

# Kathleen McKeown

Department of Computer Science  
Columbia University  
New York, NY. 10027 U.S.A.

Phone: 212-939-7118  
Fax: 212-666-0140

email: [kathy@cs.columbia.edu](mailto:kathy@cs.columbia.edu)  
website: <http://www.cs.columbia.edu/~kathy>

## Current position

*Henry and Gertrude Rothschild Professor of Computer Science, Columbia University, New York*

## Research Interests

**Computational Linguistics/Natural-Language Processing:** Text Summarization; Language Generation; Social Media Analysis; Open-ended question answering; Sentiment Analysis

## Education

1982 PH.D., Computer and Information Science, University of Pennsylvania  
1979 M.S., Computer and Information Science, University of Pennsylvania  
1976 A.B., Comparative Literature, Brown University

## Appointments held

2012-2018 Founding Director, Columbia Data Science Institute  
2011-2012 Vice Dean for Research, School of Engineering and Applied Science  
2005-present Henry and Gertrude Rothschild Professor of Computer Science, Columbia University  
1997-present Professor, Columbia University  
7/2003-12/2003 Acting Chair, Department of Computer Science, Columbia University  
1997-2002 Chair, Department of Computer Science, Columbia University  
1987-1997 Associate Professor, Columbia University  
1982-1987 Assistant Professor, Columbia University

## Honors & awards

2018 Keynote Speaker, ROCLING-18 speaker, Hsinchu, Taiwan.  
2016 Keynote Speaker, International Semantic Web Conference, Kobe, Japan  
2016 Grace Hopper Distinguished Lecture, University of Pennsylvania Philadelphia, Pa., Nov. 2016. Distinguished Lecturer, University of Edinburgh, Launch of the Centre for Doctoral Training in Data Science, Edinburgh, Scotland  
2014 Keynote Speaker, 10th IEEE International Conference on eScience, Guarujá, Sao Paulo, Brazil, October  
2013 Keynote Speaker, The 2013 Conference of the North American Association for Computational Linguistics

- 2012 Best paper award, Kristen Parton, Nizar Habash, Kathleen McKeown, Gonzalo Iglesias, and Adria de Gispert. Can Automatic Post-Editing Make MT More Meaningful? In Proceedings of the 16th Conference of the European Association for Machine Translation (EAMT), Trento, Italy
- 2011 Founding Fellow, Association of Computational Linguistics
- 2010 Columbia Great Teacher Award
- 2010 Anita Borg Woman of Vision Award for Innovation
- 2010 Best student paper award, Elson, D. and McKeown, K., “Automatic Social Network Extraction for Literary Fiction,” 48th Annual Conference of the Association for Computational Linguistics
- 2003 Fellow, Association for Computing Machinery
- 2000 Outstanding Woman Scientist, New York Association of Women in Science
- 1994 Fellow, American Association of Artificial Intelligence
- 2007 Invited speaker, CICLING-07, Mexico
- 2006 Invited speaker, INLG-06, Australia
- 2000 Invited speaker, INLG-2000, Israel
- 1997 Invited speaker, IJCAI-97, Japan
- 1997 Invited speaker, ACSC-97 (Australian Computer Science Conference, Australia)
- 1993 Invited speaker, PACLING-93, British Columbia
- 1991 NSF Faculty Award for Women (5 years)
- 1990 The Instructional Technology Research Award, presented by the American Society for Training and Development for an outstanding research contribution which impacts instructional technology practice, jointly with Steven K. Feiner, for “COMET: Generating Coordinated Multimedia Explanations,” a video
- 1990 Best Paper Award, Feiner, S. and McKeown, K.R., Generating Coordinated Multimedia Explanations, IEEE Conference on AI Applications
- 1985-1986 IBM Faculty Development Award
- 1985 NSF Presidential Young Investigator Award (5 years)
- 1983 American Society of Information Science award for outstanding Doctoral Dissertation
- 1982 Rubinoff Award for outstanding achievement in the Doctoral Dissertation
- 1978, 1979 IBM PhD Research Fellowship, awarded twice

## Professional Activities

- 2018-2019 Co-Chair, Workshop on Interaction for AI, part of the CCC Roadmap for Artificial Intelligence
- 2016-2019 Co-Chair, National Academies Roundtable on Data Science Post-secondary Education
- 2017 Organizing Committee, Workshop on Challenges in Machine Generation of Automated Products from Multi-Source Data, National Academies of Science, Engineering and Medicine
- 2018 Area Chair, Summarization and Generation, Association for Computational Linguistics
- 2018 National Science Foundation Panel in CyberSystems
- 2014-2109 Member, ACM Prize Award Committee
- 2014 Member, Microsoft Research Faculty Award Committee
- 2014 Organizing Committee, ACL Workshop on Language and Computational Social Science, Baltimore, MD
- 2008 General Conference Chair, Association for Computational Linguistics Annual Conference
- 2005-2009 Member, AAAI Newell Awards Committee
- 2006 Co-Chair, AAAI-06 Senior Member Presentations
- 2003-2004 CRA Undergraduate Research Awards Committee
- 2001-2004 Secretary, Computing Research Association Board
- 1999-2004 Computing Research Association Board Member
- 1995-1997 Secretary/Treasurer, Association for Computational Linguistics
- 1996-1999 Editorial Board, Free Speech Journal
- 1993-1996 Editorial Board, Journal of Artificial Intelligence Research

1993-1995 Nominating Committee, Association for Computational Linguistics  
 1992-1993 President, Association for Computational Linguistics

## Research Activities

Currently directing seven PhD graduate students. PhD students are Emily Allaway, Noura Farra, Chris Hidey, Chris Kedzie, Fei-Tzin Lee, Jessica Ouyang, Elsbeth Turcan.

### Graduated Students

- Or Biran, 2016, NJoin, NY NY
- Sara Rosenthal, 2014, IBM Watson, White Plains, NY
- Yves Petinot, 2014, startup, Paris, France.
- Kapil Thadani, 2013, Yahoo Research Labs, NY NY.
- Kristen Parton, 2013, Facebook, NY NY.
- David Elson, 2011, Google, NY, NY
- Michel Galley, 2007, Microsoft Research, Seattle, Wa.
- Elena Filatova, 2007, Assistant Professor, New York Institute of Technology, NY, NY.
- Sasha Blair-Goldensohn, 2007, Technical Staff, Google, New York, NY.
- Ani Nenkova, 2006, Associate Professor, University of Pennsylvania, Philadelphia, Pa.
- Noemie Elhadad, 2006, Associate Professor, Department of Biomedical Informatics, Columbia University.
- Barry Schiffman, 2005, Consultant.
- David Evans, 2005, Amazon.com, Tokyo, Japan.
- Pablo Duboue, 2005, start-up.
- Regina Barzilay, 2003, Professor, MIT, Boston, MA.
- Carl Sable, 2002, Associate Professor, Cooper Union, NY NY
- Min Yen Kan, 2002, Associate Professor, National University of Singapore.
- Shimei Pan, 2001, University of Maryland at Baltimore County, Baltimore, Md.
- James Shaw, 2001, IHeart Radio, NY NY.
- Hongyan Jing, 2001, Context, New York, NY.
- Dragomir Radev, 1999, Professor, Yale University, New Haven, Ct.
- Vasileios Hatzivassiloglou, 1998, retired.
- Eric Siegel, October 1997, Predictive Analytics.
- Pascale Fung, 1996, Professor, Hong Kong Institute of Science and Technology.

- Jacques Robin, 1994. University of Paris 1, Paris, France.
- Michael Elhadad, 1992. Professor, Ben Gurion Univ., Israel.
- Ursula Wolz, 1992. Professor, The College of New Jersey, Trenton, N.J.
- Andrea Danyluk, 1992. Professor Williams College.
- Frank Smadja, 1991. Consulting, Israel.
- Galina Datskovsky Moerdler, 1990. VaporStream, N.Y., N.Y.
- Cecile Paris, 1987, Senior Research Leader in the area of Knowledge Based Systems and AI, CSIRO, Sydney, Australia.

## Grants

### Recent Funding

"A System for Cross lingual Information Processing, Translation and Summarization: SCRIPTS", McKeown (PI), Hirschberg (co-PI), Muresan (co-PI), Rambow (co-PI), Oard (co-PI, UMD), Carpuat (co-PI, UMD), Daume (co-PI, UMD), Resnik (co-PI, UMD), Radev (co-PI, Yale), Renals (co-PI, Edinburgh), Heafield (co-PI, Edinburgh), Gales (co-PI, Cambridge), IARPA, \$14,000,00, Oct 1st, 2017 Sept 30th, 2021 (4 years provided all evaluations passed). My share: \$4,187,469.

"GAIA", Chang (PI), McKeown (co-PI), subcontract from ISI, DARPA, \$900,00, Jan 1st, 2018 Dec. 31st, 2022. My share \$675,000.

