

## JOEL TETREAULT

---

CONTACT INFORMATION	400 Beale Street Apt. 2035 San Francisco, CA, 94105 USA	Phone: 310.612.0315  tetreaul@gmail.com <a href="http://www.cs.rochester.edu/u/tetreaul">http://www.cs.rochester.edu/u/tetreaul</a>
EDUCATION	<b>University of Rochester</b> PhD, Computer Science <i>Advisor:</i> James Allen <i>Thesis:</i> Empirical Evaluations of Pronoun Resolution	September 1998 – December 2004 Rochester, NY
	<b>University of Rochester</b> MA, Computer Science <i>Advisor:</i> James Allen	September 1998 - June 1999 Rochester, NY
	<b>Harvard University</b> BA, Computer Science <i>Advisor:</i> Barbara Grosz <i>Thesis:</i> A Centering Approach to Plural Pronouns	September 1994 - June 1998 Cambridge, MA
RESEARCH INTERESTS	Natural Language Processing, Discourse Processing, Automated Essay Scoring, Grammatical Error Detection, Plagiarism Detection, Text Quality Assessment, Spoken Dialog (Tutoring) Systems, Reference Resolution, Crowdsourcing	
WORK EXPERIENCE	<b>Senior Principal Manager, Core NL R&amp;D</b> Nuance Communications, Inc.	November 2012 – present Sunnyvale, CA, USA
	Lead NLP group of Nuance long-term research team on conversational user interfaces. Duties include developing short and long term strategic direction of NLP group, management and hiring of scientists and interns, collaborate with other research groups in Nuance. Research focused on developing high-quality, flexible and generalizable NLP tools for the development of an intelligent dialog system. Specific research areas for the group include syntactic and semantic parsing of noisy data, morphological analysis, natural language dialog and generation.	
	<b>Managing Senior Research Scientist</b> Educational Testing Service	February 2007 – October 2012 Princeton, NJ, USA
	Lead Automated Scoring NLP group. Duties include developing long-term strategic direction of NLP group, management of 5 scientists, mentoring of summer and year-long interns, recruitment of scientists, engineers, and research assistants, associate editor for company-wide internal review process, NLP project leader for two worldwide and national level assessments (managing 8 people per project), and representing ETS to external clients. Research projects include developing NLP grammatical error detection systems, plagiarism and paraphrase detection, identification of discourse coherence in student essays, and effective use of crowdsourcing for improving educational applications. Started as an Associate Research Scientist in 2007 and was promoted to current position in 2010.	
	<b>Postdoctoral Research Fellow</b> <i>Mentor:</i> Diane Litman	November 2004 – February 2007 Learning Research and Development Center, University of Pittsburgh, PA, USA
	Developed Reinforcement Learning techniques to improve the ITSPOKE spoken dialogue tutoring system. Investigated which features should be used to best represent the student state in our tutoring domain. Assisted with corpora development and data collection.	

## Dissertation Research

Mentor: James Allen

1999 – 2004

University of Rochester,  
Rochester, NY, USA

Researched factors (semantic information, discourse and dialogue structure) that contribute to proper interpretation of pronouns and implicit roles in a natural language system. Built a reference resolution system that can compare different pronoun resolution methods. Also developed a refined reference model that works well computationally and is psycholinguistically plausible.

## Research Intern

Mentor: David Traum

Summer 2002

Institute for Creative Technologies,  
Los Angeles, CA, USA

Worked on building a reference resolution module in the Mission Rehearsal Exercise (MRE) natural language system. Work also included investigating best practices and tools for transcribing and annotating dialogs as well as training annotators.

## SELECTED PUBLICATIONS

### Books

Claudia Leacock, Martin Chodorow, Michael Gamon and Joel Tetreault. *Automated Grammatical Error Detection for Language Learners*. Synthesis Lectures on Human Language Technologies, 2010, Vol 3., No. 1.

### Journal Articles and Book Chapters

Joel Tetreault, Martin Chodorow and Nitin Madnani. Bucking the trend: improved evaluation and annotation practices for ESL error detection systems, *Language Resources and Evaluation*. Springer Netherlands: July, 2013.

Jill Burstein, Joel Tetreault and Nitin Madnani. The E-rater Automated Essay Scoring System. In M. D. Shermis and J. Burstein (Eds.), *Handbook of Automated Essay Scoring: Current Applications and Future Directions*. New York: Routledge, 2013.

