

Sebastian Schuster

CONTACT INFORMATION	Center for Data Science New York University 60 Fifth Avenue New York, NY 10011	<i>E-mail:</i> schuster@nyu.edu <i>Tel:</i> +1 650 862 3589 <i>Website:</i> http://www.sebschu.com
EMPLOYMENT	Postdoctoral Scholar Center for Data Science & Department of Linguistics, New York University	since January 2021
	Postdoctoral Scholar Center for the Study of Language and Information, Stanford University	September - December 2020
EDUCATION & TRAINING	Stanford University , Stanford, CA, USA <i>PhD</i> , Linguistics Dissertation: Semantic-pragmatic adaptation Committee: Judith Degen (chair), Christopher Potts, Daniel Lassiter	September 2020
	Summer School on Statistical Methods for Linguistics and Psychology , Potsdam, Germany	September 2018
	2015 LSA Linguistic Institute , Chicago, Illinois, USA	July 2015
	Stanford University , Stanford, CA, USA <i>Master of Science</i> , Computer Science Advisor: Christopher Manning	December 2014
	University of Vienna , Vienna, Austria <i>Bachelor of Science</i> , Computer Science Thesis: The influence of training data selection on the quality of a model Advisor: Ingo Ebersberger	February 2012
NON-ACADEMIC RESEARCH	Research Intern , Conversational AI Team, Facebook Inc., Menlo Park, CA, USA	June 2018 - December 2018
AWARDS AND FELLOWSHIPS	Computing Innovation Fellow (2-year postdoctoral fellowship) CRA, CCC and NSF	July 2020
	GRO Research Grant School of Humanities and Sciences, Stanford University	December 2019
	Community Impact Award , Stanford Alumni Association	May 2019
	Pigott Fellowship , School of Humanities and Sciences Stanford University	March 2018
	Travel award , 14th International Congress for the Study of Child Language	July 2017
	Goodan Family Graduate Fellowship School of Humanities and Sciences, Stanford University	2015 - 2020

Best rated poster presentation award, 2014 IEEE Spoken Language Technology Workshop December 2014
Merit scholarship, University of Vienna 2011 - 2012

WORK IN
PROGRESS

Jennifer Hu, Roger Levy, and **Sebastian Schuster**. (to appear). *Predicting scalar diversity with context-driven uncertainty over alternatives*.

Sebastian Schuster and Tal Linzen. (under review). *When a sentence does not introduce a discourse entity, Transformer-based models still refer to it*.

Sebastian Schuster and Najoung Kim. (in prep.). *Can pre-trained language models track the state of entities in text?*

Sebastian Schuster, Matthew Loder, and Judith Degen. (in prep.). *Explaining away in semantic/pragmatic adaptation*.

Masoud Jasbi and **Sebastian Schuster** (in prep.). *Open Linguistics: A platform for sharing informal judgment data and analysis*.

PEER-REVIEWED
PUBLICATIONS

Aaron Mueller, Robert Frank, Tal Linzen, Luheng Wang, and **Sebastian Schuster**. 2022. *Coloring the Blank Slate: Pre-training Imparts a Hierarchical Inductive Bias to Sequence-to-sequence Models*. In *Findings of the Association of Computational Linguistics (ACL 2022)*.

Soo-Hwan Lee and **Sebastian Schuster**. 2022. *Can language models capture syntactic associations without surface cues? A case study of reflexive anaphor licensing in English control constructions*. In *Proceedings of the Society for Computation in Linguistics 5 (SCiL 2022)*.

Alicia Parrish*, **Sebastian Schuster***, Alex Warstadt*, Omer Agha, Zhuoye Zhang, Soo-Hwan Lee, Samuel R. Bowman, and Tal Linzen. 2021. *NOPE: A corpus of naturally-occurring presuppositions in English*. In *Proceedings of the 24th Conference on Computational Natural Language Learning (CoNLL 2021)*.

Elissa Li, **Sebastian Schuster** and Judith Degen. 2021. *Predicting scalar inferences from “or” to “not both” using neural sentence encoders*. In *Proceedings of the Society for Computation in Linguistics 4 (SCiL 2021)*.