"SEEM: Inferring SEntiment and EMotion from Low Resource Languages," McKeown (PI), Hirschberg (co-PI), Diab (Co-PI, GWU), DARPA, \$1,785,792. My share: \$774,528. July 1, 2015 June 30.

"Reducing plug-load electricity footprint of residential buildings through low-cost, non-intrusive sub-metering and personalized feedback technology," Culligan (PI), McKeown (co-PI), Meinreken (co-PI), DOE, total of \$1,999,782 over three years starting September 2016. Goal: to change consumer behavior in electricity usage through the use of personalized feedback. McKeown supervises one student who is developing language generation for feedback. My share: \$350,000.

"Northeast Big Data Innovation Hub," McKeown (PI), including a total of 99 universities, companies, and non-profits in the region, NSF, \$1,250,000. Oct 1, 2015 Sept 30, 2018.

"RI:Small: Describing Disasters and the Ensuing Personal Toll," Kathleen McKeown (PI), NSF, Sept. 2014 - Aug. 2017, \$543,000.

"DRATS: Detecting Relations and Anomalies in Text and Speech," DARPA, Owen Rambow (PI), Mona Diab (George Washington Univ.), Julia Hirschberg, Smaranda Muresan (Rutgers), Kathleen McKeown, and Mari Ostendorf (University of Washington), \$5,499,999, 11/13/12 - 5/12/17.

"Broad Operational Language Translation/IBM Team - Algorithmic Development and Integrated Systems", DARPA, Subcontract to IBM, Kathleen McKeown (PI), \$550,000 Jan 1, 2013- Dec. 31, 2014.

"RI: Large: Collaborative Research: Richer Representations for Machine Translation," Kathleen McKeown, jointly with Martha Palmer (PI), Colorado University, Kevin Knight, University of Southern California, Daniel Gildea, Rochester University, and Bert Nangwen, Brandeis University, \$593,045 (Columbia's

portion), Sept. 1 2009 - Dec. 31st 2014.

"Discovering and explaining technical emergence through analysis of the language and structure of scientific publications," Kathleen McKeown (PI), Michael Collins, Luis Gravano, Owen Rambow (Columbia University); Simone Teufel (Cambridge University), Bonnie Dorr, Kenneth Fleischman, Hal Daume (University of Maryland), Dragomir Radev (University of Michigan), Fei Xia (University of Washington), \$6,000,000, July 7, 2011 - May 31, 2014.

"Broad Operational Language Translation/SRI Team", DARPA, Subcontract to IBM, Julia Hirschberg (PI), Michael Collins, Mona Diab and Kathleen McKeown, 10/1/2011-9/30/2012, \$219,206 (McKeown's share).

"Predicting Power, Influence and Rifts in Social Groups Through Linguistic Analysis: A Study of English, Arabic, and Urdu," Kathleen McKeown, Mona Diab, Dragomir Radev, Owen Rambow, IARPA, \$672,472 (year 1) option for \$1,309,568 (Nov 2010 - Oct 2012). Option for year 2 approved.

"Distillation for ROSETTA," Kathleen McKeown (Subcontract PI), subcontract to IBM, DARPA GALE, \$180,000, July 2009 - June 2011.

"EAGER: Corpus-Based Narrative Semantics," Kathleen McKeown, NSF, \$253,888 total, July 1, 2009 - December 31st, 2010.

"Task-Embedded Machine Translation," Kathleen McKeown, sub-contract to IBM, DARPA GALE, \$107,149, 9/1/2009-5/31/2010.

"Contextually Sensitive Semantic Relationships," Kathleen McKeown and Barry Schiffman, NGA HM1582-06-1-2015, \$119,058/year for 2 years, 09/08/06-09/07/08.

"A Multimodal Interface for Generating Physician Notes," Steven Feiner, Kathleen McKeown, and Desmond Jordan, IBM, \$143,000/year for 3 years, Sept 1, 2006.

"Text-to-Text Generation for Summarizing Informal Genres," Kathleen McKeown, NSF, \$300,000, Jan 2006 - Dec 2009.

\$25,000, IBM Faculty Development Award, January 2005.

\$20,000, Microsoft Research Gift, May 2005.

"NIGHTINGALE: Novel Information Gathering and Harvesting Techniques", \$4,884,605, Kathleen McKeown, Julia Hirschberg and Dan Ellis, DARPA, September 1 2005 - August 31, 2010 (Evaluation milestones were met in year 1, grant was renewed for year 2. Renewal for Year 3 is dependent on meeting year 2 evaluation milestones).

"Text Summarization: News and Beyond," DARPA, \$200,000, March 1 2005-August 31, 2005.

"Tools for Monitoring Online Information Sources," Kathleen McKeown, Luis Gravano, Julia Hirschberg and Owen Rambow, NSF KDD program, \$1,795,000, renewed yearly June 1, 2002 - January 31, 2006.

"Multilingual Multidocument Information Tracking and Summarization," Kathleen McKeown, Vasileios Hatzivassiloglou, Judith Klavans, DARPA, \$1,800,000, March 1, 2000 - February 28, 2005.

“ITR/PE+SY: Mapping Meetings: Language Technology to make Sense of Human Interaction,” Dan Ellis (Dept of Electrical Engineering) and Kathleen McKeown, \$269,904 (Columbia’s subcontract), September 1, 2001 - August 31, 2005 (awarded to Univ of California at Berkeley, Nelson Morgan, PI, Univ of Washington, Mari Ostendorf, PI).

“Question Answering - Phase II”, Vasileios Hatzivassiloglou, Kathleen McKeown (Columbia University), Wayne Ward, Daniel Jurafsky (University of Colorado), ARDA AQUAINT, \$1,900,000, April, 2004-March 2006.

“Question Answering”, Vasileios Hatzivassiloglou, Kathleen McKeown (Columbia Univeristy), Wayne Ward, Daniel Jurafsky (University of Colorado), ARDA AQUAINT, \$1,800,000, November, 2001-November 2003.

”A Patient Care Digital Library: Personalized Search and Summarization over Multimedia Information,” Kathleen McKeown, Shih-Fu Chang, James Cimino, George Hripcsak, and Judith Klavans, NSF, \$5,000,000, September 1, 1999 - August 31, 2004.

## Publications

### Journals

- [1] Desmond Upton Patton, Jamie MacBeth, Sarita Schoenebeck, Katherine Shear, and Kathleen McKeown. Accommodating grief on twitter: An analysis of expressions of grief among gang involved youth on twitter using qualitative analysis and natural language processing. *Biomedical Informatics Insight*, 10, Jan 2018.
- [2] Mohammad Sadegh Rasooli, Noura Farra, Axinia Radeva, Tao Yu, and Kathleen McKeown. Cross-lingual sentiment transfer with limited resources. *Machine Translation*, Nov 2017.
- [3] Sara Rosenthal and Kathleen McKeown. Detecting influencers in multiple online genres. *ACM Transactions on Internet Technology (TOIT): Special Issue on Advances in Social Computing*, 17(2), May 2017.
- [4] Kathleen McKeown, Hal Daume III, Snigdha Chaturvedi, John Paparrizos, Kapil Thadani, Pablo Barrio, Or Biran, Suvarna Bothe, Michael Collins, Kenneth R. Fleischmann, Luis Gravano, Rahul Jha, Ben King, Kevin McInerney, Taesun Moon, Arvind Neelakantan, Diarmuid O’Seaghdha, Dragomir Radev, Clay Templeton, and Simone Teufel. Predicting the impact of scientific concepts using full text features. *Journal of the Association for Information Science and Technology (JASIST)*, January 2016.
- [5] Wei-Yun Ma and Kathleen McKeown. Detecting and correcting syntactic errors in machine translation using feature-based lexicalized tree adjoining grammars. *International Journal of Computational Linguistics and Chinese Language Processing (IJCLCLP)*, 17(4):1–14, December 2012.
- [6] Ani Nenkova and Kathleen McKeown. Automatic summarization. *Foundations and Trends in Information Retrieval*, 52(2-3):103–233, 2011. <http://dx.doi.org/10.1561/15000000015>.
- [7] Advaith Siddharthan, Ani Nenkova, and Kathleen McKeown. Information status distinctions and referring expressions: An empirical study of references to people in news summaries. *Journal of Computational Linguistics*, 37(4), 2011.
- [8] Ani Nenkova, Rebecca Passonneau, and Kathleen McKeown. The pyramid method: incorporating human content selection variation in summarization evaluation. *ACM Transactions on Speech and Language Processing*, 4(2), 2007.
- [9] Regina Barzilay and Kathleen McKeown. Sentence fusion for multidocument summarization. *Computational Linguistics*, 31, 2005.
- [10] Noemie Elhadad, Min-Yen Kan, Judith Klavans, and Kathleen McKeown. Using different levels of customization in a unified framework for summarizing medical literature. *Journal of Artificial Intelligence in Medicine*, 33(2):179–98, December 2004.
- [11] Shimei Pan, Kathleen McKeown, and Julia Hirschberg. Exploring features from natural language generation for prosody modeling. *Computer Speech and Language (special issue on spoken language generation)*, 16(3):457–90, October 2002.
- [12] Desmond Jordan, Kathleen McKeown, Kristian Concepcion, Steven Feiner, and Vasileios Hatzivasiloglou. Generation and evaluation of medical inferences for automated healthcare briefings. *Journal of American Medical Informatics Association*, 8, may 2001.
- [13] Eric Siegel and Kathleen McKeown. Learning methods to combine linguistic indicators: Improving aspectual classification and revealing linguistic insights. *Computational Linguistics*, 26, December 2000.

