Ryan David Cotterell

CONTACT INFORMATION

Department of Computer Science
Johns Hopkins University
Hackerman 321
3400 North Charles Street
Baltimore, Maryland 21218, USA

mobile: (213) 905-2260
email: ryan.cotterell@jhu.edu
www: ryancotterell.github.io

EDUCATION

Johns Hopkins University — Spring 2018 (Expected)
Ph.D. in Computer Science
Advisors: Jason Eisner and David Yarowsky

Ludwig-Maximilians-Universität München — 2014-2016
Visiting Ph.D. Student
Advisor: Hinrich Schütze

Johns Hopkins University — Spring 2018
M.S.E. in Applied Mathematics and Statistics

Johns Hopkins University — Spring 2017
M.S.E. in Computer Science
Advisor: Chris Callison-Burch
GPA: 4.0

Johns Hopkins University — Spring 2013
B.A. in Cognitive Science
Minor: Linguistics
Advisor: Colin Wilson
GPA: 3.87 (*General Honors*)
Major GPA: 4.0 (*Departmental Honors*)

Faculty of Liberal Arts and Sciences of St. Petersburg State University — Fall 2009
Study Abroad, St. Petersburg, Russia

Friends School — Spring 2007
Baltimore, Maryland

EMPLOYMENT

Google Research, New York, NY — June-September 2017
Software Engineering Intern
Hosts: Brian Roark and Vlad Schogol

Human Language Technology Center of Excellence, Baltimore, MD — June-August 2012
Participant in the Summer Camp for Applied Language Exploration (SCALE)
Supervisor: Benjamin Van Durme

TEACHING

Johns Hopkins University
Role: Co-Instructor — Spring 2018
Course: Machine Learning: Linguistic and Sequence Modeling
Primary Instructor: Jason Eisner
I am co-developing a new course in advanced machine learning—mostly structured prediction—with natural language processing as the primary application. We will touch upon a wide variety of subjects: Bayesian non-parametrics, approximate inference and deep learning.

Johns Hopkins University
Role: Teaching Assistant — Fall 2016
Course: Machine Learning (600.475)
Professor: Mark Dredze
I held discussion sessions with students to prepare them for homework problem sets.

Johns Hopkins University
Role: Teaching Assistant — Spring 2014
Course: Automata and Computation Theory (600.271)
Professor: Stephen Checkoway
I managed three course assistants and held weekly office hours.

Johns Hopkins University
Role: Teaching Assistant — Fall 2013
Course: Natural Language Processing (600.465)
Professor: Jason Eisner
I led weekly discussion sections to cement concepts and improve problem solving skills. I supervised three course assistants in grading the assignments.

GRANTS

PURA (Provost Undergraduate Research Award)
Awarding body: Johns Hopkins University
Amount: $1,000
Awarded to investigate phonological opacity in Portuguese and Turkish in an optimality-theoretic framework.

PAPER AWARDS

Best Paper at ACL	2017
Outstanding Paper at EACL	2017
Runner-up for Best Short Paper at NAACL	2016
Honorable Mention for Best Short Paper at EMNLP	2015

OTHER AWARDS

Facebook Ph.D. Fellowship	2018
Fredrick Jelinek Fellowship	2017
National Defense Science and Engineering Fellowship (NDSEG)	2016-2018
DAAD Long-term Research Grant, Germany	2015-2016
Fulbright Research Grant, Germany	2014-2015
George M. L. Sommerman Engineering Graduate Teaching Assistant Award Finalist	2014
Computer Science Department Outstanding Teaching Assistant	2014
Cognitive Science Undergraduate Research Award	2013

PUBLICATIONS

Refereed Journal Papers

1. Ryan Cotterell, Christo Kirov, Mans Hulden, and Jason Eisner. 2018b. On the complexity and typology of inflectional morphological systems. *Transactions of the Association for Computational Linguistics (TACL)*, 6.
2. Ryan Cotterell and Hinrich Schütze. 2018. Joint semantic synthesis and morphological analysis of the derived word. *Transactions of the Association for Computational Linguistics (TACL)*, 6:33–48.
3. Ryan Cotterell, Nanyun Peng, and Jason Eisner. 2015b. Modeling word forms using latent underlying morphs and phonology. *Transactions of the Association for Computational Linguistics (TACL)*, 3:433–447.

Refereed Conference Papers

4. Ryan Cotterell, Christo Kirov, Mans Hulden, and Jason Eisner. 2018c. On the diachronic stability of irregularity in inflectional morphology. In *Proceedings of the 2018 Conference of the North*

American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), New Orleans, Louisiana. Association for Computational Linguistics.

5. Ryan Cotterell and Jason Eisner. 2018. A deep generative model of vowel formant typology. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, New Orleans, Louisiana. Association for Computational Linguistics.

