

Alex Warstadt

warstadt@nyu.edu



PhD student
New York University
Department of Linguistics
10 Washington Place, Rm. 603
New York, NY 10003

Date: 24 October 2021

Education

- 2017–pres. PhD Student in Linguistics, New York University**
Advisor: Samuel R. Bowman
- 2011–2015 A.B. Linguistics, Brown University**
magna cum laude, with honors, GPA – 3.97
Thesis title: *Right-node Wrapping* [\[pdf\]](#), Advisor: Pauline Jacobson
- 2011–2015 A.B. Music, Brown University**
magna cum laude, GPA – 3.97
Music theory, composition, and history of Western music

Peer Reviewed Papers

^J=journal publication

- 2021 NOPE: A Corpus of Naturally-Occurring Presuppositions in English.**
{Alicia Parrish, Sebastian Schuster, Alex Warstadt}, Omar Agha,
Soo-Hwan Lee, Zhuoye Zhao, Samuel R. Bowman, Tal Linzen.
Proceedings of CONLL. [\[link\]](#)
- 2021 What Ingredients Make for an Effective Crowdsourcing Protocol for
Difficult NLU Data Collection Tasks?** Nikita Nangia, Saku Sugawara,
Harsh Trivedi, Alex Warstadt, Clara Vania, Samuel R. Bowman.
Proceedings of ACL. [\[link\]](#)

- 2021** **Does Putting a Linguist in the Loop Improve NLU Data Collection?**
Alicia Parrish, William Huang, Omar Agha, Soo-Hwan Lee, Nikita Nangia, [Alex Warstadt](#), Karmanya Aggarwal, Emily Allaway, Tal Linzen, Samuel R. Bowman. *Findings of EMNLP*. [[link](#)]
- 2021** **CLiMP: A Benchmark for Chinese Language Model Evaluation.** Beilei Xiang, Changbing Yang, Yu Li, [Alex Warstadt](#), Katharina Kann. *Proceedings of EACL*. [[link](#)]
- 2020** **When Do You Need Billions of Words of Pretraining Data?** {Yian Zhang, [Alex Warstadt](#)}, Xiaocheng Li, Samuel R Bowman. *Proceedings of ACL*. [[link](#)]
- 2020** **Learning which features matter: RoBERTa acquires a preference for linguistic generalizations (eventually).** [Alex Warstadt](#), Yian Zhang, Xiaocheng Li, Haokun Liu, Samuel R Bowman. *Proceedings of EMNLP*. [[link](#)]
- 2020** **"Just" don't ask: Exclusives and potential questions.** [Alex Warstadt](#) *Proceedings of Sinn und Bedeutung 24*. [[link](#)]
- 2020** **Non-resolving responses to polar questions: A revision to the QUD theory of relevance.** Omar Agha, [Alex Warstadt](#). *Proceedings of Sinn und Bedeutung 24*. [[link](#)]
- 2020** **Can neural networks acquire a structural bias from raw linguistic data?** [Alex Warstadt](#), Samuel R. Bowman. *Proceedings of the Annual Meeting of the Cognitive Science Society*. [[link](#)]
- 2020** **Are natural language inference models IMPPRESSive? Learning IMPLicature and PRESupposition.** {Paloma Jeretič, [Alex Warstadt](#)}, Suvrat Bhooshan, Adina Williams. *Proceedings of ACL*. [[link](#)]
- 2020** **BLiMP: The Benchmark of Minimal Pairs for English.** [Alex Warstadt](#), Alicia Parrish, Haokun Liu, Anhad Mohananey, Wei Peng, Sheng-Fu Wang, Samuel R Bowman. *Transactions of the Association for Computational Linguistics (TACL)*. [[link](#)]
- 2019** **Investigating BERT's Knowledge of Language: Five Analysis Methods with NPIs.** {[Alex Warstadt](#), Yu Cao, Ioana Grosu, Wei Peng, Hagen Blix, Yining Nie, Anna Alsop, Shikha Bordia, Haokun Liu, Alicia Parrish, Sheng-Fu Wang, Jason Phang, Anhad Mohananey, Phu Mon

Htut, Paloma Jeretič}, and Samuel R Bowman. *Proceedings of EMNLP*. [\[link\]](#)