- [14] Kathleen McKeown and Shimei Pan. Prosody modeling in concept-to-speech generation: methodological issues. *Phil. Trans. Royal Society London*, 358(1769):1419–1431, April 2000.
- [15] Alfred Aho, Shih-Fu Chang, Kathleen McKeown, Dragomir Radev, John Smith, and Kazi Zaman. Columbia digital news project. *International Journal of Digital Libraries*, 1(4):377–385, 1998.
- [16] Dragomir Radev and Kathleen McKeown. Generating natural language summaries from multiple on-line sources. *Computational Linguistics*, 24(3):469–500, September 1998.
- [17] Kathleen McKeown, Steven Feiner, Mukesh Dalal, and Shih-Fu Chang. Generating multimedia briefings: Coordinating language and illustration. *Artificial Intelligence Journal*, 103:95–116, 1998.
- [18] Pascale Fung and Kathleen McKeown. Technical word and term translation from noisy parallel corpora across language groups. *Machine Translation*, November 1997.
- [19] Michael Elhadad, Kathleen McKeown, and Jacques Robin. Floating constraints in lexical choice. *Journal of Computational Linguistics*, June 1997.
- [20] Jacques Robin and Kathleen McKeown. Empirically designing and evaluating a new revision-based model for summary generation. *Artificial Intelligence Journal*, 85, August 1996.
- [21] Frank Smadja, Kathleen McKeown, and Hatzivassiloglou. Translating collocations for bilingual lexicons: A statistical approach. *Journal of Computational Linguistics*, 22(1), March 1996.
- [22] Kathleen McKeown, Karen Kukich, and Jacques Robin. Generating concise natural language summaries. *Journal of Information Processing and Management*, 31(5), 1995.
- [23] Cole et. al. The challenge of spoken language systems: Research directions for the nineties. *IEEE Transactions on Speech and Audio Processing*, 3(1):1–22, January 1995.
- [24] Steven Feiner and Kathleen McKeown. Automating the generation of coordinated multimedia explanations. *IEEE Computer*, 24(10):33–41, October 1991.
- [25] Frank Smadja and Kathleen McKeown. Using collocations for language generation. *Computational Intelligence*, 7(4), December 1991.
- [26] Ursula Wolz, Kathleen McKeown, and Gail Kaiser. Automated tutoring in interactive environments: A task centered approach. *Journal of Machine Mediated Learning*, December 1989.
- [27] Kathleen McKeown. Generating goal oriented explanations. *International Journal of Expert Systems*, I, Issue 4,;377–95, December 1988.
- [28] Kathleen McKeown. Language generation: Applications, issues, and approaches. *Proceedings of the IEEE*, July 1986.
- [29] Kathleen McKeown. Paraphrasing questions using given and new information. *American Journal of Computational Linguistics*, January-March 1983.
- [30] Kathleen McKeown. Discourse strategies for generating natural language text. *Artificial Intelligence Journal*, 27 No. 1., 1985.
- [31] Bonnie L. Webber, Aravind K. Joshi, Eric Mays, and Kathleen McKeown. Extended natural language data base interaction. *International Journal of Computers and Mathematics*, December 1982.
- [32] Kathleen McKeown and Norman I. Badler. Creating polyhedral stellations. *Computer Graphics*, 14, No. 3:19–24, July 1980.



## Conference Proceedings

- [33] Christopher Hidey and Kathleen McKeown. Fixed that for you: Generating contrastive claims with semantic edits. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technology (NAACL-HLT)*, Minneapolis, Minnesota, June 2019.
- [34] Jessica Ouyang, Boya Song, and Kathleen McKeown. A robust abstractive system for cross-lingual summarization. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technology (NAACL-HLT)*, Minneapolis, Minnesota, June 2019.
- [35] Tuhin Chakrabarty, Christopher Hidey, and Kathleen McKeown. Imho fine-tuning improves claim detection. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technology (NAACL-HLT)*, Minneapolis, Minnesota, June 2019.
- [36] Philipp Blandfort, Desmond Patton, William R. Frey, Svebor Karaman, Surabhi Bhargava, Fei-Tzin Lee, Siddharth Varia, Chris Kedzie, Michael B. Gaskell, Rossano Schifanella, Kathleen R. McKeown, and Shih-Fu Chang. Multimodal social media analysis for gang violence prevention. *CoRR*, to appear in *ICWSM-2019*, abs/1807.08465, 2018.
- [37] Serina Chang, Ruiqi Zhong, Ethan Adams, Fei-Tzin Lee, Siddharth Varia, Desmond Patton, William Frey, Chris Kedzie, and Kathleen McKeown. Detecting gang-involved escalation on social media using context. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 46–56. Association for Computational Linguistics, 2018.
- [38] Chris Kedzie, Kathleen McKeown, and Hal Daume III. Content selection in deep learning models of summarization. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1818–1828. Association for Computational Linguistics, 2018.
- [39] Christopher Thomas Hidey and Kathleen McKeown. Persuasive influence detection: The role of argument sequencing. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, Louisiana, USA, February 2018.
- [40] Or Biran and Kathleen McKeown. Human-centric justification of machine learning predictions. In *Proceedings of IJCAI*, Melbourne, Australia, August 2017.
- [41] Or Biran and Kathleen McKeown. Domain-adaptable hybrid generation of rdf entity descriptions. In *Proceedings of IJCNLP*, Melbourne, Australia, August 2017.
- [42] Noura Farra and Kathleen McKeown. Smarties: Sentiment models for arabic target entities. In *Proceedings of EACL*, Valencia, Portugal, April 2017.
- [43] Jessica Ouyang, Serina Chang, and Kathleen McKeown. Crowd-sourced iterative annotation for narrative summarization corpora. In *Proceedings of EACL*, Valencia, Portugal, April 2017.
- [44] Terra Blevins, Robert Kwiatkowski, Jamie Macbeth, Kathleen McKeown, Desmond Patton, and Owen Rambow. Automatically processing tweets from gang-involved youth: Towards detecting loss and aggression. In *Proceedings of COLING*, Mumbai, India, December 2016.
- [45] Using natural language processing and qualitative analysis to intervene in gang violence: A collaboration between social work researchers and data scientists. In *Proceedings of Bloomberg Conference on Social Good*, New York, NY, October 2016.

- [46] Or Biran, Terra Blevins, and Kathleen McKeown. Mining paraphrasal typed templates from a plain text corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, Berlin, Germany, August 2016.
- [47] Christopher Hidey and Kathleen McKeown. Identifying causal relations using parallel wikipedia articles. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, Berlin, Germany, August 2016.
- [48] Gavin Saldanha, Or Biran, Kathleen McKeown, and Alfio Gliozzo. An entity-focused approach to generating company descriptions. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, Berlin, Germany, August 2016.
- [49] Chris Kedzie, Fernando Diaz, and Kathleen McKeown. Real-time web scale event summarization using sequential decision making. In *Proceedings of IJCAI*, New York, NY, July 2016.
- [50] Or Biran and Kathleen McKeown. Pdtb discourse parsing as a tagging task: The two taggers approach. In *Proceedings of the 16th Annual SIGdial Meeting on Discourse and Dialogue*, SIGDIAL 2015, Prague, Czech Republic, 2015.
- [51] Sara Rosenthal and Kathleen McKeown. I couldn't agree more: The role of conversational structure in agreement and disagreement detection in online discussions. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 168–177, Prague, Czech Republic, September 2015. Association for Computational Linguistics.
- [52] Wei-Yun Ma and Kathleen McKeown. System combination for machine translation through paraphrasing. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1053–1058, Lisbon, Portugal, September 2015. Association for Computational Linguistics.
- [53] Jessica Ouyang and Kathleen McKeown. Modeling reportable events as turning points in narrative. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2149–2158, Lisbon, Portugal, September 2015. Association for Computational Linguistics.
- [54] Chris Kedzie, Kathleen McKeown, and Fernando Diaz. Predicting salient updates for disaster summarization. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing*, pages 1608–1617, 2015.
- [55] Or Biran and Kathleen McKeown. Aggregated word pair features for implicit discourse relation disambiguation. In *Proceedings of ACL*, Sofia, Bulgaria, 2013.
- [56] Wei-Yun Ma and Kathleen McKeown. Using a supertagged dependency model to select a good translation in system combination. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technology (NAACL-HLT)*, Atlanta, Georgia, June 2013.
- [57] Yves Petinot, Kathleen McKeown, and Kapil Thadani. Cluster-based web summarization. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pages 1124–1128, Nagoya, Japan, October 2013. Asian Federation of Natural Language Processing.
- [58] Kapil Thadani and Kathleen McKeown. Sentence compression with joint structural inference. In *Proceedings of CoNLL*, Sofia, Bulgaria, 2013.
- [59] Kapil Thadani and Kathleen McKeown. Supervised sentence fusion with single-stage inference. In *Proceedings of IJCNLP*, Nagoya, Japan, 2013.

