

# Graeme Hirst

## Basic information

---

**St George campus address:** Department of Computer Science, 40 St George Street (room 4283), University of Toronto, Toronto, Ontario, Canada M5S 2E4.  
**Email:** gh@cs.toronto.edu; **Web:** <http://graemehirst.com>

## Education

---

- 1979–83:** Department of Computer Science, Brown University, Providence, Rhode Island.
- PhD degree; Thesis: *Semantic interpretation against ambiguity*. (Supervisor: Eugene Charniak.)
- 1975–79:** Department of Engineering Physics, Research School of Physical Sciences, Australian National University, Canberra, A.C.T., Australia *and* Department of Computer Science, University of British Columbia, Vancouver, B.C., Canada.
- MSc degree (ANU); Thesis: *Anaphora in natural language understanding*. Thesis written partly at ANU (supervisors: Stephen Kaneff and Iain Macleod) and partly at UBC (supervisor: Richard Rosenberg).
- 1970–73:** Monash University, Clayton, Victoria, Australia.
- Majors in Information Science and Pure Mathematics; minor in Psychology. BSc degree with First-Class Honours in Information Science, 1974. BA degree, 1973.

## Other qualifications

---

**2007** University of Toronto, Certificate of Continuing Studies in Dispute Resolution. [7-course sequence, 196 hours.]

## Employment

---

- 1984– :** University of Toronto: Department of Computer Science; also (from January 1986) University of Toronto, Scarborough: Department of Computer and Mathematical Sciences (formerly, Division of Mathematical Sciences; formerly, Division of Physical Sciences).
- Professor Emeritus (2020– ); Professor (1995–2020); Associate Professor [with tenure] (1990–95); Assistant Professor (1984–90). Member, School of Graduate Studies (from November 1987).
  - Associate Chair, Graduate Studies, 2019– .
  - Gastprofessor, Universität Mannheim, Fakultät für Wirtschafts-informatik und -mathematik (Data and Web Science Group), and Gastwissenschaftler, Universität Heidelberg, Institut für Computerlinguistik, March 2019.
  - Faculty affiliate (2018–20), Vector Institute.
  - Visiting Professor (2005–06), University of Colorado at Boulder, Center for Spoken Language Research.
  - Adjunct Professor (1995–98), University of Waterloo, Department of Computer Science.
  - Visiting Professor (1990–91), University of Rochester, Department of Computer Science.

## **Awards and honours**

---

Distinguished Service Award, Association for Computational Linguistics, 2020.

Fellow of the Association for Computational Linguistics, 2019.

Lifetime Achievement Award, Canadian Artificial Intelligence Association, 2017.

Distinguished Service Award and Fellow, Canadian Society for Computational Studies of Intelligence [now the Canadian Artificial Intelligence Association], 2008.

Awards for excellence in teaching, Computer Science Students' Union, University of Toronto, September 1988 and October 1994.

Publisher's Prize, Best paper by student or recent student, National Conference on Artificial Intelligence (AAAI-84), Austin, Texas, August 1984.

## **Research interests**

---

Computational linguistics; natural language processing; machine translation; knowledge representation; cognitive science; computer models of language and cognition; semantics; lexical semantics; pragmatics; linguistic style; artificial intelligence; applications of computational linguistics and natural language processing in writing, in health and medicine, and in the humanities.

## **Research grants and contracts**

---

Ontario Institute for Cancer Research and Cancer Care Ontario Health Services Research Network, 2019–2021, "Building an artificial intelligent system to enhance online support groups in cancer." \$299,923 [Principal investigators: Yvonne Leung and Mary Jane Epstein; co-investigators: Graeme Hirst and 10 others].

Social Sciences and Humanities Research Council of Canada, Insight Grant, 2018–2022, "Text as image, image as text: charter integrity and topic modelling". \$286,050 [Principal investigator: Michael Gervers; co-investigator: Graeme Hirst.]

Connaught Fund, Global Challenge Award, 2017–2019, "Innovations to close the global mortality divide". \$250,000 [Lead applicant: Prabhat Jha; Co-applicants: Marie-Josée Fortin, Graeme Hirst, Anita McGahan, John McLaughlin, Erica di Ruggiero.]

Canadian Institutes of Health Research, Collaborative Health Research Projects grant, 2017–2019, "Automatic remote screening of speech features associated with Alzheimer's disease", \$437,664 [Lead applicant: Frank Rudzicz; Principal applicants: Graeme Hirst, Regina Jokel; Co-applicants: Carmela Tartaglia, Khai Truong; Knowledge users: Sandra Black, Sheila Dunn, Alex Mihailidis.]

