# KEVIN KNIGHT

## Contact

Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292
(310) 448-8716
knight@isi.edu

## Current

- Research Director and Fellow, USC/Information Sciences Institute.

- Dean's Professor of Computer Science (2014-), Research Professor (2011-2014), Research Associate Professor (2002-2010), Research Assistant Professor (1993-2002), Computer Science Department, University of Southern California.

- Chief Technical Officer (2017-), Specifio.

- Chief Scientist (2001-2010), SDL Language Weaver.

## Education

- **Carnegie Mellon University**. PhD, Computer Science, 1992. Dissertation title: "Integrating Knowledge Acquisition and Language Acquisition."

- **Harvard University**. BA, Computer Science, 1986.

## Professional

- **ACL Executive Council member** (Association for Computational Linguistics), 2009-2013.

- **ACL President** (Association for Computational Linguistics), 2011.

- **AAAI Council member** (Association for the Advancement of Artificial Intelligence), 2005-2007.

- **Editorial boards.** *ACM Transactions on Speech and Language Processing*, 2004-2007. *Computational Linguistics*, 1999-2001. *Journal of Artificial Intelligence Research*, 1999-2001. *Machine Translation*, 2007-.

- **General chair.** Conference of the Association for Computational Linguistics (ACL), 2005. Conference of the North American Association for Computational Linguistics (NAACL), 2016.

- **Program chair.** Conference of the North American Association for Computational Linguistics (NAACL), 2001.

- **Book review editor.** *AI Magazine*, 1994-1997.

- **Invited to lead** six-week intensive workshop on statistical machine translation (Language Engineering for Students and Professionals Integrating Research and Education, Johns Hopkins University, funded by NSF and DoD), Summer 1999.

# Invited talks

- "Statistical Machine Translation: Where it Went," Conference on Empirical Methods in Natural Language Processing (EMNLP), 1997.

- "Every Time I Fire a Statistician, I Get a Warm Fuzzy Feeling," Conference of the Association for Machine Translation in the Americas (AMTA), Cuernavaca, Mexico, 2000.

- "Deeper Representations for Machine Translation: Ready or Not?," Conference on Empirical Methods in Natural Language Processing (EMNLP), 2003.

- "What's New in Statistical Machine Translation," Conference on Uncertainty in AI (UAI), Acapulco, Mexico, 2003.

- "What's New in Statistical Machine Translation," International Joint Conference on AI (IJCAI), Edinburgh, Scotland, 2005.

- "Results on Decipherment Problems," ESSLLI, Heriot-Watt, Scotland, 2005.

- "Isn't Linguistic Structure Important, Asked the Engineer?" Conference of the European Association for Computational Linguistics (EACL), Trento, 2006.

- "Can a Machine Translate Without Knowing What a Verb Is?" Conference of the North American Association for Computational Linguistics, Workshop on Statistical Machine Translation (WMT), New York City, 2006.

- "No More Strings, Please," IEEE/ACL Spoken Language Technology Workshop, Aruba, 2006.

- "Capturing Practical Natural Language Transformations," Mathematics of Language Conference (MOL), Los Angeles, 2007.

- "Automated Language Translation Needs Generation Help Badly," International Conference on Natural Language Generation (NLG), Prague, 2007.

- "Weighted Transducers for Natural Language Processing," Weighted Automata: Theory and Applications (WATA), Dresden, 2008.

- "Morphology in Statistical Machine Translation: Integrate-in, or Tack-on?" NAACL Workshop on Syntax and Parsing of Morphologically Rich Languages (SPMRL), Los Angeles, 2010.

- "Theory and Practice of Translation," ACL Workshop on Applications of Tree Automata in Natural Language Processing, Stockholm, 2010.

- "What's on the Horizon? The Research Agenda," Translation Automation Users Society (TAUS) User Conference, Portland, 2010.

- "Why do We Call it Decoding?" Conference on Empirical Methods in Natural Language Processing (EMNLP), Cambridge, USA, 2010.

- "Human Language Technology," NITRD 20th Anniversary, 2012.

- "Transformation Frameworks for Machine Translation: Strings, Trees, and Graphs," Eleventh International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+11), 2012; Weighted Automata: Theory and Applications (WATA), 2012.

- "Human Language Technology," AAAS conference (Boston, MA), 2013.

- "Structures for Machine Translation: Strings, Trees, and Graphs," International Workshop on Computational Semantics (IWCS), Potsdam, 2013.

- "The DARPA Big Mechanisms Project," Assocation for Computational Linguistics, BioNLP Workshop (ACL-BioNLP), 2015.

- "What Is Neural Machine Translation Learning?" Association for Computational Linguistics, Workshop on Neural Machine Translation (ACL-NMT), 2017.

- "What Is Neural Machine Translation Learning?" Conference on Technologies and Applications of Artificial Intelligence (TAAI), 2017.

- Keynote talk, Conference of the North American Association for Computational Linguistics (NAACL), 2018.

# Teaching

- **University of Southern California**. *Natural Language Processing*, annually from 1993 to 1998. *Advanced Natural Language Processing*, annually from 2000.

- **Stanford University**. *Statistical Machine Translation*, short course, Linguistic Society of America Summer School (2007).

- **University of California, Los Angeles**. *Machine Translation*, short course, UCLA Extension (1996, 1997).

- **University of Aberdeen**. *Machine Translation*, short course (2005).