- 2019 **Verb Argument Structure Alternations in Word and Sentence Embeddings.** {Katharina Kann, [Alex Warstadt](#), Adina Williams}, Samuel R. Bowman. *Proceedings of SCiL 2*. [\[link\]](#)
- 2019 **Neural Network Acceptability Judgments.** [Alex Warstadt](#), Amanpreet Singh and Samuel R. Bowman. *Transactions of the Association for Computational Linguistics (TACL)*. [\[link\]](#)
- 2015 **Right-node wrapping: A combinatory account.** [Alex Warstadt](#). In Proceedings for ESSLI 2015, Workshop 'Empirical Advances in Categorical Grammar'. [\[link\]](#)

Manuscripts

- 2019 **Linguistic Analysis of Pretrained Sentence Encoders with Acceptability Judgments.** [Alex Warstadt](#) & Samuel R. Bowman. *ArXiv preprint*. [\[link\]](#)
- 2015 **The Syntax of Coordination and Discontinuity in Categorical Grammar.** [Alex Warstadt](#). Unpublished manuscript, presented at Tri-College Undergraduate Linguistics Conference. [\[link\]](#)

Presentations (Refereed)

- 2021 **NOPE: A Corpus of Naturally-Occurring Presuppositions in English.** Co-presented with Alicia Parrish, Sebastian Schuster. Work also with Omar Agha, Soo-Hwan Lee, Zhuoye Zhao, Samuel R. Bowman, Tal Linzen. *Talk given at CoNLL (Punta Cana)*.
- 2021 **Testing gradient measures of relevance in discourse.** Co-presented with Omar Agha. *Talk given at Sinn und Bedeutung 26 (Cologne)*
- 2021 **The Dynamics of Refusal: A revision to the QUD theory of discourse structure.** *Talk given at Semantics and Linguistic Theory (SALT) 31 (Brown University)*.
- 2021 **When Do You Need Billions of Words of Pretraining Data?** Co-presented with Yian Zhang. Work also with Xiaocheng Li, Samuel R Bowman. *Poster presented at ACL (Bangkok/Online)*.

- 2020** **Learning which features matter: RoBERTa acquires a preference for linguistic generalizations (eventually).** Work with Yian Zhang, Xiaocheng Li, Haokun Liu, Samuel R Bowman. *Talk given at EMNLP (Punta Cana/Online).*
- 2020** **Can neural networks acquire a structural bias from raw linguistic data?** Work with Samuel R. Bowman. *Poster presented at the 42nd Annual Meeting of the Cognitive Science Society (Toronto/Online).*
- 2020** **Are natural language inference models IMPPRESsive? Learning IMPIicature and PRESupposition.** Co-presented with Paloma Jeretic. Work also with Suvrat Bhooshan, Adina Williams. *Talk given at ACL (Seattle/Online).*
- 2020** **BLIMP: The Benchmark of Minimal Pairs for English.** [\[pdf\]](#) Work also with Alicia Parrish Haokun Liu, Anhad Mohananey, Wei Peng, Sheng-Fu Wang, Samuel R Bowman *Talk given at EMNLP 2020 (Punta Cana/Online).*
- 2020** **BLIMP: The Benchmark of Minimal Pairs for English.** Co-presented with Alicia Parrish. Work also with Haokun Liu, Anhad Mohananey, Wei Peng, Sheng-Fu Wang, Samuel R Bowman. *Talk given at SCiL 3 (New Orleans).*
- 2019** **Investigating BERT's Knowledge of Language: Five Analysis Methods with NPIs.** Work with Yu Cao, Ioana Grosu, Wei Peng, Hagen Blix, Yining Nie, Anna Alsop, Shikha Bordia, Haokun Liu, Alicia Parrish, Sheng-Fu Wang, Jason Phang, Anhad Mohananey, Phu Mon Htut, Paloma Jeretič, and Samuel R Bowman. *Talk given at EMNLP 2019 (Hong Kong).*
- 2019** **Neural Network Acceptability Judgments.** Work with Amanpreet Singh & Samuel R Bowman. *Talk given at EMNLP 2019, Hong Kong.*
- 2019** **“Just”: Don’t Answer the Question.** *Talk given at Sinn und Bedeutung 24 (Osnabrück).*
- 2019** **Non-Resolving Responses to Polar Questions.** Co-presented with Omar Agha. *Poster presented at Sinn und Bedeutung 24 (Osnabrück).*
- 2019** **Verb Argument Structure Alternations in Word and Sentence Embeddings.** Co-presented with Katharina Kann and Adina Williams, and Samuel R Bowman. *Poster presented at SCiL 2 (New York).*
- 2019** **Non-Resolving Responses to Polar Questions.** Co-presented with