6. Ryan Cotterell, Sebastian J. Mielke, Jason Eisner, and Brian Roark. 2018e. Are all languages equally hard to language-model? In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, New Orleans, Louisiana. Association for Computational Linguistics.

7. Ryan Cotterell, Christo Kirov, Sebastian J. Mielke, and Jason Eisner. 2018d. Unsupervised disambiguation of syncretism in inflected lexicons. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, New Orleans, Louisiana. Association for Computational Linguistics.

8. Christo Kirov, Ryan Cotterell, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, Patrick Xia, Manaal Faruqui, Arya McCarthy, Sandra Kübler, David Yarowsky, Jason Eisner, and Mans Hulden. 2018. Unimorph 2.0: Universal morphology. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*. European Language Resources Association (ELRA).

9. Ryan Cotterell, Christo Kirov, Jason Eisner, and Mans Hulden. 2018a. The Pareto complexity of inflectional systems. In *Proceedings of the 1st Annual Meeting of the Society for Computation in Linguistics (SCiL)*, Salt Lake City.

10. Ryan Cotterell and Kevin Duh. 2017. Low-resource named entity recognition with cross-lingual, character-level neural conditional random fields. In *Proceedings of 8th International Joint Conference on Natural Language Processing (IJCNLP)*, Taipei, Taiwan. Asian Federation of Natural Language Processing. 6 pages.

11. Ryan Cotterell and Georg Heigold. 2017. Cross-lingual, character-level neural morphological tagging. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 759–770, Copenhagen, Denmark. Association for Computational Linguistics.

12. Ryan Cotterell, Katerina Vylomova, Huda Khayrallah, Christo Kirov, and David Yarowsky. 2017d. Paradigm completion for derivational morphology. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 725–731, Copenhagen, Denmark. Association for Computational Linguistics.

13. Francis Ferraro, Adam Poliak, Ryan Cotterell, and Benjamin Van Durme. 2017. Frame-based continuous lexical semantics through exponential family tensor factorization and semantic proto-roles. In *Proceedings of the Sixth Joint Conference on Lexical and Computational Semantics (*SEM)*, Vancouver, Canada. Association for Computational Linguistics.

14. Ryan Cotterell and Jason Eisner. 2017. Probabilistic typology: Deep generative models of vowel inventories. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1182–1192, Vancouver, Canada. Association for Computational Linguistics. Best Paper Award.

15. Katharina Kann, Ryan Cotterell, and Hinrich Schütze. 2017b. One-shot neural cross-lingual transfer for paradigm completion. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1182–1192, Vancouver, Canada. Association for Computational Linguistics.

16. Ryan Cotterell, Adam Poliak, Benjamin Van Durme, and Jason Eisner. 2017b. Explaining and generalizing skip-gram through exponential family principal component analysis. In

Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), pages 175–181, Valencia, Spain. Association for Computational Linguistics.

17. Ryan Cotterell, John Sylak-Glassman, and Christo Kirov. 2017c. Neural graphical models over strings for principal parts morphological paradigm completion. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 759–765, Valencia, Spain. Association for Computational Linguistics. Outstanding Paper Award.

18. Arun Kumar, Ryan Cotterell, Lluís Padró, and Antoni Oliver. 2017. Morphological analysis of the Dravidian language family. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 217–222, Valencia, Spain. Association for Computational Linguistics.

19. Christo Kirov, John Sylak-Glassman, Rebecca Knowles, Ryan Cotterell, and Matt Post. 2017. A rich morphological tagger for English: Exploring the cross-linguistic tradeoff between morphology and syntax. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 112–117, Valencia, Spain. Association for Computational Linguistics.

20. Ekaterina Vylomova, Ryan Cotterell, Timothy Baldwin, and Trevor Cohn. 2017. Context-aware prediction of derivational word-forms. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 118–124, Valencia, Spain. Association for Computational Linguistics.

21. Katharina Kann, Ryan Cotterell, and Hinrich Schütze. 2017a. Neural multi-source morphological reinflection. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 514–524, Valencia, Spain. Association for Computational Linguistics.

22. Ryan Cotterell, Arun Kumar, and Hinrich Schütze. 2016b. Morphological segmentation inside-out. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2325–2330, Austin, Texas. Association for Computational Linguistics.

23. Katharina Kann, Ryan Cotterell, and Hinrich Schütze. 2016. Neural morphological analysis: Encoding-decoding canonical segments. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 961–967, Austin, Texas. Association for Computational Linguistics.

24. Tim Vieira*, Ryan Cotterell*, and Jason Eisner. 2016. Speed-accuracy tradeoffs in tagging with variable-order CRFs and structured sparsity. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1973–1978, Austin, Texas. Association for Computational Linguistics.

25. Ryan Cotterell, Hinrich Schütze, and Jason Eisner. 2016c. Morphological smoothing and extrapolation of word embeddings. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1651–1660, Berlin, Germany. Association for Computational Linguistics.