- **Conference tutorials. ACL 1996**: *Machine Translation*. **NAACL 2000**: *Machine Translation*. **ACL 2000**: *Statistical Machine Translation*. **AMTA 2000**: *Statistical Machine Translation*. **MT Summit IX 2003**: *Statistical Machine Translation*. **AMTA 2004**: *Statistical Machine Translation*. **NAACL 2004**: *Machine Translation*. **ACL 2006**: *Synchronous Grammars and Tree Transducers*. **MT Summit XI 2007**: *Statistical Machine Translation*. **ACL, 2009**: *Topics in Statistical Machine Translation*. **NAACL, 2009**: *Writing Systems, Transliteration, and Decipherment*. **ACL, 2013**: *Decipherment*.

- **Harvard University**. Teaching Assistant (1983-1985). Certificate of Distinction in Teaching, Committee on Undergraduate Education.

- **Beijing, People's Republic of China**. English instructor at the Mining College in Beijing (1981).

# Grants Awarded

- "Communicating with Intelligent Computers." 2015-2019. DARPA-funded.

- "Exploiting Language Information for Situational Awareness." 2015-2019. DARPA-funded.

- "Decipherment of Historical Manuscripts." 2015-2018. Funded by Swedish equivalent of NSF.

- "Machine Translation of Informal Text and Speech." 2011-2016. (with D. Marcu). DARPA-funded.

- "MURI: Machine Translation for Low-Density Languages." 2010-2014. (with J. Carbonell, R. Barzilay, and J. Baldridge). Army-funded.

- "Richer Representations for Machine Translation." 2009-2013. (with M. Palmer, N. Xue, D. Gildea, and K. McKeown). NSF-funded.

- "Deciphering Natural Language." 2009-2013. (with R. Barzilay). NSF-funded.

- "Improvements to Speech-to-Speech Translation." 2008-09. DARPA-funded.

- "Decipherment Techniques Applied to Machine Translation." 2007-08. DARPA-funded.

- "Improving Fundamental Accuracy of Speech-to-Speech Translation." 2006-07. (with D. Marcu). DARPA-funded.

- "Human Informed Statistical Machine Translation." 2005-2011. (with D. Marcu). DARPA-funded.

- "Advanced Language Modeling for Machine Translation." 2005-06.

- "Iraqi-English Speech-to-Speech Deployment." 2005-06. DARPA-funded.

- "Treeworld: Probabilistic Tree Transducers for Machine Translation and Natural Language Processing." 2004-09. (with D. Gildea and D. Marcu). NSF-funded.

- "Translation Technology for Multilingual Translation Modeling." 2003-07. (with M. Ostendorf, J. Bilmes, K. Kirchoff, and D. Marcu). NSF-funded.

- "Machine Translation for a Handheld Speech Translator." 2002-04. (with D. Marcu). DARPA-funded.

- "Transmodel: US-Egypt Cooperative Project on Machine Translation." (with A. Rafea). 2003-04. NSF-funded.

- "Advanced Generation for Question Answering." 2002-03. (with D. Marcu). ARDA-funded.

- "ReWRITE: A Probabilistic Approach to Rewriting for Machine Translation, Language Generation, and Abstracting." 2000-02. (with D. Marcu). DARPA-funded.

- "Robust, Scalable Language Generation Using Symbolic and Statistical Techniques." 1999-2001. NSF-funded.

- "Designing Robust Technology for Automatic Language Generation." 1998-99. (with E. Hovy). DARPA-funded.

- "Building a Standard Platform for Statistical Machine Translation." 1999. NSF-funded.

- "Arabic-to-English Machine Translation." 1996-99. (with E. Hovy). DoD-funded.

- "FSTRANS: Performance- and User-Oriented Enhancements to Machine Translation." 1996. (with E. Hovy). DoD-funded.

## PhD Students

- Dr. Qing Dou, PhD thesis (2015), "Beyond Parallel Data—A Decipherment Approach for Better Quality Machine Translation."

- Dr. Steven DeNeefe, PhD thesis (2011), "Tree-Adjoining Machine Translation."

- Dr. Sujith Ravi, PhD thesis (2011), "Deciphering Natural Language."

- Dr. Victoria Fossum, PhD thesis (2010), "Integrating Parsing and Word Alignment in Syntax-Based Machine Translation." (*University of Michigan PhD, advised by Steven Abney*).

- Dr. Jonathan May. PhD thesis (2010), "Weighted Tree Automata and Transducers for Syntactic Natural Language Processing."

- Dr. Liang Huang. PhD thesis (2008), "Forest-Based Algorithms in Natural Language Processing." (*University of Pennsylvania PhD, co-advised with Aravind Joshi*).

- Dr. Philipp Koehn. PhD thesis (2003), "Noun Phrase Translation."

- Dr. Kenji Yamada. PhD thesis (2002), "A Syntax-Based Statistical Translation Model."

- Dr. Yaser Al-Onaizan. PhD thesis (2002), "Translating Named Entities Using Bilingual and Monolingual Resources."

- Dr. Irene Langkilde. PhD thesis (2002), "Automatic Sentence Generation Using a Hybrid Statistical Model of Lexical Collocations and Syntactic Relations."

- Dr. Ishwar Chander. PhD thesis (1998), "Automated Postediting of Documents."

