

Dan Jurafsky Publications

Books and Edited Collections:

- Jurafsky, Dan. 2014. *The Language of Food: A Linguist Reads the Menu*. Norton.
- Jurafsky, Daniel, and James H. Martin. 2009. [Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics](#). 2nd edition. Prentice-Hall.
- E. Fosler-Lussier, W. Byrne, and D. Jurafsky, eds. 2005. [Speech Communication Special Issue on Pronunciation Modeling and Lexicon Adaptation](#), 46:2, June 2005.
- Fox, Barbara A., Dan Jurafsky, and Laura A. Michaelis (Eds.). 1999. *Cognition and Function in Language*. CSLI Publications, Stanford, CA.

Journal Articles, Book Chapters, and Conference Papers (see very bottom for Corpora and Data)

- 2014** Sebastian Schuster, Stephanie Pancoast, Milind Ganjoo, Michael C. Frank, and Dan Jurafsky. 2014. [Speaker-Independent Detection of Child-Directed Speech](#). Proceedings of the IEEE Workshop on Spoken Language Technology (SLT-2014). [\[pdf\]](#) [\[bib\]](#) [\[github\]](#)

Dan Jurafsky, Victor Chahuneau, Bryan R. Routledge, and Noah A. Smith. 2014. [Narrative framing of consumer sentiment in online restaurant reviews](#). *First Monday* 19:4. [\[bib\]](#)

[Press: NPR All Things Considered, Wash. Post, SF Chronicle, etc.]

Tim Althoff, Cristian Danescu-Niculescu-Mizil, Dan Jurafsky. 2014. [How to Ask for a Favor: A Case Study on the Success of Altruistic Requests](#). AAAI ICWSM 2014 [\[pdf\]](#) [\[bib\]](#) [\[data\]](#)

[Press: Technology Review, Huffington Post, Fast Company, Gizmodo, ABC News]

Heeyoung Lee, Mihai Surdeanu, Bill MacCartney, and Dan Jurafsky. 2014. [On the Importance of Text Analysis for Stock Price Prediction](#). Proceedings of the Language Resources and Evaluation Conference (LREC), 2014. [\[pdf\]](#) [\[bib\]](#) [\[data\]](#)

Elie Bursztein, Angelique Moscicki, Celine Fabry, Steven Bethard, John C. Mitchell, and Dan Jurafsky. 2014. [Easy Does It: More Usable CAPTCHAs](#). Proceedings of ACM CHI 2014. [\[pdf\]](#) [\[bib\]](#)

[Press: Help Net Security]

Kevin Reschke, Martin Jankowiak, Mihai Surdeanu, Christopher D. Manning, and Daniel Jurafsky. 2014. [Event Extraction Using Distant Supervision](#). Proceedings of the Language Resources and Evaluation Conference (LREC), 2014. [\[pdf\]](#) [\[bib\]](#)

Rob Voigt, Robert J. Podesva and Dan Jurafsky. 2014. [Body Movement Correlates](#)

with Prosodic Indicators of Engagement. *Proceedings of Speech Prosody 2014*. [pdf] [bib]

- 2013** Valentin I. Spitkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2013. [Breaking Out of Local Optima with Count Transforms and Model Recombination: A Study in Grammar Induction](#). EMNLP 2013. [pdf, bib; **best paper award**]

Heeyoung Lee, Angel Chang, Yves Peirsman, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. 2013. [Deterministic Coreference Resolution Based on Entity-Centric, Precision-Ranked Rules](#). *Computational Linguistics* 39:4, 885-916.

Dan Jurafsky. 2013. [Why Ice Cream Sounds Fat and Crackers Sound Skinny](#). *Stanford Magazine* July/August 2013

Daniel Cer, Christopher D. Manning and Dan Jurafsky. 2013. [Positive Diversity Tuning for Machine Translation System Combination](#). Proceedings of the Eighth Workshop on Statistical Machine Translation (WMT 2013) [pdf, bib]

McFarland, Daniel A., Christopher D. Manning, Daniel Ramage, Jason Chuang, Jeffrey Heer, and Dan Jurafsky. 2013. [Differentiating Language Usage Through Topic Models](#). *Poetics* 41 (6), 607-625

McFarland, Daniel A., Dan Jurafsky, and Craig M. Rawlings. 2013. [Making the Connection: Social Bonding in Courtship Situations](#). *American Journal of Sociology* Vol. 118, No. 6, 1596-1649

Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, Christopher Potts. 2013. [No country for old members: User lifecycle and linguistic change in online communities](#). Proceedings of WWW 2013. **Best paper award**. [pdf, bib]

Adam Vogel, Christopher Potts, and Dan Jurafsky. 2013. [Implicatures and Nested Beliefs in Approximate Decentralized-POMDPs](#). In Proceedings of ACL 2013. [pdf, bib]

Kevin Reschke, Adam Vogel, and Dan Jurafsky. 2013. [Generating Recommendation Dialogs by Extracting Information from User Reviews](#). ACL 2013 [pdf, bib]

Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. [A computational approach to politeness with application to social factors](#). Proceedings of ACL, 2013. [pdf, bib]

Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. [Linguistic Models for Analyzing and Detecting Biased Language](#). 2013. Proceedings of ACL 2013. [pdf, bib]

Marta Recasens, Matthew Can, and Dan Jurafsky. 2013. [Same Referent, Different Words: Unsupervised Mining of Opaque Coreferent Mentions](#). Proceedings of NAACL 2013. [pdf, bib; data]

Vogel, Adam, Max Bodoia, Dan Jurafsky, and Christopher Potts. 2013. [Emergence of Gricean maxims from multi-agent decision theory](#). Proceedings of HLT NAACL

2013. [[pdf](#), [bib](#)]

Rob Voigt and Dan Jurafsky. 2013. [Tradition and Modernity in 20th Century Chinese Poetry](#). NAACL Second Workshop on Computational Linguistics for Literature.