26. Ryan Cotterell, Tim Vieira, and Hinrich Schütze. 2016d. A joint model of orthography and morphological segmentation. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 664–669, San Diego, California. Association for Computational Linguistics. Runner-up for Best Paper.

27. Pushpendre Rastogi, Ryan Cotterell, and Jason Eisner. 2016. Weighting finite-state transductions with neural context. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 623–633, San Diego, California. Association for Computational Linguistics.

28. John Sylak-Glassman and Ryan Cotterell. 2016. Contrastive morphological typology and logical hierarchies. In Jessica Kantarovich, Tran Truong, and Orest Xherija, editors, *Proceedings of the 52nd Meeting of the Chicago Linguistic Society (CLS52)*. Chicago Linguistic Society. 15 pages.

29. Nanyun Peng, Ryan Cotterell, and Jason Eisner. 2015. Dual decomposition inference for graphical models over strings. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 917–927, Lisbon, Portugal. Association for Computational Linguistics.

30. Thomas Müller, Ryan Cotterell, Alexander Fraser, and Hinrich Schütze. 2015. Joint lemmatization and morphological tagging with LEMMING. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2268–2274, Lisbon, Portugal. Association for Computational Linguistics. Honorable Mention for Best Paper.

31. Ryan Cotterell, Thomas Müller, Alexander Fraser, and Hinrich Schütze. 2015a. Labeled morphological segmentation with semi-Markov models. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning (CoNLL)*, pages 164–174, Beijing, China. Association for Computational Linguistics.

32. Ryan Cotterell and Jason Eisner. 2015. Penalized expectation propagation for graphical models over strings. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 932–942, Denver, Colorado. Association for Computational Linguistics.

33. Ryan Cotterell and Hinrich Schütze. 2015. Morphological word embeddings. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 1287–1292, Denver, Colorado. Association for Computational Linguistics.

34. Ryan Cotterell, Nanyun Peng, and Jason Eisner. 2014a. Stochastic contextual edit distance and probabilistic FSTs. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 625–630, Baltimore, Maryland. Association for Computational Linguistics.

35. Ryan Cotterell and Chris Callison-Burch. 2014. A multi-dialect, multi-genre corpus of informal written Arabic. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*, Reykjavik, Iceland. European Language Resources Association (ELRA). 4 pages.

Refereed Workshop Papers

36. Gaurav Kumar, Yuan Cao, Ryan Cotterell, Chris Callison-Burch, Daniel Povey, and Sanjeev Khudanpur. 2014. Translations of the CALLHOME Egyptian Arabic corpus for conversational speech translation. In *Proceedings of the International Workshop on Spoken Language Translation (IWSLT)*, Lake Tahoe, USA. Association for Computational Linguistics.

37. Ryan Cotterell, Adithya Renduchintala, Naomi Saphra, and Chris Callison-Burch. 2014b. An Algerian Arabic-French code-switched corpus. In *Workshop on Free/Open-Source Arabic Corpora and Corpora Processing Tools (OSACT)*, Reykjavik, Iceland. European Language Resources Association. 5 pages.

Invited Publications

38. Ryan Cotterell, Christo Kirov, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, Patrick Xia, Manaal Faruqui, Sandra Kübler, David Yarowsky, Jason Eisner, and Mans Hulden. 2017a. The CoNLL-SIGMORPHON 2017 shared task: Universal morphological reinflection in 52 languages. In *Proceedings of the CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection*, Vancouver, Canada. Association for Computational Linguistics.

39. Ryan Cotterell, Christo Kirov, John Sylak-Glassman, David Yarowsky, Jason Eisner, and Mans Hulden. 2016a. The SIGMORPHON 2016 shared task—morphological reinflection. In *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 10–22, Berlin, Germany. Association for Computational Linguistics.

Unrefereed Publications

40. Chandler May, Ryan Cotterell, and Benjamin Van Durme. 2016. Analysis of morphology in topic modeling. *CoRR*, abs/1608.03995

Edited Volumes

41. Mans Hulden, Ryan Cotterell, Jason Eisner, Manaal Faruqui, Christo Kirov, Sandra Kübler, John Sylak-Glassman, Ekaterina Vylomova, Géraldine Walther, Patrick Xia, and David Yarowsky, editors. 2017. *Proceedings of the CoNLL SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection*. Association for Computational Linguistics, Vancouver. 125 pages.

Technical Reports

42. David Etter, Francis Ferraro, Ryan Cotterell, Olivia Buzek, and Benjamin Van Durme. 2013. Nerit: Named entity recognition for informal text. Technical Report 11, Human Language Technology Center of Excellence, Johns Hopkins University.

