

# 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

- Senior Research Scientist (1992-) and ISI Fellow (2006-), USC/Information Sciences Institute.
- Research Professor (2011-), Research Associate Professor (2002-2010), Research Assistant Professor (1993-2002), Computer Science Department, University of Southern California.
- Chief Scientist (2001-), 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.
- **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.

## Grants Awarded (2006-2011)

- “Machine Translation of Informal Text and Speech.” 2011-2016. (with D. Marcu). DARPA-funded, subcontract to BBN, USC portion \$7m.
- “MURI: Machine Translation for Low-Density Languages.” 2010-2014. (with J. Carbonell, R. Barzilay, and J. Baldridge). Army-funded, USC portion \$2m.
- “Richer Representations for Machine Translation.” 2009-2013. (with M. Palmer, N. Xue, D. Gildea, and K. McKeown). NSF-funded, USC portion \$480k.
- “Deciphering Natural Language.” 2009-2013. (with R. Barzilay). NSF-funded, USC portion \$800k.
- “Improvements to Speech-to-Speech Translation.” 2008-09. DARPA-funded, subcontract to SRI, USC portion \$125k.
- “Decipherment Techniques Applied to Machine Translation.” 2007-08. DARPA-funded, subcontract to SRI, USC portion \$250k.
- “Improving Fundamental Accuracy of Speech-to-Speech Translation.” 2006-07. (with D. Marcu). DARPA-funded, subcontract to SRI, USC portion \$370k.
- “Human Informed Statistical Machine Translation.” 2005-2011. (with D. Marcu). DARPA-funded, subcontract to BBN, USC portion \$12m. (with D. Marcu).

## Publications (2006-2011)

### Journal Articles

- “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.

## Refereed Conference Papers

- “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. DeNeeffe 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.
- “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.