Canadian Cancer Society, 2017–2019, "Can online physician ratings be used to provide meaningful and actionable patient feedback? A mixed methods feasibility study", \$188,895 [Primary applicant: Jill Tinmouth, Sunnybrook Research Institute; other co-applicants: Nancy Baxter, Caroline Bravo, Caitlin Crystoja, Graeme Hirst, Diego Llovet, David Urbach.]

Google, Faculty Research Award, 2016–2017, "Natural language processing to determine causes of death in developing countries", US\$41,500 [in collaboration with Prabhat Jha, Centre for Global Health Research].

Riksbankens Jubileumsfond (Sweden), 2016-01-01 to 2019-12-31, “Språkliga och extra-lingvistiska parametrar för tidig upptäckt av kognitiv svikt [Linguistic and extra-linguistic parameters for early detection of cognitive impairment]”, SEK 10,460,000 (approx \$1,660,000). [Primary applicant: Dimitrios Kokkinakis, University of Gothenburg; other co-applicants: Simon Dobnik, Christos Koniaris, Arto Nordlund (University of Gothenburg).]

NSERC, Engage grant, 2015-03-01 to 2015-08-31, “Automatically categorizing forum posts”, \$24,904.

Vetenskapsrådet (Swedish Research Council), 2015-01-01 to 2018-12-31, “The advantage of country comparisons — Towards a new method for estimating language effects in cross-cultural surveys”, SEK 9,200,000 (approx \$1,400,000). [Primary applicant: Stefan Dahlberg, University of Gothenburg; other co-applicants: Sören Holmberg (University of Gothenburg), Magnus Sahlgren (Gavagai AB, Stockholm), Jonas Linde (University of Bergen).]

NSERC Discovery Grant, 2014–2020, “Applied computational models of discourse, argument, and text”, \$54,000 p.a. × 6 years.

NSERC, Discovery Frontiers — Digging Into Data, 2014–2015, “Digging into linked parliamentary data”, \$39,250 + \$41,000. A component of a multi-agency multinational award totalling approximately \$580,000 to 6 PIs.

U.S. DARPA, 2012–2013, “Explorations on semantic equivalence”, US\$220,058.

NSERC, Research Tools and Instruments Grant, 2010, “Massively parallel computing in natural language technology”, \$140,843 [co-applicant with Gerald Penn and Suzanne Stevenson].

Google Research Award, 2009, “Natural language processing methods for the early detection of Alzheimer’s disease in patients’ on-line writing”, US\$36,000 [with Ian Lancashire and Regina Jokel].

NSERC, Discovery Grant, 2009–2014, “Nuances of meaning, paraphrase, and argument identification in applications of natural language processing”, \$50,000 p.a. × 5 years.

NSERC, Collaborative Research and Development Grant, 2008–09, 2009–10, “Towards articulatory-based adaptation in recognition of dysarthric speech”, \$47,713 + 129,820 = 177,533 [with Fraser Shein, Gerald Penn, Pascal van Lieshout, and Willy Wong, co-applicants].

Bell University Laboratories, 2007–2008, “Towards articulatory-based adaptation in recognition of dysarthric speech”, \$75,000 [with Fraser Shein, Gerald Penn, and Pascal van Lieshout, co-applicants].

NSERC, Research Tools and Instruments Grant, 2005, “Compute server for research in computational linguistics and natural language Processing”, \$73,011 [with Suzanne Stevenson and Gerald Penn, co-applicants].

NSERC, Discovery Grant, 2004–2009, “Paraphrase and semantic distance in applications of natural language processing”, \$59,000 p.a. × 5 years.

SSHRC, Major Collaborative Research Initiative, 2002–2006, “International Research on Permanent Authentic Records in Electronic Systems 2: Experiential, Interactive, and Dynamic Records”, \$2,499,400. [P.I., Luciana Duranti, UBC; 24 co-investigators.]

Bell University Laboratories, 2001–2003, \$199,640 p.a. × 2 years [P.I., John Mylopoulos; 5 other co-investigators]; renewed 2003–04, \$215,000.

Bell University Laboratories, 2001–2003, \$114,582 [P.I., monica schraefel].

Bell University Laboratories, 2000–2001, \$86,250 [P.I., monica schraefel].

