Ryan 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 2019 (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 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

EXPERIENCE

Google Research, New York, NY — June-September 2017
Software Engineering Intern
Host: Keith Hall

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

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

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

Teaching Assistant
Johns Hopkins University — 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.

Awards

Runner-up for Best Paper at NAACL	2016
Runner-up for Best Paper at EMNLP	2015
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 and Hinrich Schütze. 2017. Joint semantic synthesis and morphological analysis of the derived word. *Transactions of the Association for Computational Linguistics (TACL)* 5. https://arxiv.org/abs/1701.00946.

2. Ryan Cotterell, Nanyun Peng, and Jason Eisner. 2015. Modeling word forms using latent underlying morphs and phonology. *Transactions of the Association for Computational Linguistics (TACL)* 3:433–447. https://tacl2013.cs.columbia.edu/ojs/index.php/tacl/article/view/480.

Refereed Conference Papers

3. Ryan Cotterell, Adam Poliak, Benjamin Van Durme, and Jason Eisner. 2017. 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)*. Association for Computational Linguistics, Valencia, Spain. https://ryancotterell.github.io/papers/cotterell+alb.eacl17.pdf

4. Ryan Cotterell, John Sylak-Glassman, and Christo Kirov. 2017. 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)*. Association for Computational Linguistics, Valencia, Spain. https://ryancotterell.github.io/papers/cotterell+ala.eacl17.pdf

5. Arun Kumar, Ryan Cotterell, Luís 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)*. Association for Computational Linguistics, Valencia, Spain. https://ryancotterell.github.io/papers/etter+al.tr13.pdf

6. Christo Kirov, John Sylak-Glassman, Rebecca Knowles 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)*. Association for Computational Linguistics, Valencia, Spain. https://ryancotterell.github.io/papers/kirov+al.eacl17.pdf

7. 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). Association for Computational Linguistics, Valencia, Spain. https://ryancotterell.github.io/papers/vylomova+al.eacl17.pdf

8. Katharina Kann, Ryan Cotterell, and Hinrich Schütze. 2017. Neural multi-source morphological reinflection. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. Association for Computational Linguistics, Valencia, Spain. https://ryancotterell.github.io/papers/kann+al.eacl17.pdf.

9. Ryan Cotterell, Arun Kumar, and Hinrich Schütze. 2016. Morphological segmentation inside-out. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Austin, Texas, pages 2325–2330. https://aclweb.org/anthology/D16-1256.

10. 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)*. Association for Computational Linguistics, Austin, Texas, pages 961–967. https://aclweb.org/anthology/D16-1097.

11. 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)*. Association for Computational Linguistics, Austin, Texas, pages 1973–1978. https://aclweb.org/anthology/D16-1206.

12. Ryan Cotterell, Hinrich Schütze, and Jason Eisner. 2016. Morphological smoothing and extrapolation of word embeddings. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*. Association for Computational Linguistics, Berlin, Germany, pages 1651–1660. http://www.aclweb.org/anthology/P16-1156.

13. Ryan Cotterell, Tim Vieira, and Hinrich Schütze. 2016. 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)*. Association for Computational Linguistics, San Diego, California, pages 664–669. Runner-up for Best Paper. http://www.aclweb.org/anthology/N16-1080.

14. 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)*. Association for Computational Linguistics, San Diego, California, pages 623–633. http://www.aclweb.org/anthology/N16-1076.

15. John Sylak-Glassman and Ryan Cotterell. 2015. 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. https://ryancotterell.github.io/papers/sylak-glassman+cotterell.cls16.pdf.

16. 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)*. Association for Computational Linguistics, Lisbon, Portugal, pages 917–927. http://aclweb.org/anthology/D15-1108.

17. 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)*. Association for Computational Linguistics, Lisbon, Portugal, pages 2268–2274. Runner-up for Best Paper. http://aclweb.org/anthology/D15-1272.

18. Ryan Cotterell, Thomas Müller, Alexander Fraser, and Hinrich Schütze. 2015. Labeled morphological segmentation with semi-markov models. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning (CoNLL)*. Association for Computational Linguistics, Beijing, China, pages 164–174. http://www.aclweb.org/anthology/K15-1017.

19. 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)*. Association for Computational Linguistics, Denver, Colorado, pages 932–942. http://www.aclweb.org/anthology/N15-1094.