Rajesh Ranganath, Dan Jurafsky, and Daniel A. McFarland. 2013. [Detecting friendly, flirtatious, awkward, and assertive speech in speed-dates](#). Computer Speech and Language. 27:1, 89-115. [[pdf](#), [bib](#)]

2012 J. J. McAuley, J. Leskovec, D. Jurafsky. 2012. [Learning attitudes and attributes from multi-aspect reviews](#). International Conference on Data Mining

Valentin I. Spitkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2012. [Bootstrapping Dependency Grammar Inducers from Incomplete Sentence Fragments via Austere Models](#). In Proceedings of the 11th International Conference on Grammatical Inference (ICGI 2012).

Heeyoung Lee, Marta Recasens, Angel Chang, Mihai Surdeanu, and Dan Jurafsky. 2012. [Joint Entity and Event Coreference Resolution across Documents](#). In Proceedings of the Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL).

Valentin I. Spitkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2012. [Three Dependency-and-Boundary Models for Grammar Induction](#). EMNLP-CoNLL 2012.

Adam Vogel and Dan Jurafsky. 2012. [He Said, She Said: Gender in the ACL Anthology](#). ACL Workshop on Rediscovering 50 Years of Discoveries.

Ashton Anderson, Dan McFarland, and Dan Jurafsky. 2012 [Towards a Computational History of the ACL: 1980-2008](#). ACL Workshop on Rediscovering 50 Years of Discoveries.

Justine Kao and Dan Jurafsky. 2012. [A Computational Analysis of Style, Affect, and Imagery in Contemporary Poetry](#). NAACL Workshop on Computational Linguistics for Literature.

Rob Voigt and Dan Jurafsky. 2012. [Towards a Literary Machine Translation: The Role of Referential Cohesion](#). NAACL Workshop on Computational Linguistics for Literature.

Valentin I. Spitkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2012. [Capitalization Cues Improve Dependency Grammar Induction](#). In NAACL HLT 2012 Workshop on Inducing Linguistic Structure (WILS 2012)

Dan Jurafsky. 2012. [The Cosmopolitan Condiment: An exploration of ketchup's Chinese origins](#) *Slate*, May 30, 2012.

Gabor Angeli, Chris Manning, Dan Jurafsky. 2012. [Parsing Time: Learning to Interpret Time Expressions](#). North American Chapter of the Association for Computational Linguistics (NAACL) 2012.

Michael Levin, Stefan Krawczyk, Steven Bethard, and Dan Jurafsky. 2012. [Citation-based bootstrapping for large-scale author disambiguation](#). Journal of the American

Society for Information Science and Technology 63:5, 1030-1047.

- 2011** Joshua Freedman and Dan Jurafsky. 2011. [Authenticity in America: Class Distinctions in Potato Chip Advertising](#). *Gastronomica* 11, 4: 46-54.

Dan Jurafsky. 2011. [Macarons, Macaroons, Macaroni: The curious history](#). *Slate*, Nov 16, 2011

Valentin I. Spitkovsky, Hiyan Alshawi, Angel X. Chang, and Daniel Jurafsky. 2011. [Unsupervised Dependency Parsing without Gold Part-of-Speech Tags](#). Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP 2011). [PDF]

Valentin I. Spitkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2011. [Lateen EM: Unsupervised Training with Multiple Objectives, Applied to Dependency Grammar Induction](#). Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP 2011). [PDF]

Nathanael Chambers and Dan Jurafsky. 2011. [Template-Based Information Extraction without the Templates](#). ACL-2011, Portland, OR. [PDF]

Valentin I. Spitkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2011. [Punctuation: Making a Point in Unsupervised Dependency Parsing](#). In Proceedings of the Fifteenth Conference on Computational Natural Language Learning (CoNLL-2011) [PDF]

Heeyoung Lee, Yves Peirsman, Angel Chang, Nathanael Chambers, Mihai Surdeanu, Dan Jurafsky. 2011. [Stanford's Multi-Pass Sieve Coreference Resolution System at the CoNLL-2011 Shared Task](#). Proceedings of CoNLL 2011.

Nikhil Johri, Daniel Ramage, Daniel A. McFarland, Dan Jurafsky. 2011. [A Study of Academic Collaborations in Computational Linguistics using a Latent Mixture of Authors Model](#). In ACL 2011 Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities.

Ramesh Nallapati, Xiaolin Shi, Dan McFarland, Jure Leskovec and Daniel Jurafsky. 2011 [LeadLag LDA: Estimating Topic Specific Leads and Lags of Information Outlets](#). Proceedings of ICWSM.