Jill Burstein, Joel Tetreault, Martin Chodorow, Daniel Blanchard and Slava Andreyev. Automated Evaluation of Discourse Coherence Quality in Essay Writing. In M. D. Shermis and J. Burstein (Eds.), *Handbook of Automated Essay Scoring: Current Applications and Future Directions*. New York: Routledge, 2013.

Jill Burstein, Joel Tetreault and Martin Chodorow. Holistic Annotation of Discourse Coherence Quality in Noisy Essay Writing. In the *Special issue of Dialogue and Discourse on: Beyond semantics: the challenges of annotating pragmatic and discourse phenomena*. (Eds. S. Dipper, H. Zinsmeister, and B. Webber), 2013.

Martin Chodorow, Michael Gamon and Joel Tetreault. The Utility of Article and Preposition Error Correction Systems for English Language Learners: Feedback and Assessment. *Language Testing*, 2009, Vol. 27, No. 3, pp. 419-436.

Yoko Futagi, Paul Deane, Martin Chodorow and Joel Tetreault. A computational approach to detecting collocation errors in the writing of non-native speakers of English. *Computer Assisted Language Learning Journal*, 2008.

Joel Tetreault and Diane Litman. A Reinforcement Learning Approach to Evaluating State Representations in Spoken Dialogue Systems. *Special Issue of the Journal of Speech Communication*, Volume 50, Issue 8-0, August 2008.

Nathanael Chambers, Joel Tetreault, and James Allen. Approaches for Automatically Tagging Affect. *Computing Attitude and Affect in Text: Theory and Applications*. (Information Retrieval Series). Springer, December 2005.

Joel Tetreault. Decomposing Discourse. *Discourse Anaphora and Anaphora Resolution Colloquium*, 2004.

Joel Tetreault. A Corpus-Based Evaluation of Centering and Pronoun Resolution. *Computational Linguistics*, Volume 27-4, December 2001.

### Conference Papers

Mohammad Sadegh Rasooli and Joel Tetreault. (EMNLP 2013): Joint Parsing and Disfluency Detection in Linear Time. Seattle, WA, USA. October 18-21, 2013.

Hwee Tou Ng, Siew Mei Wu, Yuanbin Wu, Christian Hadiwinoto and Joel Tetreault. (CoNLL 2013): The CoNLL-2013 Shared Task on Grammatical Error Correction. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*. Sofia, Bulgaria. August 08-09, 2013.

Aoife Cahill, Nitin Madnani, Joel Tetreault and Diane Napolitano. (NAACL 2013): Robust Systems for Preposition Error Correction Using Wikipedia Revisions. In *Proceedings of the 2013 Meeting of the North American Association for Computational Linguistics: Human Language Technologies*. Atlanta, GA, USA, June 10-12, 2013.

Joel Tetreault, Daniel Blanchard, Aoife Cahill and Martin Chodorow. (COLING 2012): Native Tongues, Lost and Found: Resources and Empirical Evaluations in Native Language Identification. In *Proceedings of the 24th International Conference on Computational Linguistics*. Mumbai, India, December 08-15, 2012.

Martin Chodorow, Markus Dickinson, Ross Israel and Joel Tetreault. (COLING 2012): Problems in Evaluating Grammatical Error Detection Systems. In *Proceedings of the 24th International Conference on Computational Linguistics*. Mumbai, India, December 08-15, 2012.

Nitin Madnani, Joel Tetreault and Martin Chodorow. (NAACL 2012): Re-examining Machine Translation Metrics for Paraphrase Identification. In *Proceedings of the 2012 Meeting of the North American Association for Computational Linguistics: Human Language Technologies*. Montreal, Canada, June 03-08, 2012.

Ross Israel, Joel Tetreault and Martin Chodorow. (NAACL 2012): Correcting Comma Errors in Learner Essays, and Restoring Commas in Newswire Text. In *Proceedings of the 2012 Meeting of the North American Association for Computational Linguistics: Human Language Technologies*. Montreal, Canada, June 03-08, 2012.