- **Served on PhD dissertation committee:** Simon Woo (USC), Jason Riesa (USC), Sudeep Gandhe (USC), Markos Mylonakis (UvA), Vamshi Ambati (CMU), Markus Saers (Uppsala), Gholamreza Haffari (SFU), Rahul Bhagat (USC), Matt Michelson (USC), Antonio Roque (USC), Ashish Venugopal (CMU), Hao Zhang (Rochester), Alexander Fraser (USC), Qing Luo (USC), Colin Cherry (Alberta), Yuk Wah Wong (UTexas), Feng Pan (USC), Chris Callison-Burch (Edinburgh), Dragos Stefan Munteanu (USC), Yuan Ding (Penn), Deepak Ravichandran (USC), Noemie Elhadad (Columbia), David Evans (Columbia), Hal Daume (USC), Soomin Kim (USC), Radu Soricut (USC), Chad Langley (CMU), Liang Zhou (USC), Fei Huang (CMU), Shou-de Lin (USC), Greg Kondrak (Toronto), Sheila Tejada (USC), Michael Harm (USC).

- **Supervised summer internships: 2017:** Yining Chen (Dartmouth), Leon Cheung (UCSD), Sorcha Gilroy (Edinburgh), Nelson Liu (UW), Alexandra (Sasha) Mayn (Carleton College). **2016:** Nada Aldarrab (USC), Angeliki Laziradou (U. Trento), Xiang Li (U. Chicago), Sebastian Mielke (Dresden Univ. Technology), Ke Tran (U. Amsterdam). **2015:** Callum O'Shaughnessy (Queens University), Sudha Rao (Maryland), Wenduan Xu (Cambridge), Barret Zoph (USC). **2014:** Julian Schamper (Aachen), Eunsol Choi (Washington), Allen Schmaltz (Harvard), Matic Horvat (Cambridge). **2013:** Daniel Bauer (Columbia), Fabienne Braune (Stuttgart), Elliot Meyerson (Wesleyan), Malte Nuhn (Aachen), Arvind Neelakantan (Columbia). **2012:** Jacob Andreas (Columbia), Daniel Bauer (Columbia), Karl Moritz Hermann (Oxford), Bevan Jones (Edinburgh/Macquarrie), Nathan Schneider (CMU), Ada Wan (CUNY). **2011:** Licheng Fang (Rochester), Sravana Reddy (Chicago), Xuchen Yao (JHU). **2010:** Yoav Goldberg (Ben-Gurion), Ann Irvine (JHU), Sravana Reddy (Chicago), Sasha Rush (MIT). **2009:** Michael Auli (University of Edinburgh), Erica Greene (Haverford), Adam Pauls (UC Berkeley), Paramveer Dhillon (Penn). **2008:** Amittai Axelrod (University of Washington), John DeNero (UC Berkeley), Kyle Gorman (Penn Linguistics), Catalin Tirnauca (Universitat Rovira i Virgili). **2007:** Michael Bloodgood (Delaware), Jennifer Gillenwater (Rice University), Carmen Heger (Dresden), Wei Ho (Princeton). **2006:** Joseph Turian (NYU), Chenhai Xi (Pitt), Victoria Fossum (Michigan), Liang Huang (Penn), Jason Riesa (JHU), Oana-Diana Postolache (Saarland). **2005:** Victoria Fossum (Michigan), Mark Hopkins (UCLA), Liang Huang (Penn), Behrang Mohit (Pitt), Preslav Nakov (Berkeley), Jason Riesa (JHU), Hao Zhang (Rochester). **2004:** Madhur Ambastha (Rochester), Michel Galley (Columbia), David Kauchak (UCSD). **2003:** Michel Galley (Columbia), Mark Hopkins (UCLA), Beata Klebanov (Hebrew University), Ana-Maria Popescu (University of Washington), Lara Taylor (UCSD). **2002:** Chris Ackerman (USC), Emil Ettelaie (USC), Yuling Hsueh (USC), John Lee (Waterloo/MIT), Bo Pang (Cornell). **2001:** Abdessamad Echihabi (USC), Hal Daume (CMU), Michael Laszlo (Waterloo), Dragos Stefan Munteanu (Iowa), Rebecca Rees (BYU), Radu Soricut (Iowa).

## Honors

- Fellow of ACL (Association for Computational Linguistics), 2014-.

- Fellow of AAAI (Association for the Advancement of Artificial Intelligence), 2014-.

- USC/Information Sciences Institute: Institute Achievement Award, 2014.

- Best Demo Award, "Hafez: an Interactive Poetry Generation System" (M. Ghazvininejad, X. Shi, J. Priyadarshi, and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), Demonstrations Track, 2017.

- Best Paper Award, "11,001 New Features for Statistical Machine Translation" (D. Chiang, K. Knight, and W. Wang), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), 2009.

- Best Paper Award, "Fast Decoding and Optimal Decoding for Machine Translation" (U. Germann, M. Jahr, K. Knight, D. Marcu, and K. Yamada), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2001.

- Outstanding Paper Award, "Statistics-Based Summarization—Step One: Sentence Compression," (K. Knight and D. Marcu), National Conference on Artificial Intelligence (AAAI), 2000.

- ISI Meritorious Service Award, for outstanding work in web-based research-software distribution, 2001.

- Fellow of ISI (USC Information Sciences Institute), 2006.

- Harvard University, Certificate of Distinction in Teaching, Committee on Undergraduate Education, 1984.

# Patents Awarded

- "System and Method for Capitalizing Machine Translated Text" (W. Wang, K. Knight, and D. Marcu), US 8886518, issued November 11, 2014.

- "Modification of Annotated Bilingual Segment Pairs in Syntax-based Machine Translation" (W. Wang, J. May, and K. Knight), US 8825466, issued September 2, 2014.