Andrey Gusev, Nathanael Chambers, Divye Raj Khilnani, Pranav Khaitan, Steven Bethard, and Dan Jurafsky. 2011. [Using query patterns to learn the duration of events](#). In International Conference on Computational Semantics, 2011. [PDF]

- 2010** Xiaolin Shi, Ramesh Nallapati, Jure Leskovec, Dan McFarland and Dan Jurafsky. 2010. [Who Leads Whom: Topical Lead-Lag analysis across corpora](#). NIPS Workshop on Computational Social Science and Wisdom of Crowds.

Adam Vogel, Karthik Raghunathan, and Dan Jurafsky. [Eye Spy: Improving Vision through Dialog](#). Dialog with Robots: Papers from the AAIL Fall Symposium (FS-10-05). [PDF]

Sasha Calhoun, Jean Carletta, Jason M. Brenier, Neil Mayo, Dan Jurafsky, Mark Steedman, and David Beaver. 2010. [The NXT-format Switchboard Corpus: a rich](#)

resource for investigating the syntax, semantics, pragmatics and prosody of dialogue. *Language Resources & Evaluation* 44:387-419. DOI 10.1007/s10579-010-9120-1. [\[PDF\]](#)

Karthik Raghunathan, Heeyoung Lee, Sudarshan Rangarajan, Nathanael Chambers, Mihai Surdeanu, Dan Jurafsky, Christopher Manning. A Multi-Pass Sieve for Coreference Resolution. Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP-2010), 2010. [\[PDF\]](#)

Steven Bethard and Dan Jurafsky. 2010. Who should I cite? Learning literature search models from citation behavior. In ACM Conference on Information and Knowledge Management. [\[PDF\]](#)

Valentin I. Spitzkovsky, Daniel Jurafsky, and Hiyan Alshawi. 2010. Profiting from Mark-Up: Hyper-Text Annotations for Guided Parsing. In Proceedings of ACL-2010, Uppsala, Sweden. [\[PDF\]](#)

Adam Vogel and Dan Jurafsky. 2010. Learning to Follow Navigational Directions. In Proceedings of ACL-2010, Uppsala, Sweden. [\[PDF\]](#)

Nathanael Chambers and Dan Jurafsky. 2010. Improving the Use of Pseudo-Words for Evaluating Selectional Preferences. ACL-2010, Uppsala, Sweden. [\[PDF\]](#)

Elie Bursztein, Steven Bethard, John C. Mitchell, Dan Jurafsky, and Celine Fabry. How good are humans at solving CAPTCHAs? A large scale evaluation. In IEEE Symposium on Security and Privacy, 2010. [\[PDF\]](#)

Daniel Cer, Michel Galley, Daniel Jurafsky and Christopher Manning. 2010. Phrasal: A Toolkit for Statistical Machine Translation with Facilities for Extraction and Incorporation of Arbitrary Model Features. Proceedings of NAACL 2010 Demo Session. [\[PDF\]](#)

Daniel Cer , Daniel Jurafsky and Christopher Manning. 2010. The Best Lexical Metric for Phrase-Based Statistical MT System Optimization. Proceedings of NAACL 2010. [\[PDF\]](#)

Sharon Goldwater, Dan Jurafsky, and Christopher D. Manning. 2010. Which words are hard to recognize? Prosodic, lexical, and disfluency factors that increase speech recognition error rates. *Speech Communication* 52, 181-200. [\[PDF\]](#)

Valentin I. Spitzkovsky, Hiyan Alshawi, Daniel Jurafsky, and Christopher D. Manning. 2010. Viterbi Training Improves Unsupervised Dependency Parsing. In *Proceedings of CoNLL-2010*. [\[PDF\]](#).

Valentin I. Spitzkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2010. From Baby Steps to Leapfrog: How “Less is More” in Unsupervised Dependency Parsing. In *Proceedings of NAACL HLT 2010*. [\[PDF\]](#).

Daniel Cer, Marie-Catherine de Marneffe, Daniel Jurafsky and Christopher D. Manning. 2010. Parsing to Stanford Dependencies: Trade-offs between speed and accuracy. In Proceedings of LREC-10, Malta. [\[PDF\]](#)

Nathanael Chambers and Dan Jurafsky. 2010. [A Database of Narrative Schemas](#).

Proceedings of LREC-2010, Malta. [\[PDF\]](#)

2009 Sebastian Pado, Daniel Cer, Michel Galley, Dan Jurafsky and Christopher D. Manning. 2009. Measuring Machine Translation Quality as Semantic Equivalence: A Metric Based on Entailment Features. *Machine Translation* 23:2-3, 181-193. [doi:10.1007/s10590-009-9060-y](https://doi.org/10.1007/s10590-009-9060-y)

Yun-Hsuan Sung and Dan Jurafsky. 2009. Hidden Conditional Random Fields for Phone Recognition. ASRU 2009. [\[PDF\]](#)

Yuan Zhao and Dan Jurafsky. 2009. The effect of lexical frequency and Lombard reflex on tone hyperarticulation. *Journal of Phonetics* 27:2, 231-247. [\[PDF\]](#)

Valentin I. Spitzkovsky, Hiyan Alshawi, and Daniel Jurafsky. 2009. Baby Steps: How "Less is More" in Unsupervised Dependency Parsing. In *NIPS 2009 Workshop on Grammar Induction, Representation of Language and Language Learning*. [\[PDF\]](#).