Invited Talks

1. Musings on Linguistic Complexity and Typology — February, 2018
   Location: University of Copenhagen
   Host: Isabelle Augenstein and Anders Søgaard
2. Probabilistic Typology: Deep Generative Models of Vowel Inventories — January, 2018
   Location: Universität Hamburg
   Host: Chris Biemann and Alexander Panchenko
3. Probabilistic Typology: Deep Generative Models of Vowel Inventories — November, 2017
   Location: Massachusetts Institute of Technology
   Host: Yevgeni Berzak and Richard Futrell
4. Probabilistic Typology: Deep Generative Models of Vowel Inventories — October, 2017
   Location: The Data Science Workshop on Computational Social Science (Yale University)
   Host: Dragomir Radev
5. Neural Weighted Finite-State Machines — September, 2017
   Location: First Workshop on Subword and Character Level Models in NLP (EMNLP 2017)
   Tutorial Talk
6. Probabilistic Typology: Deep Generative Models of Vowel Inventories — August, 2017
   Location: New York University
   Host: Kyunghyun Cho
7. Neural Graphical Models over Strings — May, 2017
   Location: Universität Heidelberg
   Host: Stefan Riezler
8. Neural String-Valued Graphical Models — January, 2017
   Location: Schloss Dagstuhl
   From Characters to Understanding Natural Language (Dagstuhl Seminar 17042)
9. Graphical Models over Strings — October, 2016
   Location: University of Alberta
   Host: Greg Kondrak

10. Graphical Models over Strings — September, 2016
    Location: Johns Hopkins University
    CLSP Seminar
11. Modeling Word Forms Using Latent Underlying Morphs and Phonology — July, 2016
    Location: Universität Tübingen
    Host: Gerhard Jäger
12. Modeling Word Forms Using Latent Underlying Morphs and Phonology — February, 2016
    Location: Xerox Research Centre Europe
    Host: Xavier Carreras
13. Modeling Word Forms Using Latent Underlying Morphs and Phonology — September, 2015
    Location: Priberam Labs
    Host: André Martins
14. A Probabilistic Approach to Synchronic Phonology — November, 2014
    Institut für Phonetik und Sprachverarbeitung, LMU München
    Host: Jonathan Harrington

Shared Task Organizer

1. CoNLL-SIGMORPHON 2018 Shared Task: Cross-lingual Morphological Reinflection.
2. CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection.
3. SIGMORPHON 2016 Shared Task: Morphological Reinflection.

Service

Journal Reviewer: TACL (2017), *Computational Linguistics* (2017, 2015), *Computer Speech and Language* (2017), *Natural Language Engineering* (2017)
Conference Reviewer: ACL (2018, 2017, 2016), EMNLP (2017, 2016), NAACL (2018, 2016), EACL (2017), COLING (2018, 2016), ICLR (2018), LREC (2018), AAAI (2016 secondary)
Workshop Reviewer: Workshop on Analyzing and Interpreting Neural Networks for NLP (EMNLP 2018), Student Research Workshop (ACL 2018), Relevance of Linguistic Structure in Neural NLP (ACL 2018), Workshop on New Forms of Generalization in Deep Learning and Natural Language Processing (NAACL 2018), Ethics in NLP (NAACL 2018, EACL 2017), Workshop on Deep Structured Prediction (ICML 2017), Subword and character level models in NLP (NAACL 2018, EMNLP 2017), SIGMORPHON (EMNLP 2018, ACL 2016), Multilingual and Cross-lingual Methods in NLP (NAACL 2016)
Poster Session Chair: EMNLP (2017)
Other: SIGMORPHON Officer At-Large, CLSP Happy Hour (and Karaoke) Coordinator

References

Jason Eisner (`jason@cs.jhu.edu`), Johns Hopkins University
David Yarowsky (`yarowsky@jhu.edu`), Johns Hopkins University
Colin Wilson (`wilson@cogsci.jhu.edu`), Johns Hopkins University
Hinrich Schütze (`inquiries@cis.lmu.org`), Ludwig-Maximilians-Universität München

Skills

Programming Languages: Python, Cython, Java, Perl, Ocaml, Lisp, C, C++, R, Scala, LaTeX

Deep Learning Frameworks: PyTorch, Theano, TensorFlow

Natural Languages: English, German, Spanish, Russian, Portuguese

Graduate Coursework in Computer Science: Natural Language Processing, Speech Processing, Graphical Models, Artificial Intelligence, Programming Language Theory, Software Engineering, Representation Learning, Big Data, Causality (audit)

Graduate Coursework in Mathematics and Statistics: Nonlinear Optimization I, Stochastic Optimization, Convex Optimization, Neural Networks, Real Analysis I, Real Analysis II (Measure Theory), Matrix Analysis, Bayesian Statistics

Graduate Coursework in Linguistics: Syntax I, Semantics I, Event Semantics (audit), Phonology I, Phonology II, Morpho-Phonology, Psycholinguistics