- "Training for a Text-to-text Application which Uses String to Tree Conversion for Training and Decoding" (K. Knight, M. Galley, M. Hopkins, D. Marcu, I. Thayer), US 8600728, issued December 2, 2013.

- "Building a Translation Lexicon from Comparable, Non-parallel Corpora" (D. Marcu, K. Knight, D. Munteanu, P. Koehn), US 8234106, issued July 31, 2012.

- "Syntax-based Statistical Translation Model" (K. Yamada and K. Knight), US 8214196, issued July 3, 2012.

- "Identifying Documents which form Translated Pairs, within a Document Collection" (I. Muslea, K. Knight, D. Marcu), US 7813918, issued October 12, 2010.

- "Training Tree Transducers for Probabilistic Operations" (J. Graehl and K. Knight), US 7698125, issued April, 13, 2010.

- "Named Entity Translation" (Y. Al-Onaizan and K. Knight), US 7580830, issued August 25, 2009.

- "Adapter for Allowing both Online and Offline Training of a Text to Text System" (K. Yamada, G. Langmead), US 7624020, issued November 24, 2009.

- "Statistical Machine Translation" (P. Koehn and K. Knight), US 7624005, issued November 24, 2009.

- "Constructing a Translation Lexicon from Comparable, Non-Parallel Corpora" (D. Marcu, K. Knight, D. Munteanu, and P. Koehn), US 7620538, issued November 17, 2009.

- "Phrase to Phrase Joint Probability Model for Statistical Machine Translation" (D. Marcu, W. Wong, K. Knight, and P. Koehn), US 7454326, issued November 18, 2008.

- "Task Parallelization in a Text-to-Text System" (G. Langmead, K. Yamada, K. Knight, and D. Marcu), US 7389222, issued June 17, 2008.

- "Statistical Translation Using a Large Monolingual Corpus" (R. Soricut, D. Marcu, and K. Knight), US 7340388, issued March 4, 2008.

- "Named Entity Transliteration" (Y. Al-Onaizan and K. Knight), US 7249013, issued July 24, 2007.

- "Integer Programming Decoder for Machine Translation" (K. Knight and K. Yamada), US 7177792, issued February 13, 2007.

# Publications

**Books**

- *Artificial Intelligence, Second Edition* (E. Rich and K. Knight), McGraw-Hill Book Company, 1991.

**Journal Articles**

- "Humans Outperform Machines at the Bilingual Shannon Game" (M. Ghazvininejad and K. Knight), *Entropy* 19(1), 2017.

- "Incident-Driven Machine Translation and Name Tagging for Low-resource Languages" (U. Hermjakob, Q. Li, D. Marcu, J. May, S. J. Mielke, N. Pourdamghani, M. Pust, X. Shi, K. Knight, T. Levinboim, K. Murray, D. Chiang, B. Zhang, X. Pan, D. Lu, Y. Lin, and H. Ji), *Machine Translation*, 2017.

- "From Image to Translation: Processing the Endandered Nyushu Script" (T. Zhang, A. Chowdhury, N. Dhulekar, J. Xia, K. Knight, H. Ji, B. Yener, L. Zhao), *ACM Transactions on Asian and Low-Resource Language Information Processing*, 15(4), 2016.

- "The Secrets of the Copiale Cipher," (K. Knight, B. Megyesi, and C. Schaefer), *Journal of Research into Freemasonry and Fraternalism*, 2(2), 2012.

- "Does GIZA++ Make Search Errors?" (S. Ravi and K. Knight), *Computational Linguistics*, 36(3), 2010.

- "Re-structuring, Re-labeling, and Re-aligning for Syntax-Based Statistical Machine Translation," (W. Wang, J. May, K. Knight, and D. Marcu), *Computational Linguistics*, 36(2), 2010.

- "Binarization of Synchronous Context-Free Grammars," (L. Huang, H. Zhang, D. Gildea, and K. Knight), *Computational Linguistics*, 35(4), 2009.

- "The Power of Extended Top-Down Tree Transducers," (A. Maletti, J. Graehl, M. Hopkins, and K. Knight), *SIAM J. Comput.*, 39(2), 2009.

- "Attacking Decipherment Problems Optimally with Low-Order N-gram Models," (S. Ravi and K. Knight), *Cryptologia*, 33(4), 2009.

- "Training Tree Transducers," (J. Graehl, K. Knight, and J. May), *Computational Linguistics*, 34(3), 2008.

- "Capturing Practical Natural Language Transformations," (K. Knight), *Machine Translation*, 21(2), 2007.

- "Discovering the Linear Writing Order of a Two-Dimensional Ancient Hieroglyphic Script," (Shou-de Lin and K. Knight), *Artificial Intelligence*, 170(4-5), 2006.

- "Translation with Scarce Bilingual Resources," (Y. Al-Onaizan, U. Germann, U. Hermjakob, K. Knight, P. Koehn, D. Marcu, K. Yamada), *Machine Translation*, 2003.

- "Fast Decoding and Optimal Decoding for Machine Translation," (U. Germann, M. Jahr, K. Knight, D. Marcu, and K. Yamada), *Artificial Intelligence*, 2003.

- "Summarization Beyond Sentence Extraction: A Probabilistic Approach to Sentence Compression," (K. Knight and D. Marcu), *Artificial Intelligence*, 139(1), 2002.

- "Decoding Complexity in Word-Replacement Translation Models," (K. Knight), *Computational Linguistics*, 25(4), 1999.