Eneko Agirre, Angel X. Chang, Daniel S. Jurafsky, Christopher D. Manning, Valentin I. Spitzkovsky, and Eric Yeh. 2009. Stanford-UBC at TAC-KBP. In *Proceedings of the Second Text Analysis Conference TAC 2009*. [\[PDF\]](#).

Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant supervision for relation extraction without labeled data. *Proceedings of ACL-IJCNLP 2009*. [\[PDF\]](#)

Nathanael Chambers and Dan Jurafsky. 2009. Unsupervised Learning of Narrative Schemas and their Participants. *Proceedings of ACL-IJCNLP 2009*. [\[PDF\]](#)

Rajesh Ranganath, Dan Jurafsky, and Dan McFarland. 2009. It's Not You, it's Me: Detecting Flirting and its Misperception in Speed-Dates. *Proceedings of EMNLP 2009*. [\[PDF\]](#)

Sebastian Padó, Michel Galley, Dan Jurafsky, and Chris Manning. 2009. Robust Machine Translation Evaluation with Entailment Features. *Proceedings of ACL-IJCNLP 2009*.

[Talk not paper]. Ketchup, ceviche, and chocolate chip cookies: Travels in the language of food. [\[PDF slides\]](#)

Dan Jurafsky, Rajesh Ranganath, and Dan McFarland. 2009. Extracting Social Meaning: Identifying Interactional Style in Spoken Conversation. *Proceedings of NAACL HLT 2009*. [\[PDF\]](#)

Pi-Chuan Chang, Huihsin Tseng, Dan Jurafsky, and Christopher D. Manning. 2009. [Discriminative Reordering with Chinese Grammatical Relations Features](#). *NAACL 2009 Third Workshop on Syntax and Structure in Statistical Translation*.

Sebastian Padó, Michel Galley, Dan Jurafsky, and Christopher D. Manning. 2009. [Textual Entailment Features for Machine Translation Evaluation](#). *Proceedings of the EACL 2009 Fourth Workshop on Statistical Machine Translation*.

Pi-Chuan Chang, Dan Jurafsky and Christopher D. Manning. 2009. [Disambiguating "DE" for Chinese-English Machine Translation](#). *Proceedings of the EACL 2009*

Fourth Workshop on Statistical Machine Translation.

Alan Bell, Jason Brenier, Michelle Gregory, Cynthia Girand, and Dan Jurafsky. 2009. [Predictability Effects on Durations of Content and Function Words in Conversational English](#). Journal of Memory and Language 60:1, 92-111.

- 2008** David Hall, Daniel Jurafsky, and Christopher D. Manning. 2008. [Studying the History of Ideas Using Topic Models](#). In Proceedings of EMNLP 2008, 363-371.

Rion Snow, Brendan O'Connor, Daniel Jurafsky and Andrew Y. Ng. 2008. [Cheap and Fast - But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks](#). In Proceedings of EMNLP 2008, 254-263.

Nathanael Chambers and Dan Jurafsky. 2008. [Jointly Combining Implicit Constraints Improves Temporal Ordering](#). In Proceedings of EMNLP 2008, 698-706.

Sharon Goldwater, Dan Jurafsky, and Christopher D. Manning. 2008. [Which words are hard to recognize? Lexical, prosodic, and disfluency factors that increase ASR error rates](#). In ACL/HLT; 380-388.

Nathanael Chambers and Dan Jurafsky. 2008. [Unsupervised Learning of Narrative Event Chains](#). In Proceedings of ACL/HLT 2008.

Daniel Cer, Daniel Jurafsky, and Christopher D. Manning. 2008. Regularization and Search for Minimum Error Rate Training. Proceedings of the Third Workshop on Statistical Machine Translation. [\[PDF\]](#)

Yun-Hsuan Sung, Constantinos Boulis, and Dan Jurafsky. 2008. [Maximum Conditional Likelihood Linear Regression and Maximum A Posteriori for Hidden Conditional Random Fields Speaker Adaptation](#). IEEE ICASSP 2008, 4293-4296.

Vivek Kumar Rangarajan Sridhar, Ani Nenkova, Shrikanth Narayanan and Dan Jurafsky. 2008. [Detecting prominence in conversational speech: pitch accent, givenness and focus](#). In Proceedings of Speech Prosody, Campinas, Brazil, 453-456.

- 2007** Yun-Hsuan Sung, Constantinos Boulis, Christopher Manning and Dan Jurafsky. 2007. [Regularization, Adaptation, and Non-Independent Features Improve Hidden Conditional Random Fields for Phone Classification](#). In IEEE ASRU 2007. 347-352.

Rion Snow, Sushant Prakash, Daniel Jurafsky, and Andrew Y. Ng. [Learning to merge word senses](#). In Proceedings of EMNLP 2007

Surabhi Gupta, John Niekrasz, Matthew Purver and Dan Jurafsky. 2007. [Resolving "You" in Multi-Party Dialog](#). In Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, Belgium, September 2007.

