2021-2022 C. Michael Senko, MA in Linguistics.
 MA Thesis: *Transgressing the binary: Gendered language practices on Twitter*

TEACHING

At Northwestern

- Winter 2024 Instructor, [Introduction to Computational Linguistics](#).
 Fall 2023 Instructor, [Seminar in Linguistics: Computational Methods for Social Understanding](#).
 Spring 2023 Instructor, [Introduction to Computational Linguistics](#).
 Winter 2023 Instructor, [Text Processing for Linguists](#).
 Winter 2022 Instructor, [Seminar in Linguistics: Computation and Interaction](#).
 Fall 2021 Instructor, [Introduction to Computational Linguistics](#).
 Spring 2021 Instructor, [Introduction to Computational Linguistics](#).
 Winter 2021 Instructor, [Text Processing for Linguists](#).
 Spring 2020 Instructor, [Text Processing for Linguists](#).

At Stanford

- 2018 Teaching Assistant, *From Language to Information*, with Dan Jurafsky.
 2018 Research Mentor, [AI4ALL Summer Program](#).
 2016 Teaching Assistant, *Language and Gender*, with Penelope Eckert.
 2015-2016 Linguistics Corpus Teaching Assistant, with Christopher D. Manning.

ACADEMIC SERVICE

<i>2014-present</i>	Program Committee Member and Reviewer for ACL, EMNLP, NAACL, EACL, LREC, COLING; workshops on Computational Linguistics for Literature, Ethics in NLP, Coreference Beyond Ontonotes, Abusive Language Workshop, Online Abuse and Harms, NLP for Positive Impact
<i>2020-present</i>	Journal standing reviewer, <i>Computational Linguistics</i>
<i>2023</i>	Area Chair, Computational Social Science and Cultural Analytics Track, <i>EMNLP 2023</i> . Singapore.
<i>2021</i>	Area Chair, NLP for Social Good Theme Track, <i>ACL-IJCNLP 2021</i> . Bangkok, Thailand.
<i>2021</i>	Mentor (for Syed Haque), ACL Mentoring Reviews. Association for Computational Linguistics (ACL) 2021.
<i>2020</i>	Mentor (for Amanda Buddemeyer), <i>ACL Student Research Workshop</i> . Association for Computational Linguistics (ACL) 2020, Seattle, Washington.
<i>2018</i>	Co-organizer, <i>ACL Student Research Workshop</i> . Association for Computational Linguistics (ACL) 2018, Melbourne, Australia.
<i>2018</i>	Co-organizer, <i>Second Workshop on Abusive Language Online</i> . Empirical Methods in Natural Language Processing (EMNLP) 2018, Brussels, Belgium.
<i>2015-2016</i>	Graduate Student Representative, Stanford Linguistics
<i>2015-2016</i>	Student Member, Faculty Senate Committee on Research, Stanford University

OTHER SKILLS

Computational Python (PyTorch, scikit-learn, pandas, numpy, HuggingFace), R, Java, C++, L^AT_EX, Unix
 Natural language processing, machine learning, computer vision, neural networks, statistics
 Web development with Python, JavaScript, HTML5 (for example, nlp.stanford.edu)

Natural Languages MANDARIN CHINESE · Fluent
 JAPANESE · Advanced
 HOKKIEN CHINESE, SPANISH, SWAHILI, PORTUGUESE · Basic