Nitin Madnani, Michael Heilman, Joel Tetreault and Martin Chodorow. (NAACL 2012): Identifying High-Level Organizational Elements in Argumentative Discourse. In *Proceedings of the 2012 Meeting of the North American Association for Computational Linguistics: Human Language Technologies*. Montreal, Canada, June 03-08, 2012.

Beata Beigman Klebanov, Jill Burstein, Nitin Madnani, Adam Faulkner and Joel Tetreault. (CICLing 2012): Building Subjectivity Lexicon(s) From Scratch For Essay Data. In *Proceedings of the 13th International Conference on Intelligent Text Processing and Computational Linguistics*. New Delhi, India, March 11-17, 2012.

Wei Xu, Joel Tetreault, Martin Chodorow, Ralph Grishman and Le Zhao. (EMNLP 2011): Exploiting Syntactic and Distributional Information for Spelling Correction with Web-Scale N-gram Models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Edinburgh, Scotland, July 27-29, 2011.

Nitin Madnani, Joel Tetreault, Martin Chodorow and Alla Rozovskaya. (ACL 2011): They Can Help: Using Crowdsourcing to Improve the Evaluation of Grammatical Error Detection Systems. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, OR., June 19-24, 2011.

Joel Tetreault, Jennifer Foster and Martin Chodorow. (ACL 2010): Using Parse Features for Preposition Selection and Error Detection. In *Proceedings of the 48th Annual Meeting*

*of the Association for Computational Linguistics*. Uppsala, Sweden, July 11-16, 2010.

Jill Burstein, Joel Tetreault and Slava Andreyev. (NAACL 2010): Using Entity-Based Features to Model Coherence in Student Essays. In *Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Los Angeles, CA, USA, June 01-06, 2010.

Na-Rae Han, Joel Tetreault, Soo-Hwa Lee and Jin-Young Ha. (LREC 2010): Using an Error-Annotated Learner Corpus to Develop an ESL/EFL Error Correction System. In *Proceedings of the 7th International Conference on Language Resources and Evaluation*. Malta, May 19-21, 2010.

Joel Tetreault and Martin Chodorow. (COLING 2008): The Ups and Downs of Preposition Error Detection. In *Proceedings of the 22nd International Conference on Computational Linguistics*. Manchester, UK, August 18-22, 2008.

Kate Forbes-Riley, Diane Litman, Amruta Purandare, Mihai Rotaru, and Joel Tetreault. (AIED 2007): Comparing Linguistic Features for Modeling Learning in Computer Dialogue Tutoring. *Proceedings 13th International Conference on Artificial Intelligence in Education*. Los Angeles, CA, July.

Joel Tetreault, Dan Bohus, and Diane Litman. (NAACL 2007): Estimating the Reliability of MDP Policies: A Confidence Interval Approach. In *Proceedings of the 2007 Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Rochester, NY, April 22-27, 2007.

Hua Ai, Joel Tetreault, and Diane Litman. (NAACL 2007): Comparing User Simulation Models for Dialog Strategy Learning. In *Proceedings of the 2007 Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Rochester, NY, April 22-27, 2007.

Kate Forbes-Riley, Mihai Rotaru, Diane Litman and Joel Tetreault. (NAACL 2007): Exploring Affect-Context Dependencies for Adaptive System Development. In *Proceedings of the 2007 Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Rochester, NY, April 22-27, 2007. (Late-breaking News Award).

Kate Forbes-Riley, Diane Litman, Scott Silliman and Joel Tetreault. (FLAIRS 2006): Comparing Synthesized versus Pre-Recorded Tutor Speech in an Intelligent Tutoring Spoken Dialogue System. *Proceedings 19th International Conference of the Florida Artificial Intelligence Research Society*. Melbourne Beach, Florida, May 11-13, 2006.

Joel Tetreault and Diane Litman. (EACL 2006): Using Reinforcement Learning to Build a Better Model of Dialogue State. *Proceedings of the 11th Conference of the European Association for Computational Linguistics*. Trento, Italy, April 03-07, 2006.

Joel Tetreault and Diane Litman. (NAACL 2006): Comparing the Utility of State Features in Spoken Dialogue Using Reinforcement Learning. In *Proceedings of the 2006 Human Language Technology/North American Association for Computational Linguistics*. New York City, NY., June 05-07, 2006.