- [60] Wei-Yun Ma and Kathleen McKeown. Phrase-level system combination for machine translation based on target-to-target decoding. In *Proceedings of the Tenth Biennial Conference of the Association for Machine Translation in the Americas*, San Diego, Ca., November 2012.
- [61] Kristen Parton, Nizar Habash, and Kathleen McKeown. Lost and found in translation: The impact of machine translated results on translingual information retrieval. In *Proceedings of the Tenth Biennial Conference of the Association for Machine Translation in the Americas*, San Diego, Ca., November 2012.
- [62] Wei-Yun Ma and Kathleen McKeown. Detecting and correcting syntactic errors in machine translation using feature-based lexicalized tree adjoining grammars. In *Proceedings of Conference on Computational Linguistics and Speech Processing (ROCLING)*, ChungLi, Taiwan, September 2012.
- [63] Sara Rosenthal and Kathleen McKeown. Detecting opinionated claims in online discussions. In *Proceedings of the Sixth IEEE International Conference on Semantic Computing (IEEE ICSC2012) Special Session on Semantics and Sociolinguistics in Social Media*, Palermo, Italy, September 2012.
- [64] Jacob Andreas, Sara Rosenthal, and Kathleen McKeown. Annotating agreement and disagreement in threaded discourse. In *Proceedings of the 8th International Conference of Language Resources and Evaluation (LREC)*, Istanbul, Turkey, May 2012.
- [65] William Yang Wang, Kapil Thadani, and Kathleen McKeown. Identifying event descriptions using co-training with online news summaries. In *Proceedings of IJNLP*, Chiang-Mai, Thailand, November 2011.
- [66] Yves Petinot and Kathleen McKeown. A hierarchical model for web summarization. In *Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL-HLT)*, Portland, Oregon, June 2011.
- [67] Sara Rosenthal and Kathleen McKeown. Age prediction in blogs: A study of style, content and online behavior in pre- and post-social media generations. In *Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL-HLT)*, Portland, Oregon, June 2011.
- [68] Kapil Thadani and Kathleen McKeown. Optimal and syntactically informed decoding for monolingual phrase-based alignment. In *Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL-HLT)*, Portland, Oregon, June 2011.
- [69] Kristen Parton and Kathleen McKeown. Mt error detection for cross-lingual question answering. In *Proceedings of COLING*, Beijing, China, August 2010.
- [70] William Yang Wang and Kathleen McKeown. Got you!: Automatic vandalism detection in wikipedia with web-based shallow syntactic-semantic modeling. In *Proceedings of COLING*, Beijing, China, August 2010.
- [71] David Elson and Kathleen McKeown. Automatic social network extraction for literary fiction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, July 2010.
- [72] David Elson and Kathleen McKeown. Tense and aspect assignment in narrative discourse. In *Proceedings of INLG 2010*, Dublin, Ireland, July 2010.
- [73] David Elson and Kathleen McKeown. Automatic attribution of quoted speech in literary narrative. In *Proceedings of the 24th Annual Conference of the American Association for Artificial Intelligence*, Atlanta, Ga, July 2010.

- [74] Kathleen McKeown, Sara Rosenthal, Kapil Thadani, and Coleman Moore. Time-efficient creation of an accurate sentence fusion corpus. In *Proceedings of the 11th Annual Conference of the North American Chapter on Computational Linguistics*, Los Angeles, Ca, June 2010.
- [75] Sara Rosenthal, William J. Lipovsky, Kathleen McKeown, Kapil Thadani, and Jacob Andreas. Towards semi-automated annotation for prepositional phrase attachment. In *Proceedings of LREC 10*, Malta, May 2010.
- [76] David Elson and Kathleen R. McKeown. A tool for deep semantic encoding of narrative texts. In *Proceedings of the ACL-IJCNLP 2009 Conference*, Singapore, August 2009.
- [77] Wei-Yun Ma and Kathleen R. McKeown. Where’s the verb correcting machine translation during question answering. In *Proceedings of the ACL-IJCNLP 2009 Conference*, Singapore, August 2009.
- [78] Kristen Parton, Kathleen McKeown, Robert Coyne, Mona Diab, Ralph Grishman, Dilek Hakkani-Tur, Mary Harper, Heng Ji, Wei Yun Ma, Adam Meyers, Sara Stolbach, Ang Sun, Gokhan Tur, Wei Xu, and Sibel Yarman. Who, what, when, where, why? comparing multiple approaches to the cross-lingual 5w task. In *Proceedings of the ACL-IJCNLP 2009 Conference*, Singapore, August 2009.
- [79] Apoorv Agarwal, Fadi Biadisy, and Kathleen McKeown. Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2009)*, Athens, Greece, March, April 2009.
- [80] Kristen Parton, Kathleen McKeown, James Allan, and Enrique Henestroza. Simultaneous multilingual search for translingual information retrieval. In *Proceedings of the ACM 17th Conference on Information and Knowledge Management (CIKM)*, Napa Valley, Ca., October 2008.
- [81] Kapil Thadani and Kathleen McKeown. A framework for decreasing textual redundancy. In *Proceedings of the 22nd International Conference on Computational Linguistics (COLING)*, Manchester, England, August 2008.
- [82] David Elson and Kathleen McKeown. A platform for symbolically encoding human narratives. In *Proceedings of the AAAI 2007 Fall Symposium on Intelligent Narrative Technologies*, Arlington, Va., November 2007.
- [83] Barry Schiffman, Kathleen McKeown, Ralph Grishman, and James Allan. Question answering using integrated information retrieval and information extraction. In *Proceedings of HLT/NAACL 07*, Rochester, NY, April 2007.
- [84] Sasha Blair-Goldensohn, Kathleen McKeown, and Owen Rambow. Building and refining rhetorical-semantic relation models. In *Proceedings of HLT/NAACL 07*, Rochester, NY, April 2007.
- [85] Michel Galley and Kathleen McKeown. Lexicalized markov grammars for sentence compression. In *Proceedings of HLT/NAACL 07*, Rochester, NY, April 2007.
- [86] Kathleen McKeown, Lokesh Shrestha, and Owen Rambow. Using question-answer pairs in extractive summarization of email conversations. In *Proceedings of CICLING 07*, Mexico City, Mexico, February 2007.
- [87] Elena Filatova, Vasileios Hatzivassiloglou, and Kathleen R. McKeown. Automatic creation of domain templates. In *Proceedings of Coling/ACL 06 (Poster Session)*, pages 207–214, Sydney, Au., July 2006.
- [88] Ani Nenkova, Lucy Vandervende, and Kathleen McKeown. A compositional context sensitive multi-document summarizer. In *Proceedings of SIGIR 06*, Seattle, Wa., 2006.