Sebastian Schuster and Judith Degen. 2020. *I know what you’re probably going to say: Listener adaptation to variable use of uncertainty expressions*. *Cognition*, 203.

Sebastian Schuster*, Yuxing Chen*, and Judith Degen. 2020. *Harnessing the linguistic signal to predict scalar inferences*. In *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher

D. Manning, Sampo Pyysalo, **Sebastian Schuster**, Francis Tyers, and Daniel Zeman. 2020. *Universal Dependencies v2: An evergrowing multilingual treebank collection*. In *Proceedings of the Twelfth International Conference on Language Resources and Evaluation (LREC 2020)*.

Sebastian Schuster and Judith Degen. 2019. *Speaker-specific adaptation to variable use of uncertainty expressions*. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)*.

Sebastian Schuster, Sonal Gupta, Rushin Shah, and Mike Lewis. 2019. *Cross-lingual Transfer Learning for Multilingual Task Oriented Dialog*. In *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2019)*.

Joakim Nivre, Paola Marongiu, Filip Ginter, Jenna Kanerva, Simonetta Montemagni, **Sebastian Schuster**, and Maria Simi. 2018. *Enhancing Universal Dependency Treebanks: A Case Study*. In *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)*.

Sebastian Schuster, Joakim Nivre, and Christopher D. Manning. 2018. *Sentences with Gapping: Parsing and Reconstructing Elided Predicates*. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2018)*.

Daniel Zeman et al. 2017. *CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies (CONLL 2017)*.

Sebastian Schuster, Éric Villemonte de la Clergerie, Marie Candito, Benoît Sagot, Christopher D. Manning, and Djamé Seddah. 2017. *Paris and Stanford at EPE 2017: Downstream Evaluation of Graph-based Dependency Representations*. In *Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation (EPE 2017)*.

Sebastian Schuster, Matthew Lamm, and Christopher D. Manning. 2017. *Gapping Constructions in Universal Dependencies v2*. In *Proceedings of the First Universal Dependencies Workshop (UDW 2017)*.

Sebastian Schuster and Christopher D. Manning. 2016. *Enhanced English Universal Dependencies: An Improved Representation for Natural Language Understanding Tasks*. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*.

Sebastian Schuster, Ranjay Krishna, Angel Chang, Li Fei-Fei, and Christopher D. Manning. 2015. *Generating Semantically Precise Scene Graphs from Textual Descriptions for Improved Image Retrieval*. In *Proceedings of the Fourth*

Workshop on Vision and Language (VL15).

Sebastian Schuster,* Stephanie Pancoast,* Milind Ganjoo, Michael C. Frank, and Dan Jurafsky. 2014. *Speaker-Independent Detection of Child Directed Speech*. In *Proceedings of the 2014 IEEE Spoken Language Technology Workshop (SLT 2014)*. **Best Rated Poster Presentation Award**.

Spence Green, Sida I. Wang, Jason Chuang, Jeffrey Heer, **Sebastian Schuster**, and Christopher D. Manning. 2014. *Human Effort and Machine Learnability in Computer Aided Translation*. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*.

Julia Neidert,* **Sebastian Schuster**,* Spence Green, Kenneth Heafield, and Christopher D. Manning. 2014. *Stanford University's Submissions to the WMT 2014 Shared Task*. In *Proceedings of the Ninth Workshop On Statistical Machine Translation (WMT 2014)*.

PRESENTATIONS

Sebastian Schuster. Invited Talk: *How contextual are contextual language models?*. NYU Text-as-Data talk series. NYU, New York, NY, March 31, 2022.

Sebastian Schuster. Colloquium: *Modeling context-sensitive utterances*. Linguistics Department, Yale University, New Haven, CT, March 11, 2022.

Sebastian Schuster. Colloquium: *Modeling context-sensitive utterances*. Linguistics Department, Northwestern University, Evanston, IL, March 3, 2022.

Soo-Hwan Lee and Sebastian Schuster. *Can language models capture syntactic associations without surface cues? A case study of reflexive anaphor licensing in English control constructions*. Annual Meeting of the Society for Computation in Linguistics (SCiL 2022). online, February 7-9, 2021.

