

# Michael Hahn

CONTACT	Stanford University Department of Linguistics Margaret Jacks 117 Stanford CA 94305-2150, USA	E-mail: <a href="mailto:mhahn2@stanford.edu">mhahn2@stanford.edu</a> Website: <a href="http://www.mhahn.info/">www.mhahn.info/</a>
EDUCATION	<b>Stanford University</b> Ph.D. in Linguistics Dissertation: <i>Computational and Communicative Efficiency in Language</i> Committee: Judith Degen, Dan Jurafsky (co-advisors), Mike Frank	<b>2016–2022</b> <b>(expected)</b>
	<b>Massachusetts Institute of Technology</b> Department of Brain and Cognitive Sciences Visiting PhD student. Advisor: Edward Gibson	<b>2019</b>
	<b>University of Edinburgh</b> School of Informatics Visiting Research Postgraduate Student. Advisor: Frank Keller	<b>2015–2016</b>
	<b>University of Tübingen</b> M.Sc. in Mathematics	<b>2014–2015</b>
	26th European Summer School in Logic, Language and Information	<b>2014</b>
	<b>University of Tübingen</b> B.Sc. in Mathematics	<b>2011–2014</b>
	<b>University of Tübingen</b> B.A. in Computational Linguistics	<b>2010–2013</b>
NON-ACADEMIC RESEARCH	<b>Facebook AI Research</b> (Paris, France) Research Intern. Mentor: Marco Baroni	<b>2018</b>
FELLOWSHIPS AND AWARDS	Stanford University PhD Fellowship, 2016. German National Academic Foundation, 2011–2016. ( <i>Studienstiftung des Deutschen Volkes</i> )	
REFEREED JOURNAL ARTICLES	<b>in press.</b> Michael Hahn and Yang Xu. Crosslinguistic word order variation reflects evolutionary pressures of dependency and information locality. <i>Proceedings of the National Academy of Sciences of the United States of America (PNAS)</i> . <b>in press.</b> Richard Futrell and Michael Hahn. Information theory as a bridge between language function and language form. <i>Frontiers in Communication</i> . <b>2022.</b> Michael Hahn, Rebecca Mathew, and Judith Degen. Morpheme ordering across languages reflects optimization for processing efficiency. <i>Open Mind: Discoveries in Cognitive Science</i> . 5:208–232. <b>2021.</b> Michael Hahn, Dan Jurafsky, and Richard Futrell. Sensitivity as a complexity measure for sequence classification tasks. <i>Transactions of the Association for Computational Linguistics</i> , 9:891–908. <b>2021.</b> Michael Hahn, Judith Degen, and Richard Futrell. Modeling word and morpheme order in natural language as an efficient tradeoff of memory and surprisal. <i>Psychological Review</i> , 128(4):726–756.	

- 2020. Michael Hahn**, Dan Jurafsky, and Richard Futrell. Universals of word order reflect optimization of grammars for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 117(5):2347–2353.
- 2020. Michael Hahn**. Theoretical limitations of self-attention in neural sequence models. *Transactions of the Association for Computational Linguistics*, 8:156–171.
- 2019. Michael Hahn** and Marco Baroni. Tabula nearly rasa: Probing the linguistic knowledge of character-level neural language models trained on unsegmented text. *Transactions of the Association for Computational Linguistics*, 7:467–484.
- 2019. Michael Hahn** and Richard Futrell. Estimating predictive rate-distortion curves via neural variational inference. *Entropy*, 21(7):640.
- 2015. Michael Hahn** and Frank Richter. Henkin Semantics for reasoning with natural language. *Journal of Language Modeling*, 3(2):513–568.
- 2022 (accepted)**. Songpeng Yan, **Michael Hahn**, and Frank Keller. Modeling fixation behavior in reading with character-level neural attention. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society (CogSci)*.
- 2022 (accepted)**. Neil Rathi, **Michael Hahn**, and Richard Futrell. Explaining patterns of fusion in morphological paradigms using the memory–surprisal tradeoff. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society (CogSci)*.
- 2021**. Neil Rathi, **Michael Hahn**, and Richard Futrell. An information-theoretic characterization of morphological fusion. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)*, 10115–10120.
- 2020**. John Hewitt, **Michael Hahn**, Surya Ganguli, Percy Liang, and Christopher D. Manning. RNNs can generate bounded hierarchical languages with optimal memory. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)*, 1978–2010.
- 2019. Michael Hahn**, Frank Keller, Yonatan Bisk, and Yonatan Belinkov. Character-based surprisal as a model of human reading in the presence of errors. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci)*, 401–407.
- 2018. Michael Hahn**, Judith Degen, Noah Goodman, Dan Jurafsky, and Richard Futrell. An information-theoretic explanation of adjective ordering preferences. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society (CogSci)*, 1766–1771.
- 2018. Michael Hahn**, Andreas Krebs, and Howard Straubing. Wreath products of distributive forest algebras. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2018)*, 512–520.
- 2016. Michael Hahn** and Frank Keller. Modeling human reading with neural attention. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*, 85–95.
- 2015. Michael Hahn**, Andreas Krebs, Klaus-Jörn Lange, and Michael Ludwig. Visibly counter languages and the structure of  $NC^1$ . In *Mathematical Foundations of Computer Science 2015 – 40th International Symposium, MFCS 2015*, 384–394.
- 2014. Michael Hahn**. Predication and NP structure in an omnipredicative language: The case of Khoekhoe. In *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar*, 238–258. CSLI Publications.
- 2013. Michael Hahn**. Word order variation in Khoekhoe. In *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar*, 48–68. Stanford, CSLI Publications.