Volker Strom, Ani Nenkova, Robert Clark, Yolanda Vazquez-Alvarez, Jason Brenier, Simon King, and Dan Jurafsky. 2007. [Modelling Prominence and Emphasis Improves Unit-Selection Synthesis](#). Interspeech 2007.

Nathanael Chambers, Shan Wang and Dan Jurafsky. 2007. [Classifying Temporal Relations Between Events](#). Proceedings of ACL 2007 short papers, Prague, Czech Republic.

Surabhi Gupta, Matthew Purver and Dan Jurafsky. 2007. [Disambiguating Between Generic and Referential "You" in Dialog](#). Proceedings of ACL 2007 short papers, Prague, Czech Republic.

Surabhi Gupta, Ani Nenkova and Dan Jurafsky. 2007. [Measuring Importance and Query Relevance in Topic-focused Multi-document Summarization](#). Proceedings of ACL 2007 short papers, Prague, Czech Republic.

Ani Nenkova, Jason Brenier, Anubha Kothari, Sasha Calhoun, Laura Whitton, David Beaver, and Dan Jurafsky. 2007. [To Memorize or to Predict: Prominence Labeling in Conversational Speech](#). NAACL-HLT 2007.

Yuan Zhao and Dan Jurafsky. 2007. [The Effect of Lexical Frequency on Tone Production](#). Proceedings of ICPHS 2007, 477-480.

2006 Rion Snow, Dan Jurafsky, and Andrew Y. Ng. 2006. [Semantic taxonomy induction from heterogenous evidence](#). Proceedings of COLING/ACL 2006, Sydney. **ACL Best Paper Award**.

Constance Clarke and Dan Jurafsky. 2006. [Limitations of MLLR Adaptation with Spanish-Accented English: An Error Analysis](#). Proceedings of INTERSPEECH-2006, Pittsburgh, PA.

Filip Krsmanovic, Curtis Spencer, Daniel Jurafsky, and Andrew Y. Ng. 2006 [Have we met? MDP Based Speaker ID for Robot Dialogue](#). Proceedings of INTERSPEECH-2006, Pittsburgh, PA.

Cheng-Tao Chu, Yun-Hsuan Sung, Yuan Zhao, Dan Jurafsky. 2006. [Detection of Word Fragments in Mandarin Telephone Conversation](#). Proceedings of INTERSPEECH-2006, Pittsburgh, PA.

Jason Brenier, Ani Nenkova, Anubha Kothari, Laura Whitton, David Beaver, Dan Jurafsky. 2006. [The \(Non\)Utility of Linguistic Features for Predicting Prominence in Spontaneous Speech](#). IEEE/ACL 2006 Workshop on Spoken Language Technology, Aruba.

2005 Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. 2005. [Learning syntactic patterns for automatic hypernym discovery](#) . Proceedings of NIPS 17.

Steven Bethard, Hong Yu, Ashley Thornton, Vasileios Hatzivassiloglou, and Dan Jurafsky. [Extracting opinion propositions and opinion holders using syntactic and lexical cues](#). In James G. Shanahan, Yan Qu, and Janyce Wiebe, editors, Computing Attitude and Affect in Text: Theory and Applications. Springer, 2005

Sameer Pradhan, Kadri Hacioglu, Valerie Krugler, Wayne Ward , James H. Martin and Daniel Jurafsky. 2005. [Support Vector Learning for Semantic Argument Classification](#). *Machine Learning* 60:1-3 , 11-39

Yuan, Jiahong and Dan Jurafsky. 2005. [Detection of Questions in Chinese Conversation](#) . In *Proceedings of IEEE ASRU 2005*.

Yuan, Jiahong, Jason M. Brenier, and Dan Jurafsky. 2005. [Pitch Accent Prediction: Effects of Genre and Speaker](#). In *Proceedings of EUROSPEECH-05*.

Huihsin Tseng, Pichuan Chang, Galen Andrew, Daniel Jurafsky, Christopher Manning. 2005. [A Conditional Random Field Word Segmenter](#). Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing.

Huihsin Tseng, Daniel Jurafsky, Christopher Manning. 2005. [Morphological features help POS tagging of unknown words across language varieties](#). Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing.

Jason M. Brenier, Daniel Cer and Daniel Jurafsky. 2005. [The Detection of Emphatic Words Using Acoustic and Lexical Features](#). In *Proceedings of EUROSPEECH-05*.

Yanli Zheng, Richard Sproat, Liang Gu, Izhak Shafran, Haolang Zhou, Yi Su, Dan Jurafsky, Rebecca Starr and Su-Youn Yoon. 2005. [Accent Detection and Speech Recognition for Shanghai-Accented Mandarin](#). In *Proceedings of EUROSPEECH-05*.

Yuan Zhao and Dan Jurafsky. 2005. [A preliminary study of Mandarin filled pauses](#). Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop.

Pradhan, Sameer, Wayne Ward, Kadri Hacioglu, Jim Martin, Dan Jurafsky. 2005. [Semantic Role Labeling Using Different Syntactic Views](#). Proceedings of ACL-2005, Ann Arbor, MI.

Chen, Ying, Hongling Sun, and Dan Jurafsky. 2005. [A corrigendum to Sun and Jurafsky \(2004\) "Shallow Semantic Parsing of Chinese"](#). University of Colorado at Boulder CSLR Tech Report TR-CSLR-2005-01.