- [89] Ani Nenkova, Advaith Siddharthan, and Kathleen R. McKeown. Automatically learning cognitive status for multi-document summarization of newswire. In *Proceedings of HLT/EMNLP-05*, Vancouver, Canada, 2005.
- [90] Advaith Siddharthan and Kathleen McKeown. Improving multilingual summarization: Using redundancy in the input to correct mt errors. In *Proceedings of HLT/EMNLP-05*, Vancouver, Canada, 2005.
- [91] Barry Schiffman and Kathleen McKeown. Context and learning in novelty detection. In *Proceedings of HLT/EMNLP-05 (poster)*, Vancouver, Canada, 2005.
- [92] Kathleen McKeown, Rebecca J. Passonneau, David K. Elson, Ani Nenkova, and Julia Hirschberg. Do summaries help? a task-based evaluation of multi-document summarization. In *Proceedings of SIGIR-05*, Salvador, Brazil, 2005.
- [93] David Evans and Kathleen McKeown. Identifying similarities and differences across arabic and english news. In *Proceedings of the International Conference on Intelligence Analysis*, McLean, VA, 2005.
- [94] Kathleen McKeown, Julia Hirschberg, Michel Galley, and Sameer Maskey. From text to speech summarization. In *Proceedings of ICASSP*, Philadelphia, PA, 2005.
- [95] Sasha Blair-Goldensohn, David Evans, Vasileios Hatzivassiloglou, McKeown Kathleen, Ani Nenkova, Rebecca Passonneau, Barry Schiffman, Andrew Schlaikjer, Advaith Siddharthan, and Sergei Siegelman. Columbia university at duc 2004. In *4th Document Understanding Conference (DUC 2004) at HLT/NAACL 2004*, Boston, MA, 2004.
- [96] Michel Galley, Kathleen McKeown, Julia Hirschberg, and Elizabeth Shriberg. Identifying agreement and disagreement in conversational speech: Use of bayesian networks to model pragmatic dependencies. In *Proc. of the 42nd ACL*, Barcelona, Spain, 2004.
- [97] Desmond Jordan, Gregory Whalen, Blaine Bell, Kathleen McKeown, and Steven Feiner. An evaluation of automatically generated briefings of patient status. In *Medinfo 2004*, San Francisco, Ca., 2004.
- [98] Lokesh Shrestha and Kathleen McKeown. Detection of question-answer pairs in email conversations. In *Proceedings of Coling 2004*, Geneva, Switzerland, 2004.
- [99] Stephen Wan and Kathleen McKeown. Generating "state-of-affairs" summaries of ongoing email thread discussions. In *Proceedings of Coling 2004*, Geneva, Switzerland, 2004.
- [100] Advaith Siddharthan, Ani Nenkova, and Kathleen McKeown. Syntactic simplification for improving content selection in multi-document summarization. In *20th International Conference on Computational Linguistics (COLING 2004)*, Geneva, Switzerland, 2004.
- [101] Kathleen McKeown, Noemie Elhadad, and Vasileios Hatzivassiloglou. Leveraging a common representation for personalized search and summarization in a medical digital library. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, Houston, Tx, May 2003.
- [102] Ani Nenkova and Kathleen McKeown. References to named entities: A corpus study. In *Proceedings of the Human Language Technology Conference, Vol. II*, Edmunton, Ca., May 2003.
- [103] Kathleen McKeown, Regina Barzilay, John Chen, David Elson, David Evans, Judith Klavans, Ani Nenkova, Barry Schiffman, and Sergey Sigelman. Columbia's newsblaster: New features and future directions. In *Proceedings of the Human Language Technology Conference, Vol. II*, Edmunton, Ca., May 2003.

- [104] Michel Galley and Kathleen McKeown. Improving word sense disambiguation in lexical chaining. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, poster*, Acapulco, Mexico, August 2003.
- [105] Sasha Blair-Goldenson and Kathleen McKeown. Defscriber: A hybrid system for definitional qa. In *demo, Proceedings of SIGIR 2003*, Toronto, Ca., July 2003.
- [106] Andre Kushniruk, Men-Yen Kan, Kathleen McKeown, Judith L. Klavans, Desmond Jordan, and Vimla L. Patel. Usability evaluation of an experimental text summarization system and three search engines: Implications for the reengineering of health care interfaces. In *Proceedings of the American Medical Informatics Association*, San Antonio, Tex, November 2002.
- [107] Andre Kushniruk, Men-Yen Kan, Kathleen McKeown, Judith L. Klavans, and Vimla L. Patel. Evaluating the content and usability of an experimental text summarization system and three web-based search engines. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, Baltimore, Md, October 2002.
- [108] Carl L. Sable, Kathleen McKeown, and Kenneth W. Church. Nlp found helpful (at least for one text categorization task). In *Proceedings of Empirical Methods in Natural Language Processing*, Philadelphia, Pa., July 2002.
- [109] Pablo A. Duboue and Kathleen R. McKeown. Content planner construction via evolutionary algorithms and alignment-based fitness function. In *Proceedings of the International Conference on Natural Language Generation*, Harriman, NY, July 2002.
- [110] Min Yen Kan and Kathleen McKeown. Corpus-trained text generation for summarization. In *Proceedings of the International Conference on Natural Language Generation*, Harriman, NY, July 2002.
- [111] Judith Klavans Min Yen Kan and Kathleen R. McKeown. Using the annotated bibliography as a resource for indicative summarization. In *Proceedings of the Language Resources European Conference*, Spain, May 2002.
- [112] Kathleen McKeown, Regina Barzilay, David Evans, Vasileios Hatzivassiloglou, Judith Klavans, Ani Nenkova, Carl Sable, Barry Schiffman, and Sergey Sigelman. Tracking and summarizing news on a daily basis with columbia’s newsblaster. In *Proceedings of the Human Language Technology Conference*, San Diego, Ca., March 2002.
- [113] Barry Schiffman, Ani Nenkova, and Kathleen McKeown. Experiments in multi-document summarization. In *Proceedings of the Human Language Technology Conference*, San Diego, Ca., March 2002.
- [114] Simone Teufel, Vasileios Hatzivassiloglou, Kathleen McKeown, Desmond A. Jordan, Kathleen M. Dunn, Sergey Sigelman, and Andre Kushniruk. Personalizing retrieval of journal articles for patient care. In *Proceedings of the American Medical Informatics Association Annual Conference*, Washington D.C., November 2001.
- [115] Kathleen McKeown, Shih-Fu Chang, James Cimino, Steven K. Feiner, Carol Friedman, Luis Gravano, Vasileios Hatzivassiloglou, Desmond Jordan, Judith L. Klavans, Andre Kushniruk, and Vimla Patel. Persival, a system for personalized search and summarization over multimedia healthcare information. In *Proceedings of The First ACM+IEEE Joint Conference on Digital Libraries*, Roanoke, W. Va., June 2001.
- [116] Shimei Pan, Kathleen McKeown, and Julia Hirschberg. Semantic abnormality and its realization in spoken language. In *Proceedings of Eurospeech 2001*, pages 333–336, Aalborg, Denmark, 2001.
- [117] Regina Barzilay, Noemie Elhadad, and Kathleen McKeown. Sentence ordering in multidocument summarization. In *Proceedings of HLT*, San Diego (CA), USA, 2001.

- [118] Regina Barzilay and Kathleen McKeown. Extracting paraphrases from a parallel corpus. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL-EACL 2001)*, Toulouse, France, July 2001.
- [119] Pablo Duboue and Kathleen McKeown. Empirically estimating order constraints for content planning in generation. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL-EACL 2001)*, Toulouse, France, July 2001.
- [120] Kathleen McKeown, Desmond Jordan, Steven Feiner, James Shaw, Elizabeth Chen, Shabina Ahmad, Andre Kushniruk, and Vimla Patel. A study of communication in the cardiac surgery intensive care unit and its implications for automated briefing. In *Proceedings of the American Medical Informatics Association Annual Conference*, Los Angeles, Ca., November 2000.
- [121] Barry Schiffman and Kathleen McKeown. Experiments in automated lexicon building for text searching. In *Proceedings, COLING 2000*, Saarbruecken, Germany, 2000.
- [122] James Shaw and Kathleen McKeown. Generating referring quantified expressions. In *Proceedings of the International Natural Language Generation Conference*, Mitzpah Ramon, Israel, 2000.
- [123] Hongyan Jing and Kathleen McKeown. Cut and paste based text summarization. In *Proceedings of the First Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL'2000)*, Seattle, Washington, USA, 2000.
- [124] Vasileios Hatzivassiloglou, Olga Merport, Kathleen McKeown, and Desmond A. Jordan. Extracting patient profiles from patient records and online literature. In *Proceedings of the AMIA 1999 Annual Symposium*, Washington, DC, submitted 1999. American Medical Informatics Association.
- [125] Regina Barzilay, Kathleen McKeown, and Michael Elhadad. Information fusion in the context of multi-document summarization. In *Proceedings of the the 37th Annual Meeting of the Association for Computational Linguistics*, College Park, Maryland, June 1999.
- [126] Hongyan Jing and Kathleen McKeown. The decomposition of human-written summary sentences. In *Proceedings of the 22nd International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR'99)*, University of Berkeley, CA, August 1999.
- [127] Kathleen McKeown, Judith L. Klavans, Vasileios Hatzivassiloglou, Regina Barzilay, and Eleazor Eskin. Towards multidocument summarization by reformulation: Progress and prospects. In *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-99)*, Orlando, Florida, July 1999. American Association for Artificial Intelligence.
- [128] Shimei Pan and Kathleen McKeown. Word informativeness and automatic pitch accent modeling. In *Proc. of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, June 1999.
- [129] Hongyan Jing and Kathleen McKeown. Combining multiple, large-scale resources in a reusable lexicon for natural language generation. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics*, Montreal, Can., August 1998.
- [130] James Shaw and Kathleen McKeown. An architecture for aggregation in text generation. In *Proc. of the 15th IJCAI*, Nagoya, Japan, 1997. Poster Session.
- [131] Alfred Aho, Shih-Fu Chang, Kathleen McKeown, Dragomir Radev, John Smith, and Kazi Zaman. Columbia digital news system : An environment for briefing and search over multimedia information. In *Proceedings of IEEE International Conference on the Advances in Digital Libraries*, Washington, DC, 1997.