**2013.** Niels Ott, Ramon Ziai, **Michael Hahn**, and Detmar Meurers. CoMeT: Integrating different levels of linguistic modeling for meaning assessment. In *Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval)*, 608–616.

**2012.** **Michael Hahn**. Arabic relativization patterns: A unified HPSG analysis. In *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar*, 144–164. Stanford, CSLI Publications.

**2012.** **Michael Hahn** and Detmar Meurers. Evaluating the meaning of answers to reading comprehension questions: A semantics-based approach. In *Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA7)*, 326–336. Association for Computational Linguistics.

**2011.** **Michael Hahn** and Detmar Meurers. On deriving semantic representations from dependencies: A practical approach for evaluating meaning in learner corpora. In *Proceedings of the Int. Conference on Dependency Linguistics (Depling 2011)*, 94–103.

**2011.** **Michael Hahn**. Null conjuncts and bound pronouns in Arabic. In *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar*, 60–80. Stanford, CSLI Publications.

CONTRIBUTIONS  
TO EDITED VOL-  
UMES

**2014.** **Michael Hahn** and Detmar Meurers. On deriving semantic representations from dependencies: A practical approach for evaluating meaning in learner corpora. In Kim Gerdes, Eva Hajicová, and Leo Wanner, editors, *Dependency Theory*, Frontiers in AI and Applications Series, 94–103. IOS Press, 2014.

TEACHING AND  
MENTORING  
EXPERIENCE

**2020.** Mentor for undergraduate summer research, CSLI Summer Internship Program, CSLI, Stanford University.

**2019.** Teaching Assistant, *Natural Language Processing with Deep Learning (CS 224N)*, with Christopher D. Manning, Stanford University.

**2017–2018.** Linguistics Corpus Teaching Assistant, with Christopher D. Manning, Stanford University.

**2018.** Web-based Experimental Methods Workshop, Stanford University.

**2012.** Teaching Assistant, *Grammar Formalisms in Computational Linguistics*, with Detmar Meurers, University of Tübingen.

INVITED TALKS

**2021.** *Sensitivity as a Complexity Measure for Sequence Classification Tasks*, NLP Talk Series, Microsoft Research Lab India, November 2021.

**2021.** *Cognition Constrains Linguistic Diversity in Word Order*, Computational Psycholinguistics Lab, MIT, June 2021.

**2021.** *Memory Efficiency Predicts Ordering Universals in Language*, Colloquium, Department of Linguistics, University of Düsseldorf, Germany, April 2021.

**2021.** *Sensitivity as a Complexity Measure for Sequence Classification Tasks*, Singh Lab, Department of Computer Science, UC Irvine, March 2021.

**2021.** *Word Order as an Efficient Tradeoff of Memory and Surprisal*, Colloquium, Department of Linguistics, UT Austin, February 2021.

**2021.** *An Information-Theoretic Explanation of Adjective Ordering Preferences*, TEx-Mod2020: Theoretical and Experimental Approaches to Modification. Tübingen, Germany, January 2021.

**2020.** *A Neural Noisy-Channel Model of Structural Forgetting*, TedLab, MIT, November 2020.

**2020.** *Word Order Universals Optimize Communicative Efficiency*, Cognitive Lexicon Laboratory, University of Toronto, August 2020.

**2019.** *Crosslinguistic Word Orders Optimize Efficiency of Human Communication and Processing*, Harvard NLP, Harvard University, July 2019.

**2018.** *Explaining Syntactic Universals by Optimizing Grammars*. TedLab, MIT, November 2018.

CONTRIBUTED  
PRESENTATIONS

**2022.** *Modeling Fixations with Neural Attention* (work with Songpeng Yan and Frank Keller), Human Sentence Processing Conference, March 2022.

**2021.** *Sensitivity as a Complexity Measure for Sequence Classification Tasks* (work with Dan Jurafsky and Richard Futrell), EMNLP 2021: Conference on Empirical Methods in Natural Language Processing, November 2021.

**2020.** *Theoretical Limitations of Self-Attention in Neural Sequence Models*, 58th Annual Meeting of the Association for Computational Linguistics, July 2020.