2004 Susanne Gahl, Daniel Jurafsky, and Douglas Roland. 2004. [Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons](#). *Behavior Research Methods, Instruments, & Computers*, 36, 432-443.

Diab, Mona, Kadri Hacioglu, and Daniel Jurafsky. 2004. [Automatic Tagging of Arabic Text: From Raw Text to Base Phrase Chunks](#). In Proceedings of NAACL-HLT 2004.

Pradhan, Sameer, Honglin Sun, Wayne Ward, James H. Martin, and Daniel Jurafsky. 2004. [Parsing Arguments of Nominalizations in English and Chinese](#). In Proceedings of NAACL-HLT 2004.

Pradhan, Sameer, Wayne Ward, Kadri Hacioglu, James H. Martin, and Daniel Jurafsky. 2004. [Shallow Semantic Parsing Using Support Vector Machines](#). In Proceedings of NAACL-HLT 2004.

Steven Bethard, Hong Yu, Ashley Thornton, Vasileios Hativassiloglou, and Dan Jurafsky. 2004. [Automatic Extraction of Opinion Propositions and their Holders](#). In Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text.

Honglin Sun and Daniel Jurafsky. 2004. [Shallow Semantic Parsing of Chinese](#). In Proceedings of NAACL-HLT 2004. [see corrections in [Chen, Sun, and Jurafsky 2005](#) above]

2003 Alan Bell, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory,

and Daniel Gildea. 2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* **113** (2), 1001-1024. ([PDF](#))

Jurafsky, Daniel. 2003. Pragmatics and Computational Linguistics. In Laurence R. Horn & Gregory Ward (eds.) *Handbook of Pragmatics*. Oxford: Blackwell. ([Postscript](#)) or ([PDF](#)) [this draft is a pre-final version, there are various changes in references and copy-editing in published version]

Jurafsky, Dan. 2003. Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production. In Rens Bod, Jennifer Hay, and Stefanie Jannedy, (Eds.), *Probabilistic Linguistics* ([Postscript](#)) or ([PDF](#)).

Pradhan, Sameer, Kadri Hacioglu, Wayne Ward, James H. Martin, and Daniel Jurafsky. 2003. Semantic Role Parsing: Adding Semantic Structure to Unstructured Text. In Proceedings of the International Conference on Data Mining (ICDM-2003).

Sun, Honglin and Dan Jurafsky. 2003. [The Effect of Rhythm on Structural Disambiguation in Chinese](#). Proceedings of the Annual Conference of the Association of Computational Linguistics Special Interest Group on Chinese (SIGHAN-03). ([PDF](#))

Ikeno, Ayako, Bryan Pellom, Dan Cer, Ashley Thornton, Jason M. Brenier, Dan Jurafsky, Wayne Ward, and William Byrne. 2003. Issues in Recognition of Spanish-Accented Spontaneous English. In *Proceedings of IEEE/ISCA Workshop on Spontaneous Speech Processing and Recognition*, Tokyo, Japan. ([PDF](#)).

Gahl, Susanne, Lise Menn, Gail Ramsberger, Daniel Jurafsky, Elizabeth Elder, Molly Rewega and Audrey Holland. 2003. Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. *Brain and Cognition* 53:2, 223-228. ([PDF](#)).

2002 Daniel Gildea and Daniel Jurafsky. 2002. [Automatic Labeling of Semantic Roles](#) *Computational Linguistics* 28:3, 245-288.

Jurafsky, Daniel, Alan Bell, and Cynthia Girand. 2002. The Role of the Lemma in Form Variation. In Gussenhoven, Carlos and Natasha Warner (eds.), *Papers in Laboratory Phonology VII*. Berlin/New York: Mouton de Gruyter, 1-34. ([Postscript](#)) or ([PDF](#)).

Roland, Douglas and Daniel Jurafsky. (2002). Verb sense and verb subcategorization probabilities. In Stevenson, Suzanne, and Paola Merlo (eds.), *The Lexical Basis of Sentence Processing: Formal, Computational, and Experimental Issues*. Amsterdam: John Benjamins. 325-346. ([Postscript](#)) or ([PDF](#)).

Gildea, Daniel and Daniel Jurafsky (2002). Identifying Semantic Relations in Text. In Gerhard Lakemeyer and Bernhard Nebel (eds.) *Exploring AI in the New Millenium*. Morgan Kaufmann.

Narayanan, Sridhar and Daniel Jurafsky. 2002. A Bayesian Model Predicts Human Parse Preference and Reading Time in Sentence Processing. In T. G. Dietterich, S. Becker and Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems 14*. Cambridge, MA: MIT Press. 59-65. ([Postscript](#)) ([PDF](#))

Sameer S Pradhan, Gabriel Illouz, Sasha J Blair-Goldensohn, Andrew Hazen Schlaikjer, Valerie Krugler, Elena Filatova, Pablo A Duboue, Hong Yu, Rebecca J Passonneau, Steven Bethard, Vasileios Hatzivassiloglou, Wayne Ward, Dan Jurafsky, Kathleen R McKeown, and James H Martin. 2002. [Building a foundation system for producing short answers to factual questions](#). Eleventh Text Retrieval Conference (TREC-11). Washington, DC.