Hua Ai, Diane Litman, Kate Forbes-Riley, Mihai Rotaru, Joel Tetreault, and Amruta Purandare. (INTERSPEECH 2006): Using System and User Performance Features to Improve Emotion Detection in Spoken Tutoring Dialogs. In *Proceedings of the 9th International Conference on Spoken Language Processing*. Pittsburgh, PA., September 17-26, 2006.

Nathanael Chambers, Joel Tetreault, and James Allen. Approaches for Automatically Tagging Affect. In *Proceedings of the 2004 AAAI Spring Symposium of Exploring Affect and Attitude in Text*, Stanford, CA., March 25-27, 2004.

Mary Swift, Myrosia Dzikovska, Joel Tetreault and James F. Allen. (LREC 2004): Semi-automatic syntactic and semantic corpus annotation with a deep parser. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*. Lisbon, Portugal, May 26-28, 2004.

Joel Tetreault and James Allen. (CATALOG 2004): Semantics, Dialogue, and Pronoun Resolution. In *Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue*. Barcelona, Spain, July 19-21, 2004.

Joel Tetreault and James Allen. An Empirical Evaluation of Pronoun Resolution and Clausal Structure. In *Proceedings of the 2003 International Symposium on Reference Resolution and its Applications to Question Answering and Summarization*. Venice, Italy, June 23-24, 2003.

Joel Tetreault. Implicit Role Reference. In *Proceedings of the International Symposium on Reference Resolution for Natural Language Processing*. Alicante, Spain, June 03-04, 2002.

Joel Tetreault. (DAARC 2002): Clausal Structure and Pronoun Resolution. In *Proceedings of the 4th Discourse Anaphora and Anaphora Resolution Colloquium*. Lisbon, Portugal, September 18-20, 2002.

Joel Tetreault. (ACL 1999): Analysis of Syntax Based Pronoun Resolution Methods. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (student session)*. College Park, MD., June 20-26, 1999.

Donna Byron and Joel Tetreault. (EACL 1999): A Flexible Architecture for Reference Resolution. In *Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics*. Bergen, Norway, June 08-12, 1999.

### Select Workshop Papers

Joel Tetreault, Daniel Blanchard and Aoife Cahill. (BEA 2013): A Report on the First Native Language Identification Shared Task. In *Proceedings of the 8th Workshop on Innovative Use of NLP for Building Educational Applications*. Atlanta, GA, USA, June 13, 2013.

Michael Heilman, Aoife Cahill and Joel Tetreault. (HOO 2012): Precision Isn't Everything: A Hybrid Approach to Grammatical Error Detection. In *Proceedings of the 7th Workshop on Innovative Use of Natural Language Processing for Building Educational Applications*. Montreal, Canada, June 07, 2012.

Nitin Madnani, Joel Tetreault and Martin Chodorow. (BEA 2012): Exploring Grammatical Error Correction with Not-So-Crummy Machine Translation. In *Proceedings of the 7th Workshop on Innovative Use of Natural Language Processing for Building Educational Applications*. Montreal, Canada, June 07, 2012.

Kristen Parton, Joel Tetreault, Nitin Madnani and Martin Chodorow. (WMT 2011): E-rating Machine Translation. In *Proceedings of the 6th Workshop on Statistical Machine Translation*. Edinburgh, Scotland, July 30-31, 2011.

Joel Tetreault, Elena Filatova and Martin Chodorow. (BEA 2010): Rethinking Grammatical Error Annotation and Evaluation with the Amazon Mechanical Turk. In *Proceedings of the 5th Workshop on Innovative Use of NLP for Building Educational Applications*. Los Angeles, CA, June 05, 2010.

Joel Tetreault and Martin Chodorow. (WAC 2009): Examining the Use of Region Web Counts for ESL Error Detection. In *Proceedings of the 5th Web as Corpus Workshop*. San Sebastian, Spain. September 07, 2009

John Lee, Joel Tetreault and Martin Chodorow. (LAW 2009): Human Evaluation of Article and Noun Number Usage: Influences of Context and Construction Variability.

In *Proceedings of the 3rd Linguistic Annotation Workshop*. Suntec, Singapore, August 06-07, 2009.

Joel Tetreault and Martin Chodorow. (HJCL 2008): Native Judgments of Non-Native Usage: Experiments in Preposition Error Detection. In *Proceedings of the COLING Workshop on Human Judgments in Computational Linguistics*. Manchester, UK, August 23, 2008.