- [132] Kathleen McKeown, Shimei Pan, James Shaw, Desmond A. Jordan, and Barry Allen. Language generation for multimedia healthcare briefings. In *Proc. of the 5th ACL Conference on Applied Natural Language Processing*, pages 277–282, Washington, DC, 1997.
- [133] Vasileios Hatzivassiloglou and Kathleen McKeown. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, Madrid, Spain, July 1997.
- [134] Dragomir Radev and Kathleen McKeown. Building a generation knowledge source using internet-accessible newswire. In *Proceedings of the ACL Conference on Applied Natural Language*, pages 221–228, Washington, D.C., April 1997.
- [135] Mukesh Dalal, Steven Feiner, Kathleen McKeown, Shimei Pan, Michelle Zhou, Tobias Hoellerer, James Shaw, Y. Feng, and Jeanne Fromer. Negotiation for automated generation of temporal multimedia presentations. In *Proc. ACM Multimedia '96*, pages 55–64, Boston, MA, November 18–22 1996.
- [136] Shimei Pan and Kathleen McKeown. Spoken language generation in a multimedia system. In *Proceedings of ICSLP 96*, volume 1, pages 374–377, Philadelphia, PA, 1996.
- [137] Mukesh Dalal, Steven Feiner, Kathleen McKeown, Desmond Jordan, Barry Allen, and Yassir alSafadi. Magic: An experimental system for generating multimedia briefings about post-bypass patient status. In *Proc. 1996 AMIA Annual Fall Symp*, pages 684–688, Washington, DC, October 26–30 1996.
- [138] Eric Seigel and Kathleen McKeown. Gathering statistics to aspectually classify sentences with a genetic algorithm. In *NEMLAP Proceedings*, Ankara, Turkey., September 1996.
- [139] Kathleen McKeown and Dragomir Radev. Generating summaries of multiple news articles. In *Proceedings of SIGIR*, Seattle, Wash., July 1995.
- [140] Vasileios Hatzivassiloglou and Kathleen McKeown. Automatic prediction of semantic markedness. In *Proceedings of the Annual Conference of the Association for Computational Linguistics*, Cambridge, Mass., June 1995.
- [141] Kathleen McKeown, Karen Kukich, and James Shaw. Practical issues in automatic documentation generation. In *Proceedings of the ACL Conference on Applied Natural Language*, Stuttgart, Germany, October 1994.
- [142] Eric Siegel and Kathleen McKeown. Emergent linguistic rules from the automatic grouping of training examples. In *Proceedings of the National Conference on Artificial Intelligence*, Seattle, Wash., July 1994.
- [143] Frank Smadja and Kathleen McKeown. Translating collocations for use in bilingual lexicons. In *Proceedings of the ARPA Human Language Technology Workshop*, Princeton, N.J., March 1994.
- [144] Karen Kukich, Kathleen McKeown, James Shaw, Jacques Robin, Jong Lim, Nelson Morgan, and John Phillips. User-needs analysis and design methodology for an automated documentation generator. In *Proceedings of the Fourth Bellcore/BCC Symposium on User-Centered Design*, Piscataway, N.J., November 1993.
- [145] Vasileios Hatzivaassiloglou and Kathleen McKeown. Towards the automatic identification of adjectival scales: clustering of adjectives according to meaning. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, Columbus, Oh., June 1993.
- [146] Kathleen McKeown, Jacques Robin, and Michael Tanenblatt. Tailoring lexical choice to the user's vocabulary in multimedia explanation generation. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, Columbus, Oh., June 1993.



- [147] Jacques Robin and Kathleen McKeown. Corpus analysis for revision-based generation of complex sentences. In *Proceedings of the National Conference on Artificial Intelligence*, Washington, D.C., July 1993.
- [148] Kathleen McKeown, Steven Feiner, Jacques Robin, Doree D. Seligmann, and Michael Tanenblatt. Generating cross references for multimedia explanation. In *Proceedings of the National Conference on Artificial Intelligence*, San Jose, Ca., July 1992.
- [149] Michael Elhadad, Steven Feiner, Kathleen McKeown, and Doree D. Seligmann. Generating customized text and graphics in the comet explanation testbed. In *Proceedings of the 1991 Winter Simulation Conference*, pages 1058–65, Phoenix, Ariz., December 1991.
- [150] Michael Elhadad and Kathleen McKeown. A procedure for generating connectives. In *Proceedings of COLING*, Helsinki, Finland, July 1990.
- [151] Frank Smadja and Kathleen McKeown. Automatically extracting and representing collocations for language generation. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, Pittsburgh, Pa., June 1990.
- [152] Steven Feiner and Kathleen McKeown. Coordinating text and graphics in explanation generation. In *Proceedings of AAAI*, Boston, Mass., August 1990.
- [153] Steven Feiner and Kathleen McKeown. Generating coordinated multimedia explanations. In *Proceedings of the IEEE Conference on AI Applications*, Santa Barbara, CA., March 1990.
- [154] Galina Datskovsky Moerdler and Kathleen McKeown. Beyond semantic ambiguity. In *Proceedings of AAAI-88*, Saint Paul, Minn., August 1988.
- [155] Galina Datskovsky Moerdler, Kathleen McKeown, and Robert Ensor. Building natural language interfaces for rule-based expert systems. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, August 1987.
- [156] Kathleen McKeown and Cecile Paris. Functional unification grammar revisited. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford, Ca., July 1987. Association for Computational Linguistics.
- [157] Kathleen McKeown. Communication with expert systems. In *Proceedings of FJCC 86*, Dallas, Texas, November 1986. Fall Joint Computer Conference.
- [158] Kathleen McKeown. The need for text generation. In *Proceedings of NCC 85*, Chicago, Ill., July 1985. National Computer Conference.
- [159] Kathleen McKeown, Michael Wish, and Kevin Matthews. Tailoring explanations for the user. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, Los Angeles, Ca., August 1985.
- [160] Kathleen McKeown. Natural language for expert systems: Comparisons with database systems. In *Proceedings of COLING '84*, Stanford, Ca., July 1984. Association for Computational Linguistics.
- [161] Marsha Derr and Kathleen McKeown. Using focus to generate complex and simple sentences. In *Proceedings of COLING '84*, Stanford, Ca., July 1984. Association for Computational Linguistics.
- [162] Kathleen McKeown. Natural language systems: How are they meeting human needs? In *Proceedings of ACM '83*, New York, N.Y., October 1983. Association for Computing Machinery.

- [163] Kathleen McKeown. Recursion in text and its use in language generation. In *Proceedings of the Third Annual National Conference on Artificial Intelligence*, Washington, D.C., August 1983. American Association of Artificial Intelligence.
- [164] Kathleen McKeown. Focus constraints on language generation. In *Proceedings of the Eight International Conference on Artificial Intelligence*, Karlsruhe, Germany, August 1983.
- [165] Kathleen McKeown. The text system for natural language generation: an overview. In *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics*, pages 113–20, Toronto, Ontario, Canada, June 1982. Association for Computational Linguistics.
- [166] Kathleen McKeown. Generating relevant explanations: Natural language responses to questions about database structure. In *Proceedings of the First Annual National Conference on Artificial Intelligence*, pages 306–9, Stanford, Ca., August 1980. American Association for Artificial Intelligence.
- [167] Kathleen McKeown. Paraphrasing using given and new information in a question-answer system. In *Proceedings of the 17th Annual Meeting of the Association for Computational Linguistics*, pages 67–72, La Jolla, Ca., August 1979. Association for Computational Linguistics.

## Invited Talks and Panels

Where Natural Language Processing Meets Societal Problems, Invited talk, University of Colorado at Boulder, March 2019.

Detecting Gang-Involved Escalation on Social Media Using Context, Faculty Tech Talk, SEAS, February 2019.

Where Natural Language Processing Meets Societal Problems, Machine Learning Seminar speaker, Duke University, November 2018.

Where Natural Language Processing Meets Societal Problems, ROCLING 2018, October 2018, Taiwan.

”Natural Language for Societal Challenges,” keynote, Statistics and Data Science Conference, MIT, April 2018.

”Natural Language: Challenges and Approaches”, Moodys Investments, Feb.2018.

”Violence on the Digital Street: Detecting Aggression and Loss in Social Media Posts using Qualitative Analysis and Machine Learning” (joint talk with Desmond Patton), Computational Social Science Center, Harvard University, Feb. 2018.

”Machine Learning from Text: Applications, Challenges in Machine Generation of Automated Products from Multi-Source Data: A Workshop”, National Academies of Science, Engineering and Medicine, August 2017.

”Natural Language for Societal Challenges, Cognitive AI Meetup,” Google, New York, NY July 2017.

”At the Intersection of Language and Data Science”, talk for Huawei visit, Columbia University, NY NY, June 2017.