Wayne Ward, Holly Krech, Xiuyang Yu, Keith Herold, George Figgs, Ayako Ikeno, Dan Jurafsky, and William Byrne. 2002. [Lexicon adaptation for LVCSR: speaker idiosyncracies, non-native speakers, and pronunciation choice](#). In ISCA ITR Workshop on Pronunciation Modeling and Lexicon Adaptation, 2002.

- 2001** Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In Bybee, Joan and Paul Hopper (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 229-254. ([Postscript](#)) or ([PDF](#)).

Schone, Patrick and Daniel Jurafsky. 2001. Is Knowledge-Free Induction of Multiword Unit Dictionary Headwords a Solved Problem? Proceedings of *Empirical Methods in Natural Language Processing*, Pittsburgh, PA. ([PDF](#))

Schone, Patrick and Daniel Jurafsky. 2001. Language-Independent Induction of Part of Speech Class Labels Using Only Language Universals. In IJCAI-2001 Workshop "Text Learning: Beyond Supervision". ([Postscript](#)) ([PDF](#))

Schone, Patrick and Daniel Jurafsky. 2001. Knowledge-Free Induction of Inflectional Morphologies. Proceedings of the North American chapter of the Association for Computational Linguistics (NAACL-2001). ([PDF](#))

Jurafsky, Dan, Wayne Ward, Zhang Jianping, Keith Herold, Yu Xiuyang, and Zhang Sen. 2001. What Kind of Pronunciation Variation is Hard for Triphones to Model? *Proceedings of ICASSP-01*, 1.577-580, Salt Lake City, Utah. ([Postscript](#)) ([PDF](#))

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. The Effect of Language Model Probability on Pronunciation Reduction. In *Proceedings of ICASSP-01* II.801--804, Salt Lake City, Utah. ([Postscript](#)) ([PDF](#))

- 2000** Stolcke, Andreas, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Marie Meteer, and Carol Van Ess-Dykema. 2000. Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech. *Computational Linguistics* 26:3, 339-371 ([Postscript](#)) or ([PDF](#)).

Schone, Patrick and Daniel Jurafsky. 2000. Knowledge-Free Induction of Morphology using Latent Semantic Analysis. Proceedings of the Conference on Computational Natural Language Learning (CoNLL-2000). ([PDF](#))

Gildea, Daniel and Daniel Jurafsky. 2000. Automatic Labeling of Semantic Roles. In Proceedings of ACL 2000, Hong Kong. ([Postscript](#)) ([PDF](#))

Roland, Douglas, Daniel Jurafsky, Lise Menn, Susanne Gahl, Elizabeth Elder and Chris Riddoch. 2000. Verb Subcategorization Frequency Differences between Business-News and Balanced Corpora: the role of verb sense. In *Proceedings of*

the Association for Computational Linguistics (ACL-2000) Workshop on Comparing Corpora. ([Postscript](#)) or ([PDF](#))

- 1999** Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. Chicago Linguistics Society (CLS-99), 151-166. ([Postscript](#)) ([MS Word](#))

Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, and Daniel Gildea. 1999. Forms of English function words - Effects of disfluencies, turn position, age and sex, and predictability. Proceedings of ICPHS-99. [\[PDF\]](#)

Mark Core, Masato Ishizaki, Johanna Moore, Christine Nakatani, Nobert Reithinger, David Traum, Syun Tutiya. 1999. The Report of The Third Workshop of the Discourse Resource Initiative. Chiba University. [\[PDF\]](#)

- 1998** Shriberg, Elizabeth, Rebecca Bates, Paul Taylor, Andreas Stolcke, Daniel Jurafsky, Klaus Ries, Noah Coccaro, Rachel Martin, Marie Meteer, and Carol Van Ess-Dykema. 1998. Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech? *Language and Speech* 41:3-4, 439-487 [\[PDF\]](#)

Jurafsky, Daniel, Alan Bell, Eric Fosler-Lussier, Cynthia Girand, and William Raymond. 1998. [Reduction of English function words in Switchboard](#). Proceedings of ICSLP-98, Volume 7, 3111-3114.

Coccaro, Noah and Daniel Jurafsky. 1998. [Towards Better Integration of Semantic Predictors in Statistical Language Modeling](#) Proceedings of ICSLP-98, Volume 6, 2403-2406.

Stolcke, Andreas, Elizabeth Shriberg, Rebecca Bates, Noah Coccaro, Daniel Jurafsky, Rachel Martin, Marie Meteer, Klaus Ries, Paul Taylor, and Carol Van Ess-Dykema. (1998 draft). [Dialog Act Modeling for Conversational Speech](#). Papers from the AAAI Spring Symposium on Applying Machine Learning to Discourse Processing, TR SS-98-01, 98-105. AAAI Press.

Roland, Douglas, and Daniel Jurafsky. 1998. How Verb Subcategorization Frequencies Are Affected By Corpus Choice. In proceedings of COLING/ACL-98, 1122-1128. [\[PDF\]](#)

Jurafsky, Daniel, Elizabeth Shriberg, Barbara Fox, and Traci Curl. 1998. Lexical, Prosodic, and Syntactic Cues for Dialog Acts. Proceedings of ACL/COLING-98 Workshop on Discourse Relations and Discourse Markers, 114-120. ([Postscript](#)) or ([PDF](#)).