Alicia Parrish, Sebastian Schuster, Alex Warstadt, Omer Agha, Zhuoye Zhang, Soo-Hwan Lee, Samuel R. Bowman, and Tal Linzen. *NOPE: A corpus of naturally-occurring presuppositions in English*. 24th Conference on Computational Natural Language Learning (CoNLL 2021), Punta Cana, Dominican Republic, November 11, 2021.

Sebastian Schuster. Invited Talk: *Can neural networks learn to predict pragmatic inferences?*. Zurich NLP Meetup. online, May 5, 2021.

Elissa Li, Sebastian Schuster and Judith Degen. Poster: *Predicting scalar inferences from “or” to “not both” using neural sentence encoders*. Annual Meeting of the Society for Computation in Linguistics (SCiL 2021). online, February 14-19, 2021.

Sebastian Schuster, Matthew Loder, and Judith Degen. Poster: *Explaining away in semantic/pragmatic adaptation*. Architectures and Mechanisms for Language

Processing 2020 (AMLaP 2020). online, September 3-5, 2020.

Sebastian Schuster, Yuxing Chen, and Judith Degen. *Harnessing the linguistic signal to predict scalar inferences*. 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020), online, July 6-8, 2020.

Sebastian Schuster. Colloquium: *Modeling utterance interpretation in context*. Linguistics Department, CU Boulder, Boulder, CO, February 17, 2020.

Sebastian Schuster. Invited Talk: *Computational Models of Pragmatics*. Computation and Psycholinguistics Lab, John Hopkins University, Baltimore, MD, November 20, 2019.

Sebastian Schuster, Matthew Loder, and Judith Degen. Poster: *Explaining Away in Semantic Adaptation*. California Meeting on Psycholinguistics 2019 (CAMP 2019). Santa Cruz, CA, October 26-27, 2019.

Sebastian Schuster. Invited Talk: *Cross-lingual Transfer Learning for Multilingual Task-Oriented Dialog*. Samsung Research, Mountain View, CA, October 3, 2019.

Sebastian Schuster and Judith Degen. Poster: *Speaker-specific adaptation to variable use of uncertainty expressions*. 41st Annual Meeting of the Cognitive Science Society (CogSci 2019), Montréal, Canada, July 24-27, 2019.

Sebastian Schuster and Judith Degen. *Speaker-specific adaptation to variable use of uncertainty expressions*. 8th Experimental Pragmatics conference (XPrag 2019), Edinburgh, UK, June 19-21, 2019.

Sebastian Schuster and Judith Degen. Poster: *A computational model of listener adaptation to variable use of uncertainty expressions*. Cognitive Modeling and Computational Linguistics (CMCL 2019), Minneapolis, MN, June 7, 2019.

Sebastian Schuster, Sonal Gupta, Rushin Shah, and Mike Lewis. Poster: *Cross-lingual Transfer Learning for Multilingual Task Oriented Dialog*. 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2019), Minneapolis, MN, June 3-5, 2019.

Masoud Jasbi, Sebastian Schuster, and Philip Weiss. *Open Linguistics: A platform for informal acceptability data*. Second North American Conference in Iranian Linguistics (NACIL2), Tucson, AZ, April 19-21, 2019.

Sebastian Schuster and Judith Degen. Poster: *A computational model of listener adaptation to variable use of uncertainty expressions*. Stanford HAI Symposium, Stanford, CA, March 18, 2019.

Sebastian Schuster and Judith Degen. *Adaptation to Variable Use of Expressions*

of Uncertainty. 20th Annual SemFest, Stanford, CA, March 15-16, 2019.

Sebastian Schuster, Sonal Gupta, Rushin Shah, and Mike Lewis. *Cross-lingual Transfer Learning for Multilingual Task Oriented Dialog*. Second Conversational AI Workshop at NeurIPS 2018. Montréal, Canada, December 7, 2018.

Sebastian Schuster and Judith Degen. *Adaptation to Variable Use of Expressions of Uncertainty*. California Meeting on Psycholinguistics 2018 (CAMP 2018). Los Angeles, CA, December 1-2, 2018.

Sebastian Schuster and Judith Degen. *Listener adaptation to variable use of uncertainty expressions*. Computational Models of Language Generation and Processing in Pragmatics (CoMPPrag 2018). Bochum, Germany, September 26-28, 2018.