- "Mining Online Text," *Communications of the ACM*, (K. Knight), 42(11), 1999.

- "Machine Transliteration," (K. Knight and J. Graehl), *Computational Linguistics* 24(4), 1998.

- "Integrating Knowledge Acquisition and Language Acquisition," (K. Knight), *Applied Intelligence*, 1(1), 1992.

- "Connectionist Ideas and Algorithms," *Communications of the ACM*, 33(11), November, 1990.

- "Knowledge and Natural Language Processing," (J. Barnett, K. Knight, I. Mani, and E. Rich), *Communications of the ACM*, 33(8), August, 1990.

- "Unification: A Multidisciplinary Survey," (K. Knight), *ACM Computing Surveys*, 21(1), March, 1989.

**Refereed Conference Papers**

- "Deciphering Related Languages" (N. Pourdamghani and K. Knight), Proc. EMNLP, 2017.

- "Hafez: an Interactive Poetry Generation System" (M. Ghazvininejad, X. Shi, J. Priyadarshi, and K. Knight), Proc. ACL Demonstrations Track, 2017.

- "Speeding up Neural Machine Translation Decoding by Shrinking Run-time Vocabulary" (X. Shi and K. Knight), Proc. ACL, 2017.

- "Cross-lingual Name Tagging and Linking for 282 Languages" (X. Pan, B. Zhang, J. May, J. Nothman, K. Knight and H. Ji), Proc. ACL, 2017.

- "Team ELISA System for DARPA LORELEI Speech Evaluation 2016" (P. Papadopoulos, R. Travadi, C. Vaz, N. Malandrakis, U. Hermjakob, N. Pourdamghani, M. Pust, B. Zhang, X. Pan, D. Lu, Y. Lin, O. Glembek, M. Karthick B, M. Karafiat, L. Burget, M. Hasegawa-Johnson, H. Ji, J. May, K. Knight, and S. Narayanan), Proc. Interspeech, 2017.

- "Generating Topical Poetry" (M. Ghazvininejad, X. Shi, Y. Choi, and K. Knight), Proc. EMNLP, 2016.

- "Why Neural Translations are the Right Length" (X. Shi, K. Knight, and D. Yuret), Proc. EMNLP, 2016.

- "Does String-Based Neural MT Learn Source Syntax?" (X. Shi, I. Padhi, and K. Knight), Proc. EMNLP, 2016.

- "Transfer Learning for Low-Resource Neural Machine Translation" (B. Zoph, D. Yuret, J. May, and K. Knight), Proc. EMNLP, 2016.

- "Generating English from Abstract Meaning Representation" (N. Pourdamghani and K. Knight), Proc. INLG, 2016.

- "Grapheme-to-Phoneme Models for (Almost) Any Language" (A. Deri and K. Knight), Proc. ACL, 2016.

- "A Multi-media Approach to Cross-lingual Entity Knowledge Transfer" (D. Lu, X. Pan, N. Pourdamghani, H. Ji, S.-F. Chang and K. Knight), Proc. ACL, 2016.

- "Multi-Source Neural Translation" (B. Zoph and K. Knight), Proc. NAACL, 2016.

- "Simple, Fast Noise Contrastive Estimation for Large RNN Vocabularies" (B. Zoph, A. Vaswani, J. May, and K. Knight), Proc. NAACL, 2016.

- "Name Tagging for Low-Resource Incident Languages Based on Expectation-Driven Learning" (B. Zhang, X. Pan, T. Wang, A. Vaswani, H. Ji, K. Knight, and D. Marcu), Proc. NAACL, 2016.

- "Extracting Structured Scholarly Information from the Machine Translation Literature" (E. Choi, M. Horvat, J. May, K. Knight, D. Marcu), Proc. LREC, 2016.

- "Parsing English into Abstract Meaning Representation Using Syntax-Based Machine Translation" (M. Pust, U. Hermjakob, K. Knight, D. Marcu, and J. May), Proc. EMNLP, 2015.

- "How Much Information Does a Human Translator Add to the Original?" (B. Zoph, M. Ghazvininejad, and K. Knight), Proc. EMNLP, 2015.

- "Unifying Bayesian Inference and Vector Space Models for Improved Decipherment" (Q. Dou, A. Vaswani, K. Knight, and C. Dyer), Proc. ACL, 2015.

- "How to Make a Frenemy: Multitape FSTs for Portmanteau Generation" (A. Deri and K. Knight), Proc. NAACL, 2015.

- "How to Memorize a Random 60-Bit String" (M. Ghazvininejad and K. Knight), Proc. NAACL, 2015.

- "Unsupervised Entity Linking with Abstract Meaning Representation" (X. Pan, T. Cassidy, U. Hermjakob, H. Ji, and K. Knight), Proc. NAACL, 2015.

- "Mapping between English Strings and Reentrant Semantic Graphs" (F. Braune, D. Bauer, and K. Knight), Proc. LREC, 2014.

- "How to Speak a Foreign Language without Knowing It" (X. Shi, K. Knight, and H. Ji), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2014.

- "Be Appropriate and Funny: Automatic Entity Morph Encoding" (B. Zhang, H. Huang, X. Pan, H. Ji, K. Knight, Z. Wen, Y. Sun, J. Han, B. Yener), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2014.

- "Aligning Context-based Statistical Models of Language with Brain Activity During Reading" (L. Wehbe, A. Vaswani, K. Knight and T. Mitchell), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.