Fillmore, Charles, Nancy Ide, Daniel Jurafsky, and Catherine Macleod. 1998. An American National Corpus: A Proposal. In Proceedings of the First International Conference on Language Resources and Evaluation, Granada.

Narayanan, Srini, and Daniel Jurafsky. 1998. [Bayesian Models of Human Sentence Processing](#). In Proceedings of CogSci-98.

Jurafsky, Daniel, Rebecca Bates, Noah Coccaro, Rachel Martin, Marie Meteer, Klaus Ries, Elizabeth Shriberg, Andreas Stolcke, Paul Taylor, and Carol Van Ess-Dykema. 1998. Switchboard Discourse Language Modeling Project Report

Research Note 30, Center for Speech and Language Processing, Johns Hopkins University, Baltimore, MD. ([Postscript](#)) or ([PDF](#)).

Older Jurafsky, Daniel, Rebecca Bates, Noah Coccaro, Rachel Martin, Marie Meteer, Klaus Ries, Elizabeth Shriberg, Andreas Stolcke, Paul Taylor, and Carol Van Ess-Dykema. 1997. Automatic Detection of Discourse Structure for Speech Recognition and Understanding. In the Proceedings of the 1997 IEEE Workshop on Speech Recognition and Understanding, Santa Barbara. ([Postscript](#)) or ([PDF](#)).

Jurafsky, Daniel, Elizabeth Shriberg, and Debra Biasca. 1997a. Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13, University of Colorado, Boulder. Institute of Cognitive Science Technical Report 97-02. [Postscript](#) or [HTML](#)

Jurafsky, Daniel, Elizabeth Shriberg, and Debra Biasca. 1997b. Switchboard Dialog Act Corpus. Corpus can be downloaded here as [swb1_dialogact_annot.tar.gz](#)

Jurafsky, Daniel. (1996a). A Probabilistic Model of Lexical and Syntactic Access and Disambiguation. *Cognitive Science* **20** 137-194 [[PDF](#)]

Jurafsky, Daniel (1996b). Universal Tendencies in the Semantics of the Diminutive. *Language* **72** 533-578. [[PDF](#)]

Gildea, Daniel and Daniel Jurafsky. (1996). Learning Bias and Phonological Rule Induction. *Computational Linguistics* 22, 497-530. ([PDF](#))

Gildea, Daniel and Daniel Jurafsky. 1995. [Automatic Induction of Finite State Transducers for Simple Phonological Rules](#). In *Proceedings of ACL 95*, Cambridge, MA, pp 9-15

Tajchman, Gary and Dan Jurafsky and Eric Fosler. 1995. [Learning Phonological Rule Probabilities from Speech Corpora with Exploratory Computational Phonology](#). In *Proceedings of ACL 95*, Cambridge, MA, pp 9-15

Tajchman, Gary and Eric Fosler and Dan Jurafsky. 1995. [Building Multiple Pronunciation Models for Novel Words using Exploratory Computational Phonology](#). In *Proceedings of EUROSPEECH-95*

Jean-Pierre Koenig and Daniel Jurafsky. 1995. [Type Underspecification and On-line Type Construction in the Lexicon](#). In *Proceedings of WCCFL-94*

Jurafsky, Daniel, Chuck Wooters, Gary Tajchman, Jonathan Segal, Andreas Stolcke, Eric Fosler, and Nelson Morgan. 1995. Using a Stochastic Context-Free Grammar as a Language Model for Speech Recognition. In *Proceedings of ICASSP-95*, Detroit, MI, pp 189-192 ([PDF](#))

Jurafsky, Daniel, Chuck Wooters, Gary Tajchman, Jonathan Segal, Andreas Stolcke, Eric Fosler, and Nelson Morgan. 1994. The Berkeley Restaurant Project. In *Proceedings of ICSLP-94* Yokohama, Japan. pp 2139-2142. ([PDF](#))

Jurafsky, Daniel, Chuck Wooters, Gary Tajchman, Jonathan Segal, Andreas Stolcke, Eric Fosler, and Nelson Morgan. 1994. [Integrating Experimental Models of Syntax, Phonology, and Accent/Dialect in a Speech Recognizer](#) (in AAAI-94

workshop)

I also developed some (very old) bibtex macros for the LSA (Linguistic Society of America) style, in 1994. Here are the files: [lsalike.bst](#) and [lsalike.sty](#).

CORPORA Data related to the [Switchboard Dialog Act Tagged Corpus](#):

- Jurafsky, Daniel, Elizabeth Shriberg, and Debra Biasca. 1997a. Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13, University of Colorado, Boulder. Institute of Cognitive Science Technical Report 97-02. [Postscript](#) or [HTML](#)
- Jurafsky, Daniel, Elizabeth Shriberg, and Debra Biasca. 1997b. Switchboard Dialog Act Corpus. Corpus can be downloaded here as [swb1_dialogact_annot.tar.gz](#)
- The [training splits file ws97-train-convs.list](#) used in Stolcke et al. (2000).
- The [test splits file ws97-test-convs.list](#) used in Stolcke et al. (2000).