

## Ryan Cotterell

---

|                        |  |  |
|------------------------|--|--|
| CONTACT<br>INFORMATION | Department of Computer Science<br>Johns Hopkins University<br>Hackerman 321<br>3400 North Charles Street<br>Baltimore, Maryland 21218, USA   | <i>mobile:</i> (213) 905-2260<br><i>email:</i> ryan.cotterell@jhu.edu<br><i>www:</i> ryancotterell.github.io |
| EDUCATION              | <b>Johns Hopkins University</b><br>Ph.D. in Computer Science<br>Advisors: Jason Eisner and David Yarowsky  | Spring 2019 (Expected)   |
|                        | <b>Ludwig-Maximilians-Universität München</b><br>Visiting Ph.D. Student<br>Advisor: Hinrich Schütze  | 2014-2016  |
|                        | <b>Johns Hopkins University</b><br>M.S.E. in Computer Science<br>Advisor: Chris Callison-Burch<br>GPA: 4.0   | Spring 2017  |
|                        | <b>Johns Hopkins University</b><br>B.A. in Cognitive Science<br>Minor: Linguistics<br>Advisor: Colin Wilson<br>GPA: 3.87 ( <i>General Honors</i> )<br>Major GPA: 4.0 ( <i>Departmental Honors</i> )                  | Spring 2013  |
|                        | <b>Faculty of Liberal Arts and Sciences of St. Petersburg State University</b><br>Study Abroad, St. Petersburg, Russia   | Fall 2009  |
| EMPLOYMENT             | <b>Google Research</b> , New York, NY<br>Software Engineering Intern<br>Hosts: Brian Roark and Vlad Schogol  | June-September 2017  |
|                        | <b>Human Language Technology Center of Excellence</b> , Baltimore, MD<br>Participant in the Summer Camp for Applied Language Exploration (SCALE)<br>Supervisor: Benjamin Van Durme                                   | June-August 2012   |
| TEACHING               | <b>Teaching Assistant</b><br>Johns Hopkins University<br>Course: Machine Learning (600.475)<br>Professor: Mark Dredze<br><i>I held discussion sessions with students to prepare them for homework problem sets.</i>  | Fall 2016  |
|                        | <b>Teaching Assistant</b><br>Johns Hopkins University<br>Course: Automata and Computation Theory (600.271)<br>Professor: Stephen Checkoway<br><i>I managed three course assistants and held weekly office hours.</i> | Spring 2014  |
|                        | <b>Teaching Assistant</b><br>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

|  |           |
|--|-----------|
| Outstanding 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      |
| 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 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)*. Association for Computational Linguistics, Copenhagen, Denmark.
4. Ryan Cotterell and Georg Heigold. 2017. Paradigm completion for derivational morphology. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Copenhagen, Denmark.
5. Francis Ferraro, Adam Poliak, Ryan Cotterell, and Benjamin Van Durme. 2017. Frame-based continuous lexical semantics through exponential family tensor factorization and semantic protocols. In *Proceedings of the Sixth Joint Conference on Lexical and Computational Semantics (\*SEM)*. Association for Computational Linguistics, Vancouver, Canada.
6. 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)*. Association for Computational Linguistics, Vancouver, Canada. **Outstanding Paper Award**. <https://arxiv.org/abs/1705.01684>.
7. Katharina Kann, Ryan Cotterell, and Hinrich Schütze. 2017. [One-shot neural cross-lingual transfer for paradigm completion](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*. Association for Computational Linguistics, Vancouver, Canada. <https://arxiv.org/abs/1704.00052>.
8. 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, pages 175–181. <http://www.aclweb.org/anthology/E17-2028>.

9. 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, pages 759–765. **Outstanding Paper Award**. <http://www.aclweb.org/anthology/E17-2120>.
10. 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)*. Association for Computational Linguistics, Valencia, Spain, pages 217–222. <http://www.aclweb.org/anthology/E17-2035>.
11. 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)*. Association for Computational Linguistics, Valencia, Spain, pages 112–117. <http://www.aclweb.org/anthology/E17-2018>.
12. 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, pages 118–124. <http://www.aclweb.org/anthology/E17-2019>.
13. 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, pages 514–524. <http://aclweb.org/anthology/E/E17/E17-1049.pdf>.
14. 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>.
15. 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>.
16. 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>.
17. 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>.
18. 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>.
19. 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>.
20. 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>.
  21. 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>.
  22. 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. **Honorable Mention for Best Paper**. <http://aclweb.org/anthology/D15-1272>.
  23. 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>.
  24. 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>.
  25. 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>.
  26. 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>.
  27. 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

28. 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>.
29. 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

30. Ryan Cotterell, Christo Kirov, John Sylak-Glassman, Géraldine Walther, Ekaterina Vylomova, Patrick Xia, Manaal Faruqui, Sandra Kübler, David Yarowsky, Jason Eisner, and Mans Hulden. 2017. The CoNLL-SIGMORPHON 2017 shared task: Universal morphological reinflection in 52 languages. In *Proceedings of the CoNLL-SIGMORPHON 2017 Shared Task: Universal Morphological Reinflection*. Association for Computational Linguistics, Vancouver, Canada.
31. Ryan Cotterell, Christo Kirov, John Sylak-Glassman, David Yarowsky, Jason Eisner, and Mans Hulden. 2016. [The SIGMORPHON 2016 shared task—morphological 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>.
32. 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

33. 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. Neural Weighted Finite-State Machines Location: First Workshop on Subword and Character Level Models in NLP, held at EMNLP 2017  
Tutorial Talk
2. Neural Graphical Models over Strings May, 2017  
Location: Universität Heidelberg  
Host: Stefan Riezler
3. Neural String-Valued Graphical Models January, 2017  
Location: Schloss Dagstuhl  
From Characters to Understanding Natural Language (Dagstuhl Seminar 17042)
4. Graphical Models over Strings October, 2016  
Location: University of Alberta  
Host: Greg Kondrak
5. Graphical Models over Strings September, 2016  
Location: Johns Hopkins University  
CLSP Seminar
6. Modeling Word Forms Using Latent Underlying Morphs and Phonology July, 2016  
Location: Universität Tübingen  
Host: Gerhard Jäger
7. Modeling Word Forms Using Latent Underlying Morphs and Phonology February, 2016  
Location: Xerox Research Centre Europe  
Host: Xavier Carreras
8. Modeling Word Forms Using Latent Underlying Morphs and Phonology September, 2015  
Location: Priberam Labs  
Host: André Martins

9. A Probabilistic Approach to Synchronic Phonology  
Institut für Phonetik und Sprachverarbeitung, LMU München  
Host: Jonathan Harrington

November, 2014

SHARED TASK  
ORGANIZER

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

SERVICE

**Journal Reviewer:** *Computational Linguistics* (2017, 2015), *Computer Speech and Language* (2017)  
**Conference Reviewer:** ACL (2017, 2016), EMNLP (2017, 2016), NAACL (2016), EACL (2017), COLING (2016), AAAI (2016 secondary)  
**Workshop Reviewer:** ICML Workshop on Deep Structured Prediction (2017), 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](mailto:jason@cs.jhu.edu)), Johns Hopkins University  
David Yarowsky ([yarowsky@jhu.edu](mailto:yarowsky@jhu.edu)), Johns Hopkins University  
Colin Wilson ([wilson@cogsci.jhu.edu](mailto:wilson@cogsci.jhu.edu)), Johns Hopkins University

SKILLS

**Programming Languages:** Python, Cython, Java, Perl, Ocaml, Lisp, C, C++, R, Scala, L<sup>A</sup>T<sub>E</sub>X

**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), Bayesian Statistics

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