Sebastian Schuster and Judith Degen. *Adaptation to Variable Use of Expressions of Uncertainty*. Architectures and Mechanisms for Language Processing 2018 (AMLaP 2018). Berlin, Germany, September 6-8, 2018.

Sebastian Schuster, Joakim Nivre, and Christopher D. Manning. *Sentences with Gapping: Parsing and Reconstructing Elided Predicates*. 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2018). New Orleans, LA, June 1-8, 2018.

Sebastian Schuster, Éric Villemonte de la Clergerie, Marie Candito, Benoît Sagot, Christopher D. Manning, and Djamé Seddah. *Paris and Stanford at EPE 2017: Downstream Evaluation of Graph-based Dependency Representations*. Fourth International Conference on Dependency Linguistics (Depling 2017) and 15th International Conference on Parsing Technologies (IWPT 2017). Pisa, Italy, September 18-22, 2017.

Sebastian Schuster. Poster: *Early Marking of Standards of Comparison in German and English*. 14th International Congress for the Study of Child Language. Lyon, France, July 17-21, 2017.

Sebastian Schuster, Matthew Lamm, Christopher D. Manning. *Gapping Constructions in Universal Dependencies v2*. First Universal Dependencies Workshop (UDW 2017). Gothenburg, Sweden, May 22, 2017.

Sebastian Schuster. Invited Talk: *Universal Dependencies for Natural Language Processing*. Nuance, Inc, Sunnyvale, CA, November 30, 2016.

Sebastian Schuster and Christopher D. Manning. Poster: *Enhanced English Universal Dependencies: An Improved Representation for Natural Language Understanding Tasks*. Tenth International Conference on Language Resources and Evaluation (LREC 2016). Portorož, Slovenia, May 25-27, 2016.

Sebastian Schuster, Ranjay Krishna, Angel Chang, Li Fei-Fei, and Christopher D. Manning. *Generating Semantically Precise Scene Graphs from Textual Descriptions for Improved Image Retrieval*. Fourth Workshop on Vision and Language (VL15) at EMNLP. Lisbon, Portugal, September 18, 2015.

Sebastian Schuster. Invited Talk: *Child-Directed Speech Detection and Large-Scale Machine Translation*. Apple Inc., Cupertino, CA, October 24, 2014.

Julia Neidert, Sebastian Schuster, Spence Green, Kenneth Heafield, and Christopher D. Manning. Poster: *Stanford University's Submissions to the WMT 2014 Shared Task*. Ninth Workshop On Statistical Machine Translation (WMT 2014). Baltimore, MD, June 26-27, 2014.

TEACHING
EXPERIENCE

As Instructor:
Introductory Machine Learning January 2022
NYU AI School 2022

Running web-based experiments January 2021
(with Judith Degen) LSA 2021 Annual Meeting

Bayesian Models of Pragmatic Inference Workshop February 2018
(with Reuben Cohn-Gordon) Stanford University

SAILORS: Stanford Artificial Intelligence Lab's Summer Outreach Program (with Abigail See) June - July 2017
Stanford University

Web-based Experimental Methods Workshop May 2017
Stanford University

As Teaching Assistant:
LING145/245A/PSYCH140: **Introduction to Psycholinguistics** Fall 2018
(Instructor: Judith Degen) Stanford University

LING130A/230A: **Introduction to Semantics & Pragmatics** Winter 2018
(Instructor: Christopher Potts) Stanford University

LING145/PSYCH140: **Introduction to Psycholinguistics** Fall 2017
(Instructor: Judith Degen) Stanford University

CS22N: **Natural Language Processing** Fall 2013
(Instructor: Christopher Manning) Stanford University

Applied Bioinformatics Fall 2011
(Instructors: Ingo Ebersberger and Arndt von Haeseler) University of Vienna

As Guest Lecturer:
LING47N: **Languages, Dialects, Speakers** November 2019

Lecture: Semantic/Pragmatic Adaptation Stanford University

332: Experimental Pragmatics July 2019

Lecture: Semantic/Pragmatic Adaptation LSA Linguistic Institute 2019

LING245B: Methods in Psycholinguistics May 2019

Lecture: Open Science & Preregistration Stanford University

LING245: Methods in Psycholinguistics May 2018

Lecture: Open Science & Preregistration Stanford University

LING1: Introduction to Linguistics October 2017

Lecture: Syntax-Mediated Multilingual Question-Answering Stanford University

As Grader:

LING1: Introduction to Linguistics Fall 2016

(Instructor: Meghan Sumner) Stanford University

**MENTORING &
ADVISING**

Aaron Mueller (JHU CS PhD student): Hierarchical bias in seq2seq models (since February 2021)

Soo-Hwan Lee (NYU linguistics PhD student): Learnability of control verb constructions (since June 2021)

Dhara Yu (Stanford CS master student): Improving contextual representations for pragmatic reasoning (January 2021 - December 2021)

Sophie Regan (Stanford SymSys undergraduate/master student): Semantic-pragmatic adaptation in preschool-aged children (since December 2019)

Elissa Li (Stanford SymSys master student): Predicting the strength of exhaustivity inferences with “or” (March 2020 - Feb 2021)

Yuxing Chen (Stanford SymSys master student): Harnessing the richness of the linguistic signal in predicting pragmatic inferences (March 2019 - December 2019)

Matthew Loder (CSLI summer intern): Explaining away in semantic adaptation (June - August 2019)

Anna-Marie Sprenger (Stanford linguistics undergraduate): Development of an English enhanced UD treebank (February - June 2018)

**PROFESSIONAL
SERVICE**

Senior Area Chair, “Language Resources and Evaluation” track, NAACL 2022

Conference Reviewer, LREC 2016, ACL 2016, ACL 2017, EMNLP 2017, Universal Dependencies Workshop 2017, CoNLL Shared Task 2017, IJCNLP 2017,

EPE 2017, LREC 2018, ACL 2018, COLING 2018, ACL Student Research Workshop 2018, Universal Dependencies Workshop 2018, EMNLP 2018 (*Outstanding Reviewer*), CAMP 2018, ACL Student Research Workshop 2019, ACL 2019, EMNLP 2019, SyntaxFest 2019, CAMP 2019, AAAI 2020, LREC 2020, ACL 2020 (*Outstanding Reviewer*), COLING 2020, CogSci 2020, ACL Student Research Workshop 2020, EMNLP 2020, AACL 2020, CoNLL 2020, NAACL 2021, NAACL Student Research Workshop 2021, ACL 2021, CogSci 2021, EMNLP 2021, ACL Student Research Workshop 2021, CoNLL 2021, ACL 2022, SCiL 2022, LREC 2022, CogSci 2022

Journal Reviewer, Corpus Linguistics and Linguistic Theory (2019), Discourse Processes (2019), Cognition (2020), Natural Language Engineering (2021), Language (2021), Glossa (2021, 2022), Journal of Experimental Psychology: Learning, Memory, and Cognition (2021), Cognitive Science (2022)

Co-chair, Universal Dependencies Workshop (COLING 2020, EMNLP 2018)

Organizer, Universal Dependencies Workshop at NoDaLiDa 2017

INSTITUTIONAL
SERVICE

QP Fest Committee, Stanford Linguistics Department, 2019-2020

CSLI Summer Internship Admissions Committee, Stanford CSLI, 2019

PhD Admissions Committee, Stanford Linguistics Department, 2018 - 2019

Graduate Studies Committee, Stanford Linguistics Department, 2016 - 2018

Graduate Student Representative, Stanford Linguistics Department, 2016 - 2017

Corpus TA, Stanford Linguistics Department, 2016 - 2017

Social Committee, Stanford Linguistics Department, 2015 - 2016

PROFESSIONAL
MEMBERSHIP

Association for Computational Linguistics	2014 - present
Linguistic Society of America	2018 - present
Cognitive Science Society	2019 - present

SELECTED SKILLS

Programming: Python, Java, JavaScript, C++, R, HTML, CSS, SQL, PHP

Machine Learning: Numpy, Scikit-learn, TensorFlow, PyTorch, NLTK, Stanford CoreNLP, Spacy, Stanza, WebPPL, Stan

LANGUAGES

German (native), **English** (fluent), **French** (intermediate), **Spanish** (intermediate)

Last updated: March 2022