- "Aligning English Strings with Abstract Meaning Representation Graphs" (N. Pourdamghani, Y. Gao, U. Hermjakob and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.

- "Beyond Parallel Data: Joint Word Alignment and Decipherment Improves Machine Translation" (Q. Dou, A. Vaswani and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.

- "Cipher Type Detection" (M. Nuhn and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.

- "Dependency-Based Decipherment for Resource-Limited Machine Translation" (Q. Dou and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2013.

- "Smatch: an Evaluation Metric for Semantic Feature Structures," (S. Cai and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2013.

- "Parsing Graphs with Hyperedge Replacement Grammars," (D. Chiang, J. Andreas, D. Bauer, K. M. Hermann, B. Jones, and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2013.

- "Curating and Contextualizing Twitter Stories to Assist with Social Newsgathering," (A. Zubiaga, H. Ji, and K. Knight), Proceedings of the Conference on Intelligent User Interfaces (IUI), 2013.

- "Large Scale Decipherment for Out-of-Domain Machine Translation," (Q. Dou and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2012.

- "Decoding Running Key Ciphers," (S. Reddy and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2012.

- "Semantics-Based Machine Translation with Hyperedge Replacement Grammars," (B. Jones, J. Andreas, D. Bauer, K-M. Hermann, and K. Knight), Proceedings of the International Conference on Computational Linguistics (COLING), 2012.

- "Bayesian Inference for Zodiac and Other Homophonic Ciphers," (S. Ravi and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2011.

- "Deciphering Foreign Language," (S. Ravi and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2011.

- "Unsupervised Discovery of Rhyme Schemes," (S. Reddy and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2011.

- "Fast, Greedy Model Minimization for Unsupervised Tagging," (S. Ravi, A. Vaswani, K. Knight, and D. Chiang), Proceedings of the International Conference on Computational Linguistics (COLING), 2010.

- "A Statistical Model for Lost Language Decipherment," (B. Snyder, R. Barzilay, and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2010.

- "Efficient Inference Through Cascades of Weighted Tree Transducers," (J. May, K. Knight, and H. Vogler), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2010.

- "Minimized Models and Grammar-Informed Initialization for Supertagging with Highly Ambiguous Lexicons," (S. Ravi, J. Baldridge, and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2010.

- "Unsupervised Syntactic Alignment with Inversion Transduction Grammars," (A. Pauls, D. Klein, D. Chiang, and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), 2010.

- "Bayesian Inference for Finite-State Transducers," (D. Chiang, J. Graehl, K. Knight, A. Pauls, and S. Ravi), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), 2010.

- "Synchronous Tree Adjoining Machine Translation," (S. DeNeefe and K. Knight), Conference on Empirical Methods in Natural Language Processing (EMNLP), 2009.

- "Minimized Models for Unsupervised Part-of-Speech Tagging," (S. Ravi and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2009.

- "Fast Consensus Decoding over Translation Forests," (J. DeNero, D. Chiang, and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2009.

- "Learning Phoneme Mappings for Transliteration without Parallel Data," (S. Ravi and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), 2009.

- "11,001 New Features for Statistical Machine Translation," (D. Chiang, K. Knight, and W. Wang), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), 2009.

- "Combining Constituent Parsers," (V. Fossum and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), short paper, 2009.

- "Faster MT Decoding through Pervasive Laziness," (M. Pust and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL), short paper, 2009.

- "Probabilistic Methods for a Japanese Syllable Cipher," (S. Ravi and K. Knight), Proceedings of the International Conference on the Computer Processing of Oriental Languages (ICCPOL), 2009.

- "Name Translation in Statistical Machine Translation: Learning When to Transliterate," (U. Hermjakob, K. Knight, and H. Daume), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2008.

- "Attacking Decipherment Problems Optimally with Low-Order N-gram Models," (S. Ravi and K. Knight), Conference on Empirical Methods in Natural Language Processing (EMNLP), 2008.

- "Automatic Prediction of Parser Accuracy," (S. Ravi and K. Knight), Conference on Empirical Methods in Natural Language Processing (EMNLP), 2008.

- "Overcoming Vocabulary Sparsity in MT Using Lattices," (S. DeNeefe, U. Hermjakob, and K. Knight), Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), 2008.

- "Using Bilingual Chinese-English Word Alignments to Resolve PP-Attachment Ambiguity in English," (V. Fossum and K. Knight), Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), student session, 2008.

- "Syntactic Re-Alignment Models for Machine Translation," (J. May and K. Knight), Conference on Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2007.

- "What Can Syntax-Based MT Learn from Phrase-Based MT?" (S. DeNeefe, K. Knight, W. Wang, and D. Marcu), Conference on Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2007.

- "Binarizing Syntax Trees to Improve Syntax-Based Machine Translation Accuracy," (W. Wang, K. Knight, and D. Marcu), Conference on Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2007.

- "Unsupervised Analysis for Decipherment Problems," (K. Knight, A. Nair, N. Rathod, and K. Yamada), Proceedings of the Conference of the Association for Computational Linguistics (ACL-COLING), 2006.

- "SPMT: Statistical Machine Translation with Syntactified Target Language Phrases," (D. Marcu, W. Wang, A. Echihabi, and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2006.

- "Scalable Inference and Training of Context-Rich Syntactic Models," (M. Galley, J. Graehl, K. Knight, D. Marcu, S. DeNeefe, W. Wang, and I. Thayer), Proceedings of the Conference of the Association for Computational Linguistics (ACL-COLING), 2006.