Martin Chodorow, Joel Tetreault and Na-Rae Han. Detection of Grammatical Errors Involving Prepositions. In *Proceedings of the Fourth ACL-SIGSEM Workshop on Prepositions*. Prague, Czech Republic, June 28, 2007.

## PATENTS

### Full Patent Applications

Lei Chen, Joel Tetreault, Xiaoming Xi, Klaus Zechner, Miao Chen and Su-Youn Yoon. *Processor Implemented Systems and Methods for Measuring Syntactic Complexity on Spontaneous Non-Native Speech Data by Using Structural Event Detection*. US Patent Application #20110213610

### Provisional Patent Applications

Jill Burstein, Beata Beigman-Klebanov, Nitin Madnani, and Joel Tetreault. *Building Subjectivity Lexicons for Sentiment Analysis Systems for Use in Essay Scoring*. Provisional US Patent filing.

Ross Israel, Joel Tetreault, and Martin Chodorow. *Correcting Comma Errors in Learner Essays, and Restoring Commas in Newswire Text*. Provisional US Patent filing.

Nitin Madnani, Michael Heilman, Joel Tetreault and Martin Chodorow. *Identifying High-Level Organizational Elements in Argumentative Discourse*. Provisional US Patent filing.

## Operational ETS Software

**Preposition Error Detection (2008, 2009):** *e-rater* microfeature which detects preposition selection errors, a common error for ESL writers.

**Positive Preposition Feature (2010):** *e-rater* microfeature which rates overall preposition usage in an essay (integrated with a positive collocation microfeature into *ColPrep*).

**ESD-Lite (2011):** A flexible and portable module for plagiarism detection used primarily for the WA ProTeach contract.

**Comma Error Detection (2012):** Managed development of *e-rater* microfeature which detects comma errors in an essay.

**Shell (2012):** Managed development of new shell detection module.

## SHARED TASK COMPETITIONS

**WMT11 Shared Task on Machine Translation Evaluation:** Placed first (MTeRater-Plus) and fifth (MTeRater).

**HOO2012 Shared Task on Grammatical Error Detection:** Placed third (ETS team).

## PROFESSIONAL SERVICE

**NAACL Treasurer** (2013 – present)

**ACL 2014 Coordinating Committee** (2013 – present)

**Head Organizer:** Workshop on Innovative Use of NLP for Building Educational Applications (2008 - 2013):

<http://www.cs.rochester.edu/u/tetreault/naacl-bea8.html>

**Head Organizer:** 1st Shared Task on Native Language Identification (2013):

<http://www.nlisharedtask2013.org/>

**Organizing Committee:** CoNLL Grammatical Error Correction Shared Task (2013):  
<http://nlp.comp.nus.edu.sg/conll13st/>

**Organizing Committee:** First Workshop on Metaphor in NLP (2013):  
<https://sites.google.com/site/1stworkshoponmetaphorinnlp2013/home>

**Chair:** Area Chair and Session Chair for NAACL '10, Special Program Chair for AAAI '12, Exhibits Chair and Session Chair for NAACL '12. Session Chair for NAACL '13. Nominated for NAACL Secretary '12.

**Reviewer:** for ACL, NAACL, EACL, COLING, EMNLP, LREC, IJCAI, DAARC, SIGdial, AAAI, RANLP, Computational Linguistics

**NLP/CL Conference Calendar:** Maintain primary conference calendar for NLP and Computational Linguistics conferences and workshops (2004 - 2013):  
<http://www.cs.rochester.edu/u/tetreaul/conferences.html>

TEACHING  
EXPERIENCE

**CSC240: Computational Brain** 2003  
Head Teaching Assistant University of Rochester  
An undergraduate computational and neuroscience writing course. Managed Teaching Assistants and graded writing assignments.

**CSC242: Artificial Intelligence** 1998–2001  
Teaching Assistant University of Rochester  
Created and graded programming assignments, taught sections, created programming tests.

**CSC484: Parallel Algorithms** 1999  
Teaching Assistant University of Rochester  
Graduate course in algorithm theory and design. Graded assignments.

COMPUTER  
SKILLS

Perl, LISP, C/C++, Unix shell scripts, Python, Java, XML, HTML,  $\text{\LaTeX}$

REFERENCES

Available upon request