Panel member, ”The Machines Are Coming: Automating the Media (panel)”, Machines and Media Conference, NYC Media Lab, Bloomberg, April 2017.

At the Intersection of Data Science and Language, Elsevier Management Conference, May 2016.

At the Intersection of Data Science and Language, UBS, New York, NY August 2016.

At the Intersection of Data Science and Language, Keynote Speaker, International Semantic Web Conference, Kobe, Japan 2016.

At the Intersection of Data Science and Language, University of Pennsylvania, Grace Hopper Distinguished Lecture, Philadelphia, Pa., Nov. 2016.

At the Intersection of Data Science and Language, AIG Data Science Seminar, New York, NY Nov. 2016.

At the Intersection of Data Science and Language, Women in Data Science Day, SAS, New York, NY, Feb 2017.

Panel member: The Art of AI, Lincoln Center, June 2016.

Panel member: Data Science Education, Microsoft Research Faculty Summit, July 2016.

Panel member: Data Science, New York Academy of Science, September 2016.

Making Expert Material Accessible to Non-Experts, DARPA SMILE Workshop, Washington DC, October 2016.

"Uphill Battles in Language Generation: Sentence Planning and Lexical Choice," invited talk, Workshop on Uphill Battles in Natural Language Processing, EMNLP, November 2016.

"Detecting Influence and Aggression in Social Media", DARPA ISAT Workshop, Endicott, MA, Feb. 2017.

Panel Member, "Education for Data Science," Data Science Institute Launch, University of Massachusetts at Amherst, Amherst, Mass., April 2015.

"At the Intersection of Language and Data Science," Institute of Advanced Studies Distinguished Lecture, Hong Kong University of Science and Technology, Hong Kong, March 2015.

"Big Data and Society: What it means for your Business or Learning", Future of Learning Forum Mind, Meaning and Motivation, Columbia Business School Executive Education, New York, NY, March 2015.

"At the Intersection of Language and Data Science," Colloquium Series, Data Science Institute, Columbia University, NY February 2015.

"At the Intersection of Language and Data Science," Distinguished lecture, Launch of the Centre for Doctoral Training in Data Science, Univ of Edinburgh, Edinburgh, Scotland, November 2014.

"At the Intersection of Language, Science and Data Science," Keynote, 10th IEEE International Conference on eScience, Guarujá, Sao Paulo, Brazil, October 2014.

"Celebrating Successful Partnerships: Applied Sciences NYC Project", Panel, NACME 40th Annual Meeting, October 2014.

"Mentoring for Postdoctoral Fellows", Panel, ASCENT (Advancing Computer Science Careers through Enhanced Networking and Training) kick-off meeting, September 2014.

"Natural Language Applications From News to Novels," Microsoft Research, Seattle, Wa. March 2014.

"Natural Language Applications From News to Novels," Keynote Speech, North American Association for Computational Linguistics, Atlanta, Georgia, June 2013.

"NYC 2030: Achieving the City's Future", Panel, Columbia Business School Club of New York, April 2013.

"New York Universities", Panel, Engineering Deans Institute, American Society for Engineering Education, April 2013.

"Out of the Cloud: the Promise of Big Data," Panel Discussion with moderator Columbia Univ President Lee Bollinger, April 2013.

"Columbia's Institute for Data Sciences and Engineering," Google Tech Talk, Google, Menlo Park, Ca., March 2013.

"Natural Language Applications Across Genres: From News to Novels," CRA Distinguished Lecture Series, Department of Computer Science, University of Illinois, May 2011.

"Natural Language Applications Across Genres: From News to Novels," Distinguished speaker, Study of the Applied Language Technology and Society, Rutgers University, March 2011.

"Query-focused Summarization Using Text-to-Text Generation: When Information Comes from Multilingual Sources", Department of Computer Science Lecture Series, Johns Hopkins University, February 2011.

"Summarization and Question Answering Using Text-to-Text Generation", Brown University, November 2007.

"Text summarization: News and Beyond," CICLING invited speaker, February 2007.

"Evaluation of Text Generation Systems: Nightmares and Pleasant Surprises, INLG, July 2006, Australia.

"Text summarization: News and Beyond," The State University of New York at Stony Brook, Department of Computer Science Distinguished Lecture Series, January 2005.

"Text summarization: News and Beyond," Microsoft, August 2004.

"Generating Natural Language Summaries and its use in Tracking and Summarizing News on a Daily Basis," IBM, August 2002.

"Generating Natural Language Summaries and its use in Tracking and Summarizing News on a Daily Basis," ISI, November 2002.

"Personalized Search and Summarization over Multimedia Information," University of Michigan, Computer Science Lecture Series, October 2001.

"Personalized Search and Summarization over Multimedia Information," University of Pittsburgh, Distinguished Lecture Series, Department of Computer Science, October 2001.

"Language Generation and Summarization: Moving Away from Sentence Extraction," Computer Science Lecture Series, November 2000.

"What Can Computers do with Language and Other Media?," NSF 50th Anniversary Symposium, Columbia University, December 2000.

"Reducing Information Overload: Text Summarization and Personalized Digital Libraries," Microsoft Research Labs, October 1999.

"Generating Natural Language Summaries: What is Possible Now?," Computer Science Colloquium, Cornell University, January 1999.

"A Patient Care Digital Library: Personalized Search and Summarization over Multimedia Information," panel, IEEE Advances in Digital Libraries Conference, Baltimore, Md., May 1999.

"Generating Multimedia Briefings," Southern Methodist University Distinguished Lecture Series, Dallas, Tx., April 1999.

"Generating Natural Language Summaries: What is Possible Now?," Cornell University, Ithaca, NY, January 1999.

"Concept to Speech Generation for Multimedia Briefings," MIT, Cambridge, Ma., November 1998.

"Generating Summaries of Medical Information," Department of Medical Informatics, Columbia University, NY, NY, November 1998.

"Generating Multimedia Briefings," ISI/USC, Marina del Rey, November 1998.

"Generating Multimedia Briefings," IJCAI invited talk, August 1997.

"Generating Multimedia Briefings," AT&T, Florham Park, September 1997.

"Generating Multimedia Briefings," Harvard University, Cambridge, Mass., November 1997.

"Generating Multimedia Briefings," Univ. of Toronto, Toronto, Can., December 1997.

"Generating Multimedia Briefings," Univ. of Pennsylvania, Philadelphia, Pa., January 1998.

"Generating Multimedia Briefings," Lucent Bell Laboratory, Holmdel, NJ, February 1998.

"Automated Generation of Multimedia Briefings," ONR workshop, Woods Hole, Mass., Sept. 1997.

"Generating Natural Language Summaries for Digital Libraries," invited talk, Shanghai Univ., Shanghai, China, May 1996.

"Generating Natural Language Summaries for Digital Libraries," IBM, Hawthorne, N.Y., January 1996.

"Generating Concise Natural Language Summaries," Linguistics Colloquium, Univ. of Pennsylvania, March 1996.

"Generating Concise Natural Language Summaries," NYNEX, White Plains, N.Y., April 1995.

"Generating Concise Natural Language Summaries," Johns Hopkins University colloquium, Baltimore, Md., February 1995.

"Getting Computers to Interact with their Users in English," Columbia Univ. Deans Day, Columbia Univ., April 1994.

"The JANUS Digital Library: An Overview of Research," Columbia Univ. Partners Meeting, Columbia Univ., April 1994.

"A User Interface for the JANUS Digital Library," Rutgers Univ., April 1994.

"Automated Generation of 3D Multimedia for Healthcare User Interfaces," ARPA Human Computer Interaction Workshop, March 1994.

"Language Generation for Multimedia Explanation," Univ. of Toronto, Feb. 1994.

"Language Generation for Multimedia Explanation," invited speaker, PACLING-93, Vancouver, B.C., April 1993.

"Automatic Documentation Generation in PLANDoc," NYNEX colloquium series, NYNEX, White Plains, N.Y., February 1992. (joint talk with Karen Kukich, Bellcore)

"Extracting Constraints on Word Usage from Text Corpora," Naval Research Labs Colloquium Series, Washington D.C., November 1992.

"Constructing Lexicon Modules through Analysis of Large Text Corpora," invited talk, DARPA TIPSTER Workshop on Text and Information Retrieval, San Diego, Ca., September 1992.

"Issues for Women in Academia," Women Speaker Series, UNIX Systems Laboratory, Summit, N.J., September 1992.

"Issues for NL: An Evening of Video Entertainment," Presidential Address, ACL Banquet after dinner talk, 30th Annual Meeting of the Association for Computational Linguistics, Univ. of Delaware, Newark, Del., July 1992.

"Language Generation for Equipment Maintenance and Repair: Coordinating with Graphics," Univ. of Pennsylvania, Feb., 1990.

"Language Generation for Equipment Maintenance and Repair: Coordinating with Graphics," AT&T Bell Laboratories, Jan., 1990.