- "A Better N-Best List: Practical Determinization of Weighted Finite Tree Automata," (J. May and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2006.

- "Capitalizing Machine Translation," (W. Wang, K. Knight, D. Marcu), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2006.

- "Synchronous Binarization for Machine Translation," (H. Zhang, L. Huang, D. Gildea, K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2006.

- "Relabeling Syntax Trees to Improve Syntax-Based Machine Translation Quality," (B. Huang and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2006.

- "Transonics: A Practical Speech-to-Speech Translator for English-Farsi Medical Dialogs," (R. Belvin, E. Ettelaie, S. Gandhe, P. Georgiou, K. Knight, D. Marcu, S. Millward, S. Narayanan, H. Neely, D. Traum), ACL-2005 poster/demo.

- "Interactively Exploring a Machine Translation Model," (Steve DeNeefe, Kevin Knight, and Hayward H. Chan), ACL-2005 poster/demo.

- "Machine Translation in the Year 2004," (K. Knight, D. Marcu), Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2005.

- "An Overview of Probabilistic Tree Transducers for Natural Language Processing," (K. Knight, J. Graehl), Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Lecture Notes in Computer Science, Springer Verlag, 2005.

- "What's in a Translation Rule?" (M. Galley, M. Hopkins, K. Knight, and D. Marcu), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2004.

- "Training Tree Transducers," (J. Graehl and K. Knight), Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2004.

- "Syntax-based Language Models for Machine Translation," (E. Charniak, K. Knight, and K. Yamada), Proceedings of the Machine Translation Summit IX, 2003.

- "Empirical Methods for Compound Splitting," (P. Koehn and K. Knight), Proceedings of the Conference of the European Association for Computational Linguistics (EACL), 2003.

- "Syntax-based Alignment of Multiple Translations: Extracting Paraphrases and Generating New Sentences," (B. Pang, K. Knight, and D. Marcu, Proceedings of the Conference of the North American Association for Computational Linguistics (NAACL-HLT), 2003.

- "Using a Large Monolingual Corpus to Improve Translation Accuracy," (R. Soricut, K. Knight, and D. Marcu), Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), 2002.

- "Named Entity Translation: Extended Abstract," (Y. Al-Onaizan and K. Knight), Proceedings of the Human Language Technology Conference (HLT), 2002.

- "Translating Named Entities Using Monolingual and Bilingual Resources," (Y. Al-Onaizan and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2002.

- "A Decoder for Syntax-Based Statistical MT," (K. Yamada and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2002.

- "A Syntax-Based Statistical Translation Model," (Yamada, K. and Knight, K.), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2001.

- "Fast Decoding and Optimal Decoding for Machine Translation," (U. Germann, M. Jahr, K. Knight, D. Marcu, and K. Yamada), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 2001.

- "Knowledge Sources for Word-Level Translation Models," (P. Koehn and K. Knight), Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2001.

- "Translating with Scarce Resources," (Y. Al-Onaizan, U. Germann, U. Hermjakob, K. Knight, P. Koehn, D. Marcu, K. Yamada), National Conference on Artificial Intelligence (AAAI), 2000.

- "Preserving Ambiguities in Generation via Automata Intersection," (K. Knight and I. Langkilde), National Conference on Artificial Intelligence (AAAI), 2000.

- "Estimating Word Translation Probabilities from Unrelated Monolingual Corpora Using the EM Algorithm," (P. Koehn and K. Knight), National Conference on Artificial Intelligence (AAAI), 2000.

- "Statistics-Based Summarization—Step One: Sentence Compression," (K. Knight and D. Marcu), National Conference on Artificial Intelligence (AAAI), 2000.

- "Translation with Finite-State Devices," (K. Knight and Y. Al-Onaizan), Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), 1998.

- "Generation that Exploits Corpus-based Statistical Knowledge," (I. Langkilde and K. Knight), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 1998.

- "Learning Word Meanings by Instruction," (K. Knight), Proceedings of the National Conference on Artificial Intelligence (AAAI), 1996.

- "Two-Level, Many-Paths Generation," (K. Knight and V. Hatzivassiloglou), Proceedings of the Conference of the Association for Computational Linguistics (ACL), 1995.

- "Filling Knowledge Gaps in a Broad-Coverage MT System," (K. Knight, I. Chander, M. Haines, V. Hatzivassiloglou, E. Hovy, M. Iida, S. K. Luk, R. Whitney, and K. Yamada), Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 1995.

- "Unification-Based Glossing," (V. Hatzivassiloglou and K. Knight), Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 1995.

- "Integrating Knowledge Bases and Statistics in MT," (K. Knight, I. Chander, M. Haines, V. Hatzivassiloglou, E. Hovy, M. Iida, S. Luk, A. Okumura, R. Whitney, K. Yamada), Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), 1994.

- "Building a Large-Scale Knowledge Base for Machine Translation," (K. Knight and S. Luk), Proceedings of the National Conference on Artificial Intelligence (AAAI), 1994.

- "Automated Postediting of Documents," (K. Knight and I. Chander), Proceedings of the National Conference on Artificial Intelligence (AAAI), 1994.

- "Are Many Reactive Agents Better Than a Few Deliberative Ones?" (K. Knight), Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 1993.

- "Capturing Language-Specific Semantic Distinctions in Interlingua-Based MT," (J. Barnett, I. Mani, E. Rich, C. Aone, K. Knight, and J. C. Martinez), Proceedings of the Machine Translation Summit III, 1991.