NSERC, Discovery Grant, 2000–2004, “The determination of nuances of meaning and style and of semantic distance for applications of computational linguistics”, \$54,000 p.a. × 4 years.

CITO, “Applications of computational linguistics to communication aids for people with disabilities”, 2000–2002, \$100,500 [co-investigator, Fraser Shein].

Bell Canada University Laboratories, 1999–2000, \$55,200.

Bell Canada University Laboratories [with Robert Wright, P.I., Jutta Treviranus, and Mark Chignell], 1999–2000, \$50,000.

CITO, “Computational measures of semantic distance and the intelligent disambiguation, summarization, filtering, linking, indexing, retrieval, and searching of documents and text”, 1998–2000, \$100,000, *declined*.

NSERC, research grant, “Representation and use of nuances of meaning and style in applications of computational linguistics”, 1996–2000, \$46,500 + 46,500 + 51,150 + 53,708.

Technology Ontario [with Chrysanne DiMarco, P.I.], 1995–98, \$666,667.

NSERC, research grant, “Natural language understanding: Representation and repair”, 1992–96, \$35,000 p.a. × 4 years.

NSERC, strategic grant [with Ronald Baecker, Marilyn Mantei, and Marlene Scardamalia], 1992–95, \$120,000 p.a. × 3 years.

ITRC [with Chrysanne DiMarco], 1991–93, \$90,000

NSERC, operating grant, 1989–92: \$30,590 p.a. × 3 years.

ITRC [with Stephen Regoczei, through John Mylopoulos, P.I.] 1989–91, \$21,000.

Visiting fellowship, April 1988, £2509 from SERC, for three-month visit to the Department of Artificial Intelligence, University of Edinburgh.

NSERC, operating grant, 1986–89: \$23,732 p.a. × 3 years.

NSERC, operating grant, 1985–86: \$24,773.

IBM Canada Ltd, contract, 1985–86: \$30,300.

NSERC, strategic grant and strategic equipment grant [with John Mylopoulos, John Tsotsos, and Hector Levesque], 1984–87, \$1,495,000.

NSERC, operating grant, 1984–85: \$15,030.

Connaught Fund, University of Toronto, 1984: \$24,000.

## Patents

---

———; Jokel, Regina; Lancashire, Ian; and Le, Xuan. Method and system of longitudinal detection of dementia though lexical and syntactic changes in writing. U.S. patent US 9514281 B2, 6 December 2016; Canadian patent CA2776140C, 19 June 2018.

## Publications

---

Career publication count	
Scholarly books (authored)	2
Scholarly books (edited)	1
Chapters in books	29
Papers in refereed journals	44
Papers in refereed conferences	100
Major invited contributions	20
Other conference abstracts / posters / contributions	68
Other publications	14

## Refereed papers

Vishnubhotla, Krishnapriya; ———; and Rudzicz, Frank. “An evaluation of disentangled representation learning for texts,” *Findings of the Association for Computational Linguistics*, 2021.

Leung, Yvonne; Wouterloot, Elise; Adikari, Achini; ———; de Silva, Daswin; Wong, Jiahui; Bender, Jacqueline; Gancarz, Mathew; Gratzner, David; Alahakoon, Daminda; Esplen, Mary Jane. “Natural language processing-based virtual cofacilitator for online cancer support groups: Protocol for an algorithm development and validation study.” *JMIR Research Protocols*, 10(1), January 2021, e21453.

Hulpuş, Ioana; Kobbe, Jonathan; Stuckenschmidt, Heiner; and ———. “Knowledge graphs meet moral values.” *Proceedings, 9th Joint Conference on Lexical and Computational Semantics (\*SEM-2020)*, December 2020, 71–80.

Paul, Debjit; Optiz, Juri; Becker, Maria; Kobbe, Jonathan; ———; and Frank, Anette. “Argumentative relation classification with background knowledge.” In: Prakken, Henry; Bistarelli, Stefano; Santini, Francesco; and Taticchi, Carlo (editors), *Computational Models of Argument; Proceedings of COMMA-2020*, IOP Press (Frontiers in Artificial Intelligence and Applications, volume 326), September 2020, 319–330.

Abdalla, Mohamed; Abdalla, Moustafa; ———; and Rudzicz, Frank. “Exploring the privacy-preserving properties of word embeddings: Algorithmic validation.” *Journal of Medical Internet Research*, 22(7), July 2020, e18055.