Omar Agha. *Talk given at SYNC (SUNY Stony Brook).*

2018 ***Just Refuses to Answer the QUD.*** *Talk given at NYU Semantics Group.*

2018 **Acceptability Judgments from a Neural Network.** Joint work with Samuel R. Bowman. *Talk given at Annual Meeting of the LSA.*

2015 **Right-node wrapping: A combinatory account.** *Talk given at ESSLLI Workshop “Empirical Advances in Categorical Grammar”.*

2015 **The Syntax of Coordination and Discontinuity in a Combinatory Categorical Grammar.** *Talk given at Tri-College Undergraduate Linguistics Conference.*

Invited Talks

2021 **Testing the learnability of grammar for humans and machines: Investigations with artificial neural networks.** New York University Center for Data Science Text-as-Data series.

2020 **Learning Which Features Matter: RoBERTa Acquires a Preference for Linguistic Generalizations (Eventually).** Bar-Ilan University NLP group seminar.

2020 **Learning grammar from raw data: Case studies with acceptability judgments and neural networks.** Indiana University ClingDing discussion group.

2020 **Learning Acceptability Judgments from Text Alone.** Max Planck–NYU Center for Language, Music, and Emotion (CLaME) workshop on Language. (Stand-in for Samuel R Bowman, who was the invited speaker and wrote the talk)

Awards & Grants

2021 **Outstanding reviewer for ACL**

2019 **NSF Computer and Information Science and Engineering Research Initiation Initiative**
\$175,000, 24 months
For: Can Low-Bias Machine Learners Acquire English Grammar?

Deep Learning and Linguistic Acceptability.
Samuel R. Bowman (PI), Alex Warstadt

- 2017-2022** **McCracken Fellowship**
5 years, NYU internal
- 2015** **Outstanding Work in the Linguistics Concentration**
- 2015** **Sigma Xi scientific research society**
- 2014** **Phi Beta Kappa**
- 2014** **Undergraduate Teaching and Research Award (UTRA)**
\$3570, 3 months, Brown University internal
- 2011** **Valedictorian, Sharon High School**

Research-Related Work

- 2018** **Lambda Calculator, Java development**
under Prof. Lucas Champollion
Computer programming for automatic semantic formula parsing.
- 2016** **Sentence and Discourse Processing Lab, Brown University**
under Dr. Laura Kertz.
Assisted with experimental design and the design of stimuli, ran
experiment with participants

Teaching Experience

- 2021** **Guest Lecturer**, Natural Language Understanding and Computational
Semantics. Professor Samuel R Bowman
- 2021** **Guest Lecturer**, Machine Learning and Language Understanding.
Professor Samuel R Bowman.
- 2020** **Teaching Assistant**, Patterns in Language, NYU. Professor Lucas
Champollion.
- 2020** **Primary Instructur**, Language, NYU. Co-taught with Paloma Jeretič.

2019 **Teaching Assistant, Guest Lecturer**, Language, NYU. Professor Gary Thoms.

Service

2018-2021 **Served on Colloquium Committee, NYU Linguistics**

Reviewing

- 2021 Association for Computational Linguistics Rolling Review. October cycle.
- 2021 The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP).
- 2021 Frontiers in Psychology, section Language Sciences.
- 2021 Society for Computation in Linguistics (SCiL).
- 2021 The SIGNLL Conference on Computational Natural Language Learning (CoNLL).
- 2021 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT).
- 2020. Language and Cognition.
- 2020. The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).
- 2020. The 24th Conference on Computational Natural Language Learning (CoNLL).
- 2020. The 58th Annual Meeting of the Association for Computational Linguistics (ACL).

Industry Experience

2016-2017 **R&D Engineer, IPsoft, Inc., New York, NY**

Implemented conversation engine Amelia for Swedish and German, developed language-agnostic syntactic representation for NLP.

2015-2016 **Java Developer, IBM Watson Health, Yorktown Heights, NY**
Educated team members in and documented NLP algorithms in medical text analysis pipeline.

Technologies

Expert	Python, Java, LaTeX, Regex
Proficient	NLTK, Pytorch, Linux, Git, WebPPL
Beginner	R, Scala, ML

Languages

Native	English
Proficient	Spanish
Beginner	German
Fieldwork	Q'anjob'al (Mayan)