## Edited Volumes

- "Applications of Weighted Automata in Natural Language Processing," (K. Knight and J. May), Handbook of Weighted Automata (M. Droste, W. Kuich, and H. Vogler, eds), 2009.

- "Text Simplification for Information Seeking Applications," (B. Beigman Klebanov, K. Knight, and D. Marcu), On the Move to Meaningful Internet Systems, (R. Meersman, Z. Tari, eds), Lecture Notes in Computer Science (3290), Springer-Verlag, 2004.

- "Unification," *The Encyclopedia of Artificial Intelligence, Second Edition*, John Wiley and Sons, 1992.

**Refereed Magazine Articles**

- "Finding the WRITE Stuff: Automatic Identification of Discourse Structure in Student Essays," (J. Burstein, D. Marcu, and K. Knight), *IEEE Intelligent Systems*, Jan/Feb, 2003.

- "Automating Knowledge Acquisition for Machine Translation," (K. Knight), *AI Magazine* 18(4), 1997.

**Refereed Workshop Papers**

- "Biomedical Event Extraction using Abstract Meaning Representation" (S. Rao, D. Marcu, K. Knight, and H. Daume), Proc. BioNLP Workshop, ACL, 2017.

- "Unsupervised Neural Hidden Markov Models" (K. Tran, Y. Bisk, A. Vaswani, D. Marcu, and K. Knight), Proc. EMNLP Workshop on Structured Prediction, 2016.

- "Obfuscating Gender in Social Media Writing" (S. Reddy and K. Knight), Proc. EMNLP Workshop on Natural Language Processing and Computational Social Science, 2016.

- "Leveraging Entity Linking and Related Language Projection to Improve Name Transliteration" (Y. Lin, X. Pan, A. Deri, H. Ji, and K. Knight), Proc. ACL 2016 Workshop on Named Entities, 2016.

- "Abstract Meaning Representation for Sembanking" (L. Banarescu, C. Bonial, S. Cai, M. Georgescu, K. Griffitt, U. Hermjakob, K. Knight, P. Koehn, M. Palmer, and N. Schneider), Proc. Linguistic Annotation Workshop (LAW VII-ID), ACL, 2013.

- "Towards Probabilistic Acceptors and Transducers for Feature Structures," (D. Quernheim and K. Knight), Proceedings of the ACL Workshop on Syntax and Structure in Statistical Translation (SSST), 2012.

- "DAGGER: A Toolkit for Automata on Directed Acyclic Graphs," (D. Quernheim and K. Knight), Proceedings of the International Workshop on Finite-State Methods and Natural Language Processing, 2012.

- "The Copiale Cipher," (K. Knight, B. Megyesi, C. Schaefer), Proceedings of the ACL Workshop on Building and Using Comparable Corpora, 2011.

- "What We Know About the Voynich Manuscript," (S. Reddy and K. Knight), Proceedings of the ACL Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities, 2011.

- "A Decoder for Probabilistic Synchronous Tree Insertion Grammars," (S. DeNeefe, K. Knight, and H. Vogler), Proceedings of the ACL Workshop on Applications of Tree Automata in Natural Language Processing, 2010.

- "A New Objective Function for Word Alignment," (T. Bodrumlu, K. Knight, and S. Ravi), Proceedings of the NAACL Workshop on Integer Linear Programming for Natural Language Processing, 2009.

- "Using Syntax to Improve Word Alignment Precision for Syntax-Based Machine Translation," (V. Fossum, K. Knight, and S. Abney), Proceedings of the ACL Workshop on Statistical MT, 2008.

- "Transonics: A Speech to Speech System for English-Persian Interactions," (S. Narayanan et al), Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding, 2003.

- "Learning a Translation Lexicon from Monolingual Corpora," (P. Koehn and K. Knight), Proceedings of the ACL Workshop on Unsupervised Lexical Acquisition, 2002.

- "Machine Transliteration of Names in Arabic Text," (Y. Al-Onaizan and K. Knight), Proceedings of ACL Workshop on Computational Approaches to Semitic Languages, 2002.

- "A Computational Approach to Deciphering Unknown Scripts," (K. Knight and K. Yamada), Proceedings of the ACL Workshop on Unsupervised Learning in Natural Language Processing, 1999.

- "Translating Names and Technical Terms in Arabic Text," (B. Stalls and K. Knight), Proceedings of the ACL Workshop on Computational Approaches to Semitic Languages, 1998.

- "The Practical Value of N-Grams in Generation," (I. Langkilde and K. Knight), Proceedings of the International Natural Language Generation Workshop, 1998.

- "Building a Large Ontology for Machine Translation," Proceedings of the ARPA Workshop on Human Language Technology, 1993.

- "Automated Rationalization," (K. Knight and Y. Gil), Proceedings of the Workshop on Multistrategy Learning, 1991.

- "Integrating Language Acquisition and Knowledge Acquisition," (K. Knight, E. Rich, D. Wroblewski), Proceedings of the First International Workshop on Automated Lexical Acquisition, International Joint Conference on Artificial Intelligence (IJCAI), 1989.

**Tutorial Workbooks**

- "A Statistical Machine Translation Tutorial Workbook," (K. Knight), 1999. http://www.isi.edu/natural-language/mt/wkbk.pdf.

- "Bayesian Inference with Tears: A Tutorial Workbook for Natural Language Processing Researchers," (K. Knight), 2009. http://www.isi.edu/natural-language/people/bayes-with-tears.pdf.