Abdalla, Mohamed; Abdalla, Moustafa; Rudzicz, Frank; and ———. “Using word embeddings to improve the privacy of clinical notes.” *Journal of the American Medical Informatics Association*, 27(6), 2020, 901–907.

Xie, Jing Yi; Pinto, Renato Ferreira Jr; ———; and Xu, Yang. “Text-based inference of moral sentiment change.” *Proceedings, Conference on Empirical Methods in Natural Language Processing*, Hong Kong, November 2019, 4653–4662.

Hulpuş, Ioana; Kobbe, Jonathan; Becker, Maria; Optiz, Juri; ———; Meilicke, Christian; Nastase, Vivi; Stuckenschmidt, Heiner; and Frank, Anette. “Towards explaining natural language arguments with background knowledge.” *Joint Proceedings of the 6th International Workshop on Dataset Profiling and Search and the 1st Workshop on Semantic Explainability*, Auckland, 27 October 2019, 62–77. [CEUR Workshop Proceedings, volume 2465.]

Yan, Zhaodong; Jeblee, Serena; ———. “Can character embeddings improve cause-of-death classification for verbal autopsy narratives?” *Proceedings, BioNLP Workshop*, Florence, August 2019, 234–239.

Ethayarajh, Kawin; Duvenaud, David; and ———. “Understanding undesirable word embedding associations.” *Proceedings, 57th Annual Meeting of the Association for Computational Linguistics*, Florence, July 2019, 1696–1705.

Ethayarajh, Kawin; Duvenaud, David; and ———. “Towards understanding linear word analogies.” *Proceedings, 57th Annual Meeting of the Association for Computational Linguistics*, Florence, July 2019, 3253–3262.

Jeblee, Serena; Gomes, Mireille; Jha, Prabhat; Rudzicz, Frank; and ———. “Automatically determining cause of death from verbal autopsy narratives.” *BMC Medical Informatics and Decision Making*, 2019:127, 9 July 2019.

Vishnubhotla, Krishnapriya; Hammond, Adam; and ———. “Are fictional voices distinguishable? Classifying character voices in modern drama.” *Proceedings, 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Minneapolis, June 2019, 29–34.

Abdalla, Mohamed; Sahlgren, Magnus; and ———. “Enriching word embeddings with a regressor instead of labeled corpora.” *Proceedings, 33rd AAAI Conference on Artificial Intelligence (AAAI-19)*, Honolulu, January 2019, 6188–6195.

Jeblee, Serena and ———. “Listwise temporal ordering of events in clinical notes.” *Proceedings, LOUHI 2018: The Ninth International Workshop on Health Text Mining and Information Analysis*, Brussels, October 2018, 177–182.

Naderi, Nona and ———. “Using context to identify the language of face-saving.” *Proceedings, 5th Workshop on Argument Mining*, Brussels, November 2018, 111–120.

Naderi, Nona and ———. “Automated fact-checking of claims in argumentative parliamentary debates.” *Proceedings of the First Workshop on Fact Extraction and Verification (FEVER)*, Brussels, November 2018, 60–65.

Jeblee, Serena; Budhkar, Akshay; Milić, Saša; Pinto, Jeff; Pou-Prom, Chloé; Vishnubhotla, Krishnapriya; ———; and Rudzicz, Frank. “Toronto CL at the CLEF 2018 eHealth Challenge Task 1”. *CLEF 2018 Evaluation Labs and Workshop: Online Working Notes*, CEUR-WS, September 2018.

Jeblee, Serena; Gomes, Mireille; and ———. “Multi-task learning for interpretable cause-of-death classification using key phrase prediction.” *Proceedings, BioNLP 2018*, Melbourne, July 2018, 12–17.

Marinho, Vanessa Q.; ———; and Amancio, Diego R. “Labelled network motifs reveal stylistic subtleties in written texts.” *Journal of Complex Networks*, 6(4), 1 August 2018, 620–638 [published online 10 October 2017].

Tricco, Andrea; Zarin, Wasifa; Lillie, Erin; Jeblee, Serena; Warren, Rachel; Khan, Paul A.; Robson, Reid, Pham, Ba’; ———; Straus, Sharon E. “Utility of social media and crowd-intelligence data for pharmacovigilance: A scoping review.” *BMC Medical Informatics and Decision Making*, 18:38, 14 June 2018.