20. 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)*. Association for Computational Linguistics, Denver, Colorado, pages 1287–1292. http://www.aclweb.org/anthology/N15-1140.

21. Ryan Cotterell, Nanyun Peng, and Jason Eisner. 2014. Stochastic contextual edit distance and probabilistic FSTs. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL)*. Association for Computational Linguistics, Baltimore, Maryland, pages 625–630. http://www.aclweb.org/anthology/P14-2102.

22. Ryan Cotterell and Chris Callison-Burch. 2014. A multi-dialect, multi-genre corpus of informal written arabic. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*. European Language Resources Association (ELRA), Reykjavik, Iceland. https://ryancotterell.github.io/papers/cotterell+callison-burch.lrec14.pdf.

Refereed Workshop Papers

23. Ryan Cotterell, Christo Kirov, John Sylak-Glassman, David Yarowsky, Jason Eisner, and Mans Hulden. 2016. The SIGMORPHON 2016 shared taskmorphological reinflection. In *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*. Association for Computational Linguistics, Berlin, Germany, pages 10–22. http://anthology.aclweb.org/W16-2002.

24. 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)*. Association for Computational Linguistics, Lake Tahoe, USA. https://ryancotterell.github.io/papers/kumar+al.iwslt14.pdf.

25. Ryan Cotterell, Adithya Renduchintala, Naomi Saphra, and Chris Callison-Burch. 2014. An Algerian Arabic-French code-switched corpus. In *Workshop on Free/Open-Source Arabic Corpora and Corpora Processing Tools (OSACT)*. European Language Resources Association, Reykjavik, Iceland. https://ryancotterell.github.io/papers/cotterell+al.osact14.pdf.

Unrefereed Publications

26. Chandler May, Ryan Cotterell, and Benjamin Van Durme. 2016. Analysis of morphology in topic modeling. *CoRR* abs/1608.03995. http://arxiv.org/abs/1608.03995

Technical Reports

27. 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. https://ryancotterell.github.io/papers/etter+al.tr13.pdf.

Invited Talks

1. Graphical Models over Strings — February, 2017
   Location: Universität Heidelberg
   Host: Stefan Riezler
2. Neural String-Valued Graphical Models — January, 2017
   Location: Schloss Dagstuhl
   From Characters to Understanding Natural Language (Dagstuhl Seminar 17042)
3. Graphical Models over Strings — October, 2016
   Location: University of Alberta
   Host: Greg Kondrak
4. Graphical Models over Strings — September, 2016
   Location: Johns Hopkins University
   CLSP Seminar
5. Modeling Word Forms Using Latent Underlying Morphs and Phonology — July, 2016
   Location: Universität Tübingen
   Host: Gerhard Jäger
6. Modeling Word Forms Using Latent Underlying Morphs and Phonology — December, 2015
   Location: Xerox Research Centre Europe
   Host: Xavier Carreras
7. Modeling Word Forms Using Latent Underlying Morphs and Phonology — September, 2015
   Location: Priberam Labs
   Host: André Martins
8. A Probabilistic Approach to Synchronic Phonology — November, 2014.
   Institut für Phonetik und Sprachverarbeitung, LMU München
   Host: Jonathan Harrington

Shared Task Organizer

1. CoNLL-SIGMORPHON-2017 Shared Task: Universal Morphological Reinflection
2. SIGMORPHON 2016 Shared Task: Morphological Reinflection

Service

Journal Reviewer: *Computational Linguistics* (2015)
Conference Reviewer: ACL (2016, 2017), EMNLP (2016), NAACL (2016), EACL (2017), COLING (2016), AAAI (2016 secondary)
Workshop Reviewer: Subword and character level models in NLP (2017), Ethics in NLP (2017), SIGMORPHON (2016), Multilingual and Cross-lingual Methods in NLP (2016)
Other: SIGMORPHON Officer At-Large, CLSP Happy Hour 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

Skills

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

Languages: English, German, Spanish, Russian, Portuguese

Graduate Coursework: Natural Language Processing, Speech Processing, Machine Learning, Artificial Intelligence, Programming Language Theory, Nonlinear Optimization, Stochastic Optimization, Neural Networks, Real Analysis, Software Engineering, Representation Learning, Big Data, Bayesian Statistics