"Language Generation at Columbia: the Radio Project," GE, Schnectedy, July 1989.

"Language Generation at Columbia: the Radio Project," invited speaker, European Workshop on Language Generation, Univ. of Edinburgh, Edinburgh, Scotland, March 1989.

“Highlighting User Related Advice,” AAAI Workshop on Explanation, AAAI, St. Paul, Minn., August 1988.

“Comparison of Surface Language Generators: A Case Study in Choice of Connectives,” 4th International Language Generation Workshop, Santa Catalina Island, Ca., July 1988.

“Getting Computers to Respond in English,” invited talk, Women in Science Lecture Series, Brookhaven National Laboratory, Brookhaven, N.Y., May 1988.

“Individual Differences in Human Computer Interaction,” panel member, CHI 88, Washington D.C., May 1988.

“Constraints on Text Generation,” invited talk, University of Waterloo, Waterloo, Canada, January 1987.

“Constraints on Text Generation,” invited talk, Texas Instruments, Dallas, Texas, November 1986.

“Natural Language Generation and User Modelling,” tutorial given at COLING-86, Bonn, Germany, August 1986.

“Natural Language Generation,” tutorial given at the 24th Annual Meeting of the Association for Computational Linguistics, Columbia University, New York, N.Y., June 1986.

“Tailoring Explanations for the User,” invited talk, University of Pennsylvania, Philadelphia, Pa., March 1985.

“Text Generation: What, when, and how,” invited talk, Cognitive Studies Series, Princeton University, Princeton, N.J., February 1985.

“Tailoring Explanations for the User,” ONR Contractor’s meeting, Georgia Technical Institute, Atlanta, Ga., January 1985.

Workshop on Language Generation, invited speaker, Stanford Univ., Stanford, Ca., Doug Appelt (organizer), July 1984.

Panel member of “Natural Language and Databases Again,” COLING 84, Stanford Univ., Stanford, Ca., July 1984.

“User-Oriented Language Generation,” invited talk, ACM SIGART local chapter, New York, N.Y., May 1984.

Workshop on Knowledge Representation and Explanation, invited participant, sponsored by ISI and BBN, Santa Barbara, Ca., October 1983.

“Natural Language Systems: How are they meeting human needs?”, invited talk, ACM Annual Conference, New York, N.Y., October 1983.

“Generating Natural Language Text in Response to Questions about Database Structure”, invited talk, American Society of Information Science Annual Conference, Doctoral Forum, Washington D.C., October 1983.

“Generating Natural Language Text in Response to Questions about Database Structure”, invited talk, IBM Systems Research Institute, New York, N.Y., Oct. 1982

Panel member of “Stratgies for Explanation” panel and “Generation” panel, Explanation Workshop, sponsored by ISI and ONR, June 1982.

Panel member of “Text Generation” panel, Applied Computational Linguistics in Perspective Workshop, sponsored by ONR, June 1981, results appeared in April-June 1982 issue of the @i(American Journal of Computational Linguistics).

## Classes Taught

**Natural Language Processing**, COMS4705, Department of Computer Science, Columbia University, Spring 1983, 1984, 1985 (50 students), 1986 (50 students), 1987 (80 students), 1988 (35 students), 1989 (30 students), 1990 (15 students), 1992 (15 students), Fall 1993 (14 students), Fall 1995 (25 students), Fall 1996 (15 students), Fall 1998 (40 students), Fall 2008 (52 students), Fall 2010 (29 students), **Fall 2017** (185 students).

**Artificial Intelligence**, COMS4701, Department of Computer Science, Columbia University, Spring 2004 (55 students), Fall 2004 (39 students), Fall 2006 (45 students), Fall 2007 (68 students)

**Natural Language Processing for the Web**, ComsE6998 Section 3, Department of Computer Science, Columbia University, Spring 2008 (21 students), Spring 2010 (16 students)

**Introduction to Computers**, ComsW1001, Department of Computer Science, Columbia University, Fall 1994 (40 students)

**The Lexicon and Machine Translation**, ComsE6998 Section 37, Department of Computer Science, Columbia University, Fall 1991 (9 students)

**Independent Study in Natural Language Processing**, COMSW6901 Section 37, Department of Computer Science, Columbia University, Fall 2017 (4 MS students), Spring 2018 (7 MS students) and every semester since 1982.

**Independent Study in Natural Language Processing**, COMSW4901 Section 37, Department of Computer Science, Columbia University, Fall 2017 (1 undergraduate student), Spring 2018 (5 undergraduate students) and every semester since 1982.

**Introduction to Programming**, COMS1003, Department of Computer Science, Columbia University, Fall 1982, Fall 1984

## Departmental Service

Faculty Recruiting Committee, 2017-2018.

Advisor for MS students, NLP track, 2015-2018.

Nominations committee, 2016-2018.



Advisor for Barnard Undergraduate Students majors, 1994-00, 2011-2014.

Co-organizer, Department of Computer Science Retreat, Fall 2010 (joint with Eitan Grinspun).

Task Force on New Directions in Curriculum, Chair, 2006-2007.

Academic Committee 2004-2005, 2006-2010.

Strategic Planning Committee 2002-2005, 2009-2010, 2014-2015.

Co-Chair, CS 25th Anniversary, 2004.

Senior Faculty Recruiting 1999-00.

Internal Departmental Review Committee, 1999.

Member, Faculty Recruiting Committee, 1994-1995.

Chair, Faculty Recruiting Committee, 1991-1992, 1992-1993, 1993-1994.

Member, Chair Search Committee, 1993-1994.

Contact, discussion of combined chair in Arts and Technology, 1993-1994.

Member, Academic Committee, 1993.

Member, Faculty Recruiting Committee, 1989, 1990.

Member, PhD Committee, 1989, 1990, 1991, 1992.

PhD Admissions committee, 1989-90.

Scribe, community service records for department, 1989-90.

Chair, Committee Committee, 1987-88.

Appeals Committee, 1988.

Qualifier Reform Committee, Spring 1987.

Departmental Representative to Columbia College, 1985-86, 1986-87.

Departmental Representative to General Studies, Spring 1987.

Chairperson, AI qualifier committee, 1982 - present.

Chairperson, Committee on BA curriculum reform, 1983-84.

Participated in writing of NSF/STC proposal, fall 1987.

Participating in writing of DARPA renewal contract for March 1987.

## University Service

Member, Provost's Committee on the University Libraries, 2015-2018.

Member, Just Societies Taskforce, 2016-2017.

Member, President's University Task Force on Precision Medicine, 2014-2015.

Member, Provost's Task Force on a Standing Committee for Tenure Review, 2010-2011.

Member, Selection Committee for Columbia Graduate Student Teaching Award, 2007-2011.

Outside reader, thesis defense committee, Sophie Guitte, Spring 2011.

Member, University Task Force on Diversity in Science and Engineering, 2005-2007.

Member, Committee on Undergraduate Program in Bio-Informatics, 2003-2005.

Member, Provost's Committee on Intellectual Property Rights, 2000.

Member, University Task Force on the Libraries, 1998-1999, 1999-2000.

Chair, CRIA Research Committee, 1995-1997.

Member, Dean Search Committee, 1994-1995.

Faculty Affairs, 1993-1994.

Task Force on Sexual Assault, 1994.

Member, Columbia Univ. Senate Commission on the Status of Women, 1994-96.

Chair, Columbia Univ. Senate Commission on the Status of Women, 1991-92, 1992-93, 1993-94.

Senator, Univ. Senate, 1992-1994

Speaker and Facilitator, Issues for Women, Columbia Univ. Forum on Diversity, Freshman Orientation, September 1992.

Panelist, Career Night, Society of Women Engineers, October 1992.

Spoke at Pre-freshman Orientation for Engineering Women, Society of Women Engineers, April 1988, April 1989.

Thesis committee: Susan Alexander, Teachers College, spring 1994; Pamela Wolf, Teachers College, spring 1992; Peter Lochiel Glidden, Teachers College, spring 1987; Beth King, Teachers College, spring 1990.

Organized AI Day, Spring 1987, attended by many departments of the University.

Arranged for the 24th Annual Meeting of the Association for Computational Linguistics to be held at Columbia University in June 1986. This is a national organization and a well-known conference, with approximately 500 attending in June. Allowed Columbia University students and associates to easily attend conference. In addition, gave national recognition to Columbia as a sponsor of such a conference.

Organized Cognitive Science Seminar Series jointly with Prof. Jim Corter, held jointly with Dept. of Psychology, Dept. of Philosophy, Dept. of Linguistics (Barnard), Dept. of Psychology (Teachers College), Spring 1985.

Outside reader on the defense committee of Peter Ludlow, Dept. of Philosophy, Columbia College, Spring 1985.

Outside reader on defense committee of Elizabeth M. Reiss, November 1983, Teachers College.

Outside reader on defense committee of Thomas Eblen, Teachers College, 1984.