Naderi, Nona and ———. “Automatically labeled data generation for classification of reputation defence strategies.” In: Fišer, Darja; Eskevich, Maria; and de Jong, Franciska (eds). *Proceedings of LREC2018 Workshop ParlaCLARIN: Creating and Using Parliamentary Corpora*, Miyazaki, Japan, May 2018, 48–54.

Abdalla, Mohamed; Rudzicz, Frank; and ———. “Rhetorical structure and Alzheimer’s disease.” *Aphasiology*, 32(1), 2018, 41–60.

Abdalla, Mohamed and ———. “Cross-lingual sentiment analysis without (good) translation.” *Proceedings, 8th International Joint Conference on Natural Language Processing (IJCNLP 2017)*, Taipei, December 2017, 506–515.

Naderi, Nona and ———. “Recognizing reputation defence strategies in critical political exchanges.” *Proceedings, 11th Conference on Recent Advances in Natural Language Processing*, Varna, September 2017, 527–535.

Naderi, Nona and ———. “Classifying frames at the sentence level in news.” *Proceedings, 11th Conference on Recent Advances in Natural Language Processing*, Varna, September 2017, 536–542.

Wachsmuth, Henning; Naderi, Nona; Habernal, Ivan; Hou, Yufang; ———; Gurevych, Iryna; and Stein, Benno. “Argumentation quality assessment: Theory vs. practice.” *55th Annual Meeting of the Association for Computational Linguistics*, Vancouver, August 2017, volume 2, 250–255.

Wachsmuth, Henning; Naderi, Nona; Hou, Yufang; Bilu, Yonatan; Prabhakaran, Vinod; Alberdingk Thijm, Timothy; ———; and Stein, Benno. “Computational argumentation quality assessment in natural language.” *15th Conference of the European Chapter of the Association for Computational Linguistics*, Valencia, April 2017, 176–187.

Beelen, Kaspar; Alberdingk Thijm, Timothy; Cochrane, Christopher; Halvemaan, Kees; ———; Kimmins, Michael; Lijbrink, Sander; Marx, Maarten; Naderi, Nona; Rheault, Ludovic; Polyanovsky, Roman; Whyte, Tanya. “Digitization of the Canadian parliamentary debates”. *Canadian Journal of Political Science*, 50(3), September 2017, 849–864 [published online, 2017-01-18].

Naderi, Nona and ———. “Argumentation mining in parliamentary discourse.” In: Baldoni, Matteo; Baroglio, Cristina; Bex, Floris; Grasso, Floriana; Green, Nancy; Namazi-Rad, Mohammad-Reza; Numao, Masayuki; Suarez, Merlin Teodosia (editors), *Principles and Practice of Multi-Agent Systems*, Lecture Notes in Computer Science, volume 9935, 2016, 16–25.

Rheault, Ludovic; Beelen, Kaspar; Cochrane, Christopher; and ———. “Measuring emotion in parliamentary debates with automated textual analysis.” *PLoS ONE*, 2016, 11(12): e0168843.

Marinho, Vanessa Queiroz; Amancio, Diego Raphael; and ———. “Authorship attribution via network motifs identification”. *Proceedings, 5th Brazilian Conference on Intelligent Systems (BRACIS)*, Recife, Brazil, October 2016, 355–360.

Fraser, Kathleen C.; Rudzicz, Frank; and ———. “Detecting late-life depression in Alzheimer’s disease through analysis of speech and language.” *Proceedings, 3rd Workshop on Computational Linguistics and Clinical Psychology*, San Diego, June 2016, 1–11.

Perumal, Krish and ———. “Semi-supervised and unsupervised categorization of posts in Web discussion forums using part-of-speech information and minimal features.” *Proceedings, 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, San Diego, June 2016, 100–108.

Fraser, Kathleen C. and ———. “Detecting semantic changes in Alzheimer’s disease with vector space models.” *Proceedings, Workshop on Resources and Processing of Linguistic and Extra-Linguistic Data from People with Various Forms of Cognitive/Psychiatric Impairments (Linköping Electronic Conference Proceedings vol. 128)*, Portorož, May 2016, 1–8.

Brooke, Julian; Hammond, Adam; and ———. “Using models of lexical style to quantify free indirect discourse in modernist fiction.” *Digital Scholarship in the Humanities*, 32(2), 2017, 234–250 [online advance access 2016-02-03].

Brooke, Julian; Hammond, Adam; and ———. “Distinguishing voices in *The Waste Land* using computational stylistics.” *Linguistic Issues in Language Technology*, 12(2), 2015.

