

Department of Computer Science  
University of Rochester  
PO Box 270226  
Rochester NY 14627

gildea@cs.rochester.edu  
<http://www.cs.rochester.edu/~gildea/>  
585 275-7230

## EDUCATION

- PhD 2001 Computer Science, University of California, Berkeley.  
Directed by Nelson Morgan, Daniel Jurafsky, Charles Fillmore, Jerome Feldman.  
Dissertation: *Statistical Language Understanding Using Frame Semantics*
- MS 1999 Computer Science, University of California, Berkeley.  
Thesis: *Topic-Based Language Models Using EM*
- BA 1995 Double major in Linguistics and Computer Science,  
University of California, Berkeley.

## RESEARCH INTERESTS

- statistical machine translation
- natural language understanding, including syntactic parsing and semantic interpretation
- machine learning algorithms

## PROFESSIONAL EXPERIENCE

- 2016 – Professor, Department of Computer Science, University of Rochester.
- 2009 – 2016 Associate professor, Department of Computer Science, University of Rochester.
- 2003 – 2009 Assistant professor, Department of Computer Science, University of Rochester.
- 2001 – 2003 Postdoctoral fellow, Institute for Research in Cognitive Science and Department of Computer and Information Science, University of Pennsylvania.
- 1996 – 2001 Research assistant, International Computer Science Institute, Berkeley, California.
- 1995 – 1996 Software engineer, PA-RISC compiler group, Hewlett-Packard Corporation.

## PUBLICATIONS

### Books

*Semantic Role Labeling*. Martha Palmer, Daniel Gildea, and Nianwen Xue. 2010. Synthesis Lectures on Human Language Technology Series. Morgan and Claypool.

### Journal Articles

Ordered tree decomposition for HRG rule extraction. Daniel Gildea, Giorgio Satta, and Xiaochang Peng. 2019. *Computational Linguistics*, 45(2):339–337.

Semantic neural machine translation using AMR. Linfeng Song, Daniel Gildea, Yue Zhang, Zhiguo Wang, and Jinsong Su. 2019. *Transactions of the Association for Computational Linguistics (TACL)*, 7:19–31.

Feature-based decipherment for machine translation. Iftekhar Naim, Parker Riley, and Daniel Gildea. 2018. *Computational Linguistics*, 44(3):525–546.

Cache transition systems for graph parsing. Daniel Gildea, Giorgio Satta, and Xiaochang Peng. 2018. *Computational Linguistics*, 44(1):85–118.

Weighted DAG automata for semantic graphs. David Chiang, Frank Drewes, Daniel Gildea, Adam Lopez, and Giorgio Satta. 2018. *Computational Linguistics*, 44(1):119–186.

A notion of semantic coherence for underspecified semantic representation. Mehdi Manshadi, Daniel Gildea, and James F. Allen. 2018. *Computational Linguistics*, 44(1):39–83.

Minimizing syntactic dependency lengths: Typological/cognitive universal? David Temperley and Daniel Gildea. 2018. *Annual Reviews of Linguistics*.

Automated analysis and prediction of job interview performance. Iftekhar Naim, Mohammad Iftekhar Tanveer, Daniel Gildea, and Mohammed (Ehsan) Hoque. 2018. *IEEE Transactions on Affective Computing*, 9(2):191–204.

Scribe: Deep integration of human and machine intelligence to caption speech in real time. Walter S. Lasecki, Christopher D. Miller, Iftekhar Naim, Raja Kushalnagar, Adam Sadilek, Daniel Gildea, and Jeffrey P. Bigham. 2017. *Commun. ACM*, 60(9):93–100

Parsing linear-context free rewriting systems with fast matrix multiplication. Shay B. Cohen and Daniel Gildea. 2016. *Computational Linguistics*, 42(3):421–455.

Synchronous context-free grammars and optimal parsing strategies. Daniel Gildea and Giorgio Satta. 2016. *Computational Linguistics*, 42(2):207–243.

Synchronous context-free grammars and optimal linear parsing strategies. Pierluigi Crescenzi, Daniel Gildea, Andrea Marino, Gianluca Rossi, and Giorgio Satta. 2015. *Journal of Computer and System Sciences*, 81(7):1333–1356.

Information density and syntactic repetition. David Temperley and Daniel Gildea. 2015. *Cognitive Science*, 39(8):1802–1823.

Sampling tree fragments from forests. Tagyoung Chung, Licheng Fang, Daniel Gildea, and Daniel Štefankovič. 2014. *Computational Linguistics*, 40(1):203–229.

Bayesian tree substitution grammars as a usage-based approach. Matt Post and Daniel Gildea. 2013. *Language and Speech*, 56(3):291–308.

On the string translations produced by multi bottom-up tree transducers. Daniel Gildea. 2012. *Computational Linguistics*, 38(3):673–693.

- Grammar factorization by tree decomposition. Daniel Gildea. 2011. *Computational Linguistics*, 37(1):231–248.
- Do grammars minimize dependency length? Daniel Gildea and David Temperley. 2010. *Cognitive Science*, 34(2):286–310.
- Binarization of synchronous context-free grammars. Liang Huang, Hao Zhang, Daniel Gildea, and Kevin Knight. 2009. *Computational Linguistics*, 35(4):559–595.
- Enumeration of factorizable multi-dimensional permutations. Hao Zhang and Daniel Gildea. 2007. *Journal of Integer Sequences*, 07(5.8).
- The Proposition Bank: An annotated corpus of semantic roles. Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. *Computational Linguistics*, 31(1):71–106.
- Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. Alan Bell, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory, and Daniel Gildea. 2003. *Journal of the Acoustical Society of America*, 113(2):1001–1024.
- Automatic labeling of semantic roles. Daniel Gildea and Daniel Jurafsky. 2002. *Computational Linguistics*, 28(3):245–288.
- Learning bias and phonological rule induction. Daniel Gildea and Daniel Jurafsky. 1996. *Computational Linguistics*, 22(4):497–530.

### Refereed Conference Papers

- SemBleu: A robust metric for AMR parsing evaluation. Linfeng Song and Daniel Gildea. 2019. In *Proceedings of the 57th Conference of the Association for Computational Linguistics*, pages 4547–4552. Florence, Italy.
- N-ary relation extraction using graph-state LSTM. Linfeng Song, Yue Zhang, Zhiguo Wang, and Daniel Gildea. 2018. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2226–2235. Association for Computational Linguistics.
- Orthographic features for bilingual lexicon induction. Parker Riley and Daniel Gildea. 2018. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 390–394. Melbourne, Australia.
- A graph-to-sequence model for AMR-to-text generation. Linfeng Song, Yue Zhang, Zhiguo Wang, and Daniel Gildea. 2018. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL-18)*, pages 1616–1626.
- Sequence-to-sequence models for cache transition systems. Xiaochang Peng, Linfeng Song, Daniel Gildea, and Giorgio Satta. 2018. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL-18)*, pages 1842–1852.
- Leveraging context information for natural question generation. Linfeng Song, Zhiguo Wang, Wael Hamza, Yue Zhang, and Daniel Gildea. 2018. In *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL-18)*, pages 569–574.
- AMR parsing with cache transition systems. Xiaochang Peng, Daniel Gildea, and Giorgio Satta. 2018. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-18)*.
- AMR-to-text generation with synchronous node replacement grammar. Linfeng Song, Xiaochang Peng, Yue Zhang, Zhiguo Wang, and Daniel Gildea. 2017. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL-17)*.

- Addressing the data sparsity issue in neural AMR parsing. Xiaochang Peng, Chuan Wang, Daniel Gildea, and Nianwen Xue. 2017. In *Proceedings of the European Chapter of the ACL (EACL-17)*.
- Aligning movies with scripts by exploiting temporal ordering constraints. Iftexhar Naim, Abdullah Al Mamun, Young Chol Song, Jiebo Luo, Henry Kautz, and Daniel Gildea. 2016. In *23rd International Conference on Pattern Recognition (ICPR)*.
- AMR-to-text generation as a traveling salesman problem. Linfeng Song, Yue Zhang, Xiaochang Peng, Zhiguo Wang, and Daniel Gildea. 2016. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-16)*.
- Sense embedding learning for word sense induction. Linfeng Song, Zhiguo Wang, Haitao Mi, and Daniel Gildea. 2016. In *Fifth Joint Conference On Lexical And Computational Semantics (\*SEM 2016)*, pages 85–90.
- Unsupervised alignment of actions in video with text descriptions. Young Chol Song, Iftexhar Naim, Abdullah Al Mamun, Kaustubh Kulkarni, Parag Singla, Jiebo Luo, Daniel Gildea, and Henry Kautz. 2016. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*.
- A synchronous hyperedge replacement grammar based approach for AMR parsing. Xiaochang Peng, Linfeng Song, and Daniel Gildea. 2015. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning (CoNLL-15)*, pages 731–739. Beijing.
- Discriminative unsupervised alignment of natural language instructions with corresponding video segments. Iftexhar Naim, Young Chol Song, Qiguang Liu, Liang Huang, Henry Kautz, Jiebo Luo, and Daniel Gildea. 2015. In *Proceedings of the 2015 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-15)*.
- Automated prediction and analysis of job interview performance: The role of what you say and how you say it. Iftexhar Naim, Mohammad Iftexhar Tanveer, Daniel Gildea, and Mohammed (Ehsan) Hoque. 2015. In *IEEE Conference on Automatic Face and Gesture Recognition, (FG)*, pages 1–6.
- Comparing representations of semantic roles for string-to-tree decoding. Marzieh Bazrafshan and Daniel Gildea. 2014. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-14)*.
- Type-based MCMC for sampling tree fragments from forests. Xiaochang Peng and Daniel Gildea. 2014. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-14)*.
- Unsupervised alignment of natural language instructions with video segments. Iftexhar Naim, Young Chol Song, Qiguang Liu, Henry Kautz, Jiebo Luo, and Daniel Gildea. 2014. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-14)*.
- Sliding alignment windows for real-time crowd captioning. Mohammad Kazemi, Rahman Lavaee, Iftexhar Naim, and Daniel Gildea. 2014. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 236–240. Association for Computational Linguistics, Baltimore, Maryland.
- Semantic roles for string to tree machine translation. Marzieh Bazrafshan and Daniel Gildea. 2013. In *Association for Computational Linguistics (ACL-13) short paper*.
- Plurality, negation, and quantification: Towards comprehensive quantifier scope disambiguation. Mehdi Manshadi, Daniel Gildea, and James Allen. 2013. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL-13)*.
- Integrating programming by example and natural language programming. Mehdi Manshadi, Daniel Gildea, and James Allen. 2013. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-13)*.

- [Text alignment for real-time crowd captioning](#). Iftekhhar Naim, Daniel Gildea, Walter Lasecki, and Jeffrey Bigham. 2013. In *Proceedings of the 2013 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-13)*.
- [Simultaneous word-morpheme alignment for statistical machine translation](#). Elif Eyigöz, Daniel Gildea, and Kemal Oflazer. 2013. In *Proceedings of the 2013 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-13)*.
- [Convergence of the EM algorithm for Gaussian mixtures with unbalanced mixing coefficients](#). Iftekhhar Naim and Daniel Gildea. 2012. In *International Conference on Machine Learning (ICML)*.
- [Improving the IBM alignment models using Variational Bayes](#). Darcey Riley and Daniel Gildea. 2012. In *Association for Computational Linguistics (ACL) short paper*, pages 306–310.
- [Tuning as linear regression](#). Marzieh Bazrafshan, Tagyoung Chung, and Daniel Gildea. 2012. In *North American chapter of the Association for Computational Linguistics (NAACL) short paper*, pages 543–547.
- [Optimal head-driven parsing complexity for linear context-free rewriting systems](#). Pierluigi Crescenzi, Daniel Gildea, Andrea Marino, Gianluca Rossi, and Giorgio Satta. 2011. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL-11)*, pages 450–459.
- [Issues concerning decoding with synchronous context-free grammar](#). Tagyoung Chung, Licheng Fang, and Daniel Gildea. 2011. In *Proceedings of the ACL 2011 Conference Short Papers*. Association for Computational Linguistics, Portland, Oregon.
- [Terminal-aware synchronous binarization](#). Licheng Fang, Tagyoung Chung, and Daniel Gildea. 2011. In *Proceedings of the ACL 2011 Conference Short Papers*. Association for Computational Linguistics, Portland, Oregon.
- [Effects of empty categories on machine translation](#). Tagyoung Chung and Daniel Gildea. 2010. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-10)*. Boston.
- [A fast fertility hidden Markov model for word alignment using MCMC](#). Shaojun Zhao and Daniel Gildea. 2010. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-10)*. Boston.
- [Semantic role features for machine translation](#). Ding Liu and Daniel Gildea. 2010. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING-10)*. Beijing.
- [Optimal parsing strategies for Linear Context-Free Rewriting Systems](#). Daniel Gildea. 2010. In *Proceedings of the 2010 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-10)*, pages 769–776.
- [Weight pushing and binarization for fixed-grammar parsing](#). Matt Post and Daniel Gildea. 2009. In *International Conference on Parsing Technologies*. Paris, France.
- [Bayesian learning of phrasal tree-to-string templates](#). Ding Liu and Daniel Gildea. 2009. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-09)*. Singapore.
- [Unsupervised tokenization for machine translation](#). Tagyoung Chung and Daniel Gildea. 2009. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-09)*. Singapore.
- [Bayesian learning of a tree substitution grammar](#). Matt Post and Daniel Gildea. 2009. In *Proc. Association for Computational Linguistics (short paper)*, pages 45–48. Singapore.
- [Efficient multi-pass decoding for synchronous context free grammars](#). Hao Zhang and Daniel Gildea. 2008. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL-08)*, pages 209–217. Columbus, Ohio.

- Bayesian learning of non-compositional phrases with synchronous parsing. Hao Zhang, Chris Quirk, Robert C. Moore, and Daniel Gildea. 2008. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL-08)*, pages 97–105. Columbus, Ohio.
- Extracting synchronous grammar rules from word-level alignments in linear time. Hao Zhang, Daniel Gildea, and David Chiang. 2008. In *Proceedings of the 22nd International Conference on Computational Linguistics (COLING-08)*, pages 1081–1088. Manchester, UK.
- Parsers as language models for statistical machine translation. Matt Post and Daniel Gildea. 2008. In *The Eighth Conference of the Association for Machine Translation in the Americas (AMTA 2008)*. Honolulu, HI.
- Optimizing grammars for minimum dependency length. Daniel Gildea and David Temperley. 2007. In *Proceedings of the 45th Annual Conference of the Association for Computational Linguistics (ACL-07)*, pages 184–191. Prague.
- Worst-case synchronous grammar rules. Daniel Gildea and Daniel Štefankovič. 2007. In *Proceedings of the 2007 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-07)*, pages 147–154.
- Source-language features and maximum correlation training for machine translation evaluation. Ding Liu and Daniel Gildea. 2007. In *Proceedings of the 2007 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-07)*, pages 41–48. Rochester, NY.
- Efficient search for inversion transduction grammar. Hao Zhang and Daniel Gildea. 2006. In *2006 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 224–231. Sydney.
- Inducing word alignments with bilexical synchronous trees. Hao Zhang and Daniel Gildea. 2006. In *Proceedings of the International Conference on Computational Linguistics/Association for Computational Linguistics (COLING/ACL-06) Poster Session*, pages 953–960. Sydney.
- Factoring synchronous grammars by sorting. Daniel Gildea, Giorgio Satta, and Hao Zhang. 2006. In *Proceedings of the International Conference on Computational Linguistics/Association for Computational Linguistics (COLING/ACL-06) Poster Session*, pages 279–286. Sydney.
- Stochastic iterative alignment for machine translation evaluation. Ding Liu and Daniel Gildea. 2006. In *Proceedings of the International Conference on Computational Linguistics/Association for Computational Linguistics (COLING/ACL-06) Poster Session*, pages 539–546. Sydney.
- Synchronous binarization for machine translation. Hao Zhang, Liang Huang, Daniel Gildea, and Kevin Knight. 2006. In *Proceedings of the 2006 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-06)*, pages 256–263. New York, NY.
- Stochastic lexicalized inversion transduction grammar for alignment. Hao Zhang and Daniel Gildea. 2005. In *Proceedings of the 43rd Annual Conference of the Association for Computational Linguistics (ACL-05)*, pages 475–482. Ann Arbor, MI.
- Dependencies vs. constituents for tree-based alignment. Daniel Gildea. 2004. In *2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 214–221. Barcelona.
- Syntax-based alignment: Supervised or unsupervised? Hao Zhang and Daniel Gildea. 2004. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-04)*, pages 418–424. Geneva, Switzerland.
- Skeletons in the parser: Using a shallow parser to improve deep parsing. Mary Swift, James Allen, and Daniel Gildea. 2004. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-04)*, pages 383–389. Geneva, Switzerland.

- [A smorgasbord of features for statistical machine translation.](#) Franz Josef Och, Daniel Gildea, Sanjeev Khudanpur, Anoop Sarkar, Kenji Yamada, Alex Fraser, Shankar Kumar, Libin Shen, David Smith, Katherine Eng, Viren Jain, Zhen Jin, and Dragomir Radev. 2004. In *Proceedings of the 2004 Meeting of the North American chapter of the Association for Computational Linguistics (NAACL-04)*, pages 161–168. Boston.
- [Loosely tree-based alignment for machine translation.](#) Daniel Gildea. 2003. In *Proceedings of the 41th Annual Conference of the Association for Computational Linguistics (ACL-03)*, pages 80–87. Sapporo, Japan.
- [Identifying semantic roles using Combinatory Categorical Grammar.](#) Daniel Gildea and Julia Hockenmaier. 2003. In *2003 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 57–64. Sapporo, Japan.
- [An algorithm for word-level alignment of parallel dependency trees.](#) Yuan Ding, Daniel Gildea, and Martha Palmer. 2003. In *The 9th Machine Translation Summit of the International Association for Machine Translation*, pages 95–101. New Orleans.
- [The necessity of syntactic parsing for predicate argument recognition.](#) Daniel Gildea and Martha Palmer. 2002. In *Proceedings of the 40th Annual Conference of the Association for Computational Linguistics (ACL-02)*, pages 239–246. Philadelphia, PA.
- [Probabilistic models of verb-argument structure.](#) Daniel Gildea. 2002. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING-02)*, pages 308–314. Taipei.
- [Corpus variation and parser performance.](#) Daniel Gildea. 2001. In *2001 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 167–202. Pittsburgh, PA.
- [Automatic labeling of semantic roles.](#) Daniel Gildea and Daniel Jurafsky. 2000. In *Proceedings of the 38th Annual Conference of the Association for Computational Linguistics (ACL-00)*, pages 512–520. Hong Kong.
- [Topic-based language models using EM.](#) Daniel Gildea and Thomas Hofmann. 1999. In *Proceedings of the 6th European Conference on Speech Communication and Technology (EUROSPEECH-99)*, pages 2167–2170. Budapest.
- [Forms of English function words – Effects of disfluencies, turn position, age and sex, and predictability.](#) Alan Bell, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, and Daniel Gildea. 1999. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS-99)*, pages 395–398. San Francisco, California.
- [A dynamic model of aspectual composition.](#) Nancy Chang, Daniel Gildea, and Srinu Narayanan. 1998. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society (COGSCI-98)*, pages 226–231. Lawrence Erlbaum, Madison, WI.
- [Applications of technology in providing transit information.](#) Daniel Gildea and Mikael Sheikh. 1996. In *Transportation Research Record 1521*, pages 71–76. Washington, DC.
- [Automatic induction of finite state transducers for simple phonological rules.](#) Daniel Gildea and Daniel Jurafsky. 1995. In *Proceedings of the 33rd Annual Conference of the Association for Computational Linguistics (ACL-95)*, pages 9–15. Cambridge, MA.

## Refereed Workshop Papers

- [Neural transition-based syntactic linearization](#). Linfeng Song, Yue Zhang, and Daniel Gildea. 2018. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 431–440. Association for Computational Linguistics.
- [University of Rochester WMT 2017 NMT system submission](#). Chester Holtz, Chuyang Ke, and Daniel Gildea. 2017. In *Proceedings of the Second Conference on Machine Translation*, pages 310–314. Association for Computational Linguistics, Copenhagen, Denmark.
- [UofR at SemEval-2016 Task 8: Learning Synchronous Hyperedge Replacement Grammar for AMR Parsing](#). Xiaochang Peng and Daniel Gildea. 2016. In *NAACL-HLT 2016 Workshop on Semantic Evaluation (SemEval-2016)*.
- [Approximate inference using unimodular graphs in dual decomposition](#). Chetan Bhole, Justin Domke, and Daniel Gildea. 2013. In *Optimization for Machine Learning Workshop, NIPS 2013*.
- [Multi-rate HMMs for word alignment](#). Elif Eyigöz, Daniel Gildea, and Kemal Oflazer. 2013. In *Eighth Workshop on Statistical Machine Translation (WMT-13)*.
- [SCFG latent annotation for machine translation](#). Tagyoung Chung, Licheng Fang, and Daniel Gildea. 2011. In *Proc. of the International Workshop on Spoken Language Translation (IWSLT)*.
- [Factors affecting the accuracy of Korean parsing](#). Tagyoung Chung, Matt Post, and Daniel Gildea. 2010. In *NAACL Workshop on Statistical Parsing of Morphologically Rich Languages*.
- [Language modeling with tree substitution grammars](#). Matt Post and Daniel Gildea. 2009. In *NIPS workshop on Grammar Induction, Representation of Language, and Language Learning*. Whistler, British Columbia.
- [Improved tree-to-string transducers for machine translation](#). Ding Liu and Daniel Gildea. 2008. In *ACL Workshop on Statistical Machine Translation (ACL08-SMT)*, pages 62–69. Columbus, Ohio.
- [Factorization of synchronous context-free grammars in linear time](#). Hao Zhang and Daniel Gildea. 2007. In *NAACL Workshop on Syntax and Structure in Statistical Translation (SSST)*, pages 25–32.
- [Syntactic features for evaluation of machine translation](#). Ding Liu and Daniel Gildea. 2005. In *ACL 2005 Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, pages 25–32. Ann Arbor, MI.
- [Machine translation as lexicalized parsing with hooks](#). Liang Huang, Hao Zhang, and Daniel Gildea. 2005. In *International Workshop on Parsing Technologies (IWPT05)*, pages 65–73.
- [Online statistics for a unification-based dialogue parser](#). Micha Elsner, Mary Swift, James Allen, and Daniel Gildea. 2005. In *International Workshop on Parsing Technologies (IWPT05)*, pages 198–199. Vancouver, BC.

## Invited Papers

- [Identifying semantic roles in text](#). Daniel Gildea and Daniel Jurafsky. 2001. In *Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01)*. Seattle, Washington. Appears as chapter in *Exploring Artificial Intelligence in the New Millennium*, Gerhard Lakemeyer and Bernhard Nebel, eds., Morgan Kaufmann, 2003.



## Technical Reports

- Feature-based decipherment for large vocabulary machine translation. Iftekhhar Naim and Daniel Gildea. 2015. *arXiv preprint 1508.02142*.
- Improving the performance of GIZA++ using Variational Bayes. Darcey Riley and Daniel Gildea. 2010. Technical Report 963, University of Rochester.
- Comparing optimization criteria and methods for word alignment. Tagyoung Chung, Shaojun Zhao, and Daniel Gildea. 2010. Technical Report 958, University of Rochester.
- Bayesian learning of tokenization for machine translation. Tagyoung Chung and Daniel Gildea. 2010. Technical Report 957, University of Rochester.
- Topic models for corpus-centric knowledge generalization. Benjamin Van Durme and Daniel Gildea. 2009. Technical Report 946, University of Rochester.
- Syntax for statistical machine translation. Franz Josef Och, Daniel Gildea, Sanjeev Khudanpur, Anoop Sarkar, Kenji Yamada, Alex Fraser, Shankar Kumar, Libin Shen, David Smith, Katherine Eng, Viren Jain, Zhen Jin, and Dragomir Radev. 2003. Technical report, Center for Language and Speech Processing, Johns Hopkins University, Baltimore. Summer Workshop Final Report.
- Natural language generation in the context of machine translation. Jan Hajič, Martin Čmejrek, Bonnie Dorr, Yuan Ding, Jason Eisner, Daniel Gildea, Terry Koo, Kristen Parton, Gerald Penn, Dragomir Radev, and Owen Rambow. 2002. Technical report, Center for Language and Speech Processing, Johns Hopkins University, Baltimore. Summer Workshop Final Report.
- Topic-based novelty detection. James Allan, Hubert Jin, Martin Rajman, Charles Wayne, Daniel Gildea, Victor Lavrenko, Rose Hoberman, and David Caputo. 1999. Technical report, Center for Language and Speech Processing, Johns Hopkins University, Baltimore. Summer Workshop Final Report.

## TEACHING EXPERIENCE

- Fall 2008, 2010, 2014 Graduate seminar on Optimization, U. of Rochester.
- Spring 2005, 2006, 2007, 2008, 2011, 2012, 2013, 2014, 2015 Graduate Machine Learning, U. of Rochester.
- Fall 2006, 2012 Graduate seminar on Language, Parsing, and Complexity, U. of Rochester.
- Fall 2005, 2007, 2009, 2011, 2013 Graduate Statistical Speech and Language Processing, U. of Rochester.
- Fall 2004 Undergraduate Design and Analysis of Efficient Algorithms, U. of Rochester.
- Spring 2004 Graduate seminar on Machine Translation, U. of Rochester.
- Fall 2002 Cotaught graduate seminar on Information Extraction, U. of Pennsylvania.

## STUDENTS ADVISED

Postdocs advised: Emily Prud'hommeaux (2014-2015, now faculty at RIT)

Graduated PhD students as primary adviser: Shan He (2006, now at Google), Hao Zhang (2008, now at Google), Ding Liu (2010, now at Nuance), Matt Post (2010, now at JHU), Shaojun Zhao (2011, now at Amazon), Tagyoung Chung (2012, now at Nuance), Licheng Fang (2013, now at Google), Elif Eyigöz (2014, now at IBM), Marzieh Bazrafshan (2014, now at H5), Iftekhar Naim (2015, now at Google), Xiaochang Peng (2018, now at Facebook).

Current PhD students at Rochester as primary advisor: Linfeng Song, Parker Riley, Lisa Jin.

Past PhD Thesis Committees at Rochester: Chengliang Zhang (graduated 2007), Constantin Rothkopf (2008), Fabrizio Morbini, Benjamin Van Durme (2009), Bin Wei, Qi Ge (2011), Satyaki Mahalanabis (2012), Naushad UzZaman (2012), Daphne Liu (2012), Curtis Menton (2013), Mehdi Manshadi (2013), Yi Chu (2013), Tivadar Papai (2013), Jonathan Gordon (2014), Young Chol Song (2016), Yuncheng Li (2017), Tim Kopp (2017), Omid Bakhshandeh (2017), Nasrin Mostafazadeh (2017), Jacob Brock (2018), Adam Purtee (2018).

Current PhD Thesis Committees at Rochester: Ritwick Bose, Gene Kim, Jansen Orfan, Matt Overlan, Taylan Sen, Iftekhar Tanveer, Xiangru Lian, Haofu Liao.

Outside PhD Thesis Committees: Marta Alegre (Technical U. of Catalonia, graduated August 2004), Edward Loper (U. of Pennsylvania, graduated September 2008), Milos Stanojevic (U. of Amsterdam, graduated December 2017), Jeffrey Flanigan, (Carnegie Mellon University, graduated December 2018).

Undergraduate research advising at Rochester: Micha Elsner, Darcey Riley.

## SPONSORED RESEARCH

Feature-Rich Decipherment for Machine Translation of Low-resource Languages, Google Faculty Award, Funded 2016-2017.

Scaling Up Discriminative Learning for Natural Language Understanding and Translation, NSF EAGER funded 2014-2015, co-PI with Liang Huang, Oregon State University.

Efficient Nonterminal Refinement for Syntax-Based Machine Translation, Google Faculty Award, Funded 2014-2015.

Real-Time Captioning of Group of Non-Experts for Deaf and Hard of Hearing Students, NSF small funded 2012-2015, co-PI with Jeff Bigham.

Richer Representations for Machine Translation, NSF large funded 2009-2013, joint project between Rochester, Colorado, USC/ISI, Columbia and Brandeis.

Semantics for Statistical Machine Translation, NSF Faculty Early Career Development (CAREER) Award funded 2005-2010.

DARPA BOLT Machine Translation program, Subcontract to SRI, funded 2011-2012.

Probabilistic Tree Transducers for Machine Translation and Natural Language Processing, NSF ITR grant funded 2004-2009, Kevin Knight (USC/ISI, PI), Daniel Marcu (USC/ISI, co-PI), Daniel Gildea (U. of Rochester, co-PI).

Domain-Independent Semantic Interpretation, NSF ITR grant funded 2003-2007, Daniel Jurafsky (U. of Colorado, PI), Daniel Gildea (U. of Rochester, co-PI), Wayne Ward (U. of Colorado, co-PI), Martha Palmer (U. of Pennsylvania, co-PI), Charles Fillmore (UC Berkeley, co-PI).

## PROFESSIONAL ACTIVITIES

- Reviewer, *Transactions of Association for Computational Linguistics, Speech Communication, ACM Transactions on Speech and Language Processing, Journal of Artificial Intelligence Research, Journal of Computer and System Sciences.*
- Chair of Nominating Committee, North American Chapter of the Association for Computational Linguistics (NAACL), 2015.
- Area Chair, Conference on Empirical Methods in Natural Language Processing (EMNLP 2014, 2017)
- Area Chair, Annual Meeting of the Association for Computational Linguistics (ACL 2008)
- Reviewer, Annual Meeting of the Association for Computational Linguistics (ACL 2000, 2003, 2005, 2007, 2008, 2009, 2010, 2012, 2013, 2014, 2015, 2016)
- Program Committee, Conference on Empirical Methods in Natural Language Processing (EMNLP 2002, 2005, 2006, 2007, 2008, 2009, 2012, 2013, 2014)
- Executive Board, North American Chapter of the Association for Computational Linguistics (NAACL), 2011-2012.
- Program Committee, Workshop on Statistical Machine Translation (2012, 2013, 2014, 2016)
- Reviewer, Conference of the Association for Machine Translation in the Americas (AMTA 2008, 2012)
- Editorial Board, *Computational Linguistics* (2008-2010)
- Local Arrangements Co-chair, Annual Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL 2007)
- Reviewer, American Assoc. for Artificial Intelligence conference (AAAI 2006, 2008)
- Reviewer, Conference on Natural Language Learning (CoNLL 2008)
- Reviewer, Advances in Neural Information Processing (NIPS 2004, 2005, 2006, 2007)
- Reviewer, International Joint Conference on Natural Language Processing (IJCNLP 2004, 2007)
- Reviewer, European Chapter of the Association for Computational Linguistics (EACL 2006, 2009, 2014)
- Reviewer, Annual Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL 2004, 2006, 2013, 2016)
- Program Committee, International Workshop on Scalable Natural Language Understanding (ScaNaLU 2004, 2006)
- Co-organizer, Conference on Natural Language Learning (CoNLL 2005)
- Panelist, Doctoral Consortium, American Assoc. for Artificial Intelligence (AAAI 2005)
- Reviewer, International Conference on Computational Linguistics (COLING 2004, 2010, 2012, 2014)
- Program Committee, *Journal of Natural Language Engineering* special issue on Parallel Texts (2004)
- Reviewer, *Machine Learning Journal*, Special Issue on Learning in Speech and Language Technologies (2004)
- Reviewer, *Journal of Artificial Intelligence Research* (2004)
- Area Chair, Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)
- Program Committee, International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+ 2004, 2012, 2016)
- Reviewer, Third International Conference on Language Resources and Evaluation (LREC 2002)
- Reviewer, *Journal of Machine Learning Research* special issue on Machine Learning Approaches to Shallow Parsing (2001)
- Reviewer, *Applied Intelligence* special issue on Neural Networks and Machine Learning for Natural Language Processing (2000)
- Reviewer, International Colloquium on Grammatical Inference (ICGI 1998, 2000)
- Member, Association for Computing Machinery, Association for Computational Linguistics.

#### **DEPARTMENT AND UNIVERSITY SERVICE**

CS faculty recruiting committee (2007, 2008, 2011, 2012, 2013, 2014, 2015, 2019)

CS graduate education committee (2015, 2019)

CS graduate admissions committee (2003–2006, co-chair 2017-2019)

CS lab committee (2006–2008)

University faculty council (2003–2004, 2007–2008)

Co-organizer of the interdepartmental cognitive science colloquium (2006–2015)

#### **HONORS AND AWARDS**

NSF Faculty Early Career Development (CAREER) Award, 2005–2010

IRCS postdoctoral fellowship, U. of Pennsylvania, 2001–2003

National Defense Science and Engineering Graduate Fellowship, 1997–2000

National Science Foundation Graduate Fellowship (declined), 1997

California Microelectronics Graduate Fellowship, 1996–1997