**2020.** *Tabula nearly rasa: Probing the linguistic knowledge of character-level neural language models trained on unsegmented text* (work with Marco Baroni), 58th Annual Meeting of the Association for Computational Linguistics, July 2020.

**2020.** *Lexical Effects in Structural Forgetting: Evidence for Experience-Based Accounts and a Neural Network Model* (work with Richard Futrell and Edward Gibson), 33rd Annual CUNY Conference on Human Sentence Processing 2020, March 2020.

**2019.** *Character-based Surprisal as a Model of Human Reading in the Presence of Errors* (work with Frank Keller, Yonatan Bisk, Yonatan Belinkov), CogSci 2019, July 2019.

**2019.** *Crosslinguistic word orders enable an efficient tradeoff of memory and surprisal* (work with Judith Degen, Richard Futrell), 32nd Annual CUNY Conference on Human Sentence Processing, University of Colorado Boulder, March 2019.

**2019.** *Testing Functional Explanations of Word Order Universals* (work with Richard Futrell), 32nd Annual CUNY Conference on Human Sentence Processing, University of Colorado Boulder, USA, March 2019.

**2018.** *Testing Functional Explanations of Word Order Universals*. CAMP 2018, USC, Los Angeles, USA, November 2018.

**2018.** Poster: *An Information-Theoretic Explanation of Adjective Ordering Preferences*. (work with Judith Degen, Noah Goodman, Dan Jurafsky, and Richard Futrell), CogSci 2018, Madison, Wisconsin, USA, July 2018.

**2018.** *Wreath Products of Distributive Forest Algebras*. (work with Andreas Krebs and Howard Straubing), LICS 2018, Oxford, UK, July 2018.

**2018.** Poster: *Mutual Information Impacts Adjective Ordering Across Languages*. (work with Judith Degen, Dan Jurafsky, Noah Goodman, Richard Futrell), 31st Annual CUNY Conference on Human Sentence Processing, UC Davis, USA, March 2018.

**2018.** Poster: *Exploring Adjective Ordering Preferences via Artificial Language Learning*. (work with Judith Degen, Richard Futrell), 31st Annual CUNY Conference on Human Sentence Processing, UC Davis, USA, March 2018.

**2017.** *Exploring Adjective Ordering Preferences via Artificial Language Learning*. (work with Judith Degen, Richard Futrell), California Meeting on Psycholinguistics, 2017, UCLA, Los Angeles, USA, November 2017.

**2017.** *Modeling Task Effects in Reading with Neural Attention*. (work with Frank Keller), 30th Annual CUNY Conference on Human Sentence Processing, MIT, March 2017.

**2016.** *Modeling Human Reading with Neural Attention*. (work with Frank Keller), EMNLP 2016: Conference on Empirical Methods in Natural Language Processing, Austin, Texas, November 2016.

**2014.** *Predication and NP Structure in an Omnipredicative Language: The Case of Khoekhoe.* 21st International Conference on Head-Driven Phrase Structure Grammar, Buffalo, US, August 2014.

**2013.** *Word Order Variation in Khoekhoe.* 20th International Conference on Head-Driven Phrase Structure Grammar, Berlin, Germany, August 2013.

**2012.** *Arabic Relativization Patterns: A Unified HPSG Analysis.* 19th International Conference on Head-Driven Phrase Structure Grammar, Daejeon, South Korea, July 2012.

**2012.** (with Detmar Meurers) *Evaluating the Meaning of Answers to Reading Comprehension Questions: A Semantics-Based Approach.* 7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA7), Montreal, Canada, July 2012.

**2011.** *Null Conjuncts and Bound Pronouns in Arabic.* 18th International Conference on Head-Driven Phrase Structure Grammar, Seattle, USA, August 2011.

**2011.** (with Detmar Meurers) *On deriving semantic representations from dependencies.* International Conference on Dependency Linguistics, Barcelona, Spain, August 2011.

**2010.** *Agreement and Complex Predicates in Modern Standard Arabic.* Generative Grammatik des Nordens, Berlin, Germany, July 2010.

**2009.** *Nichtlokale Abhängigkeiten im Hocharabischen/Nonlocal Dependencies in Modern Standard Arabic.* Workshop on Grammar Theory and Grammar Implementation, Berlin, Germany, May 2009.

## SERVICE

### Reviewing

#### *Conferences*

ACL (2021, 2022), EACL (2021), EMNLP (2021), CogSci (2020, 2021, 2022), CoNLL (2020, 2021), MFCS (2019), STACS (2015).

#### *Journals*

Cognition (2020), Open Mind (2021), Neural Networks (2021), Glossa (2019, 2020, 2022), Neuropsychologia (2019), PeerJ (2019).

### Institutional Service

CSLI Summer Internship Admissions Committee, Stanford CSLI, 2020.

QP Fest Committee, Stanford Linguistics Department, 2019.

Social Committee, Stanford Linguistics Department, 2016–2017.

EMNLP Student Volunteer, 2016.