Maziero, Erick Galani; ———; and Pardo, Thiago Alexandre Salgueiro. “Adaptation of discourse parsing models for Portuguese language.” *Proceedings, 4th Brazilian Conference on Intelligent Systems (BRACIS)*, Natal, Brazil, November 2015, 140–145.

Maziero, Erick Galani; ———; and Pardo, Thiago Alexandre Salgueiro. “Semi-supervised never-ending learning in discourse parsing.” *Proceedings, Recent Advances in Natural Language Processing (RANLP)*, Hissar, Bulgaria, September 2015, 436–442.

Wang, Tong; Mohamed, Abdel-rahman; and ———. “Learning lexical embeddings with syntactic and lexicographic knowledge.” *Proceedings, 53rd Annual Meeting of the Association for Computational Linguistics*, Beijing, July 2015, volume 2, 458–463.

Zhang, Muyu; Qin, Bing; Zheng, Mao; ———; and Liu, Ting. “Encoding distributional semantics into triple-based knowledge ranking for document enrichment.” *Proceedings, 53rd Annual Meeting of the Association for Computational Linguistics*, Beijing, July 2015, volume 1, 524–533.

Brooke, Julian; Hammond, Adam; Jacob, David; Tsang, Vivian; ———; and Shein, Fraser. “Building a lexicon of formulaic language for language learners.” *Proceedings, 11th Workshop on Multiword Expressions*, Denver, June 2015, 96–104.

Brooke, Julian; Hammond, Adam; and ———. “GutenTag: An NLP-driven tool for digital humanities research in the Project Gutenberg Corpus.” *Proceedings, Fourth Workshop on Computational Linguistics for Literature*, Denver, June 2015, 42–47.

Fraser, Kathleen C.; Ben-David, Naama; ———; Graham, Naida L.; and Rochon, Elizabeth. “Sentence segmentation of aphasic speech.” *Proceedings, 2015 Conference of the North American Chapter of the Association for Computational Linguistics — Human Language Technologies*, Denver, June 2015, 862–871.

Zhang, Muyu; Feng, Vanessa Wei; Qin, Bing; ———; Liu, Ting; Huang, Jingwen. “Encoding world knowledge in the evaluation of local coherence.” *Proceedings, 2015 Conference of the North American Chapter of the Association for Computational Linguistics — Human Language Technologies*, Denver, June 2015, 1087–1096.

———. “Overcoming linguistic barriers to the Multilingual Semantic Web.” In: Paul Buitelaar and Philipp Cimiano (editors), *Towards the Multilingual Semantic Web*, Springer, 2015, 1–14.

Kolhatkar, Varada and ———. “Resolving shell nouns.” *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*, Doha, Qatar, October 2014, 499–510.

Brooke, Julian and ———. “Supervised ranking of co-occurrence profiles for acquisition of continuous lexical attributes.” *Proceedings, 25th International Conference on Computational Linguistics (COLING-2014)*, Dublin, August 2014, 2172–2183.

Brooke, Julian; Tsang, Vivian; ———; and Shein, Fraser. “Unsupervised multiword segmentation of large corpora using prediction-driven decomposition of n-grams.” *Proceedings, 25th International Conference on Computational Linguistics (COLING-2014)*, Dublin, August 2014, 753–761.

Feng, Vanessa Wei; Lin, Ziheng; and ———. “The impact of deep hierarchical discourse structures in the evaluation of text coherence.” *Proceedings, 25th International Conference on Computational Linguistics (COLING-2014)*, Dublin, August 2014, 940–949.



Feng, Vanessa Wei and ———. “Patterns of local discourse coherence as a feature for authorship attribution”. *Literary and Linguistic Computing*, 29(2), 2014, 191–198 [advance publication online, 9 April 2013].

Fraser, Kathleen C.; Meltzer, Jed A.; Graham, Naida L.; Leonard, Carol; ———; Black, Sandra E.; and Rochon, Elizabeth. “Automated classification of primary progressive aphasia subtypes from narrative speech transcripts.” *Cortex*, 55, June 2014, 43–60 [advance publication online, 21 Dec 2012].

———; Riabinin, Yaroslav; Graham, Jory; Boizot-Roche, Magali; and Morris, Colin. “Text to ideology or text to party status?” In: Kaal, Bertie; Maks, E. Isa; and van Elfrinkhof, Annemarie M.E. (editors), *From Text to Political Positions: Text analysis across disciplines*, Amsterdam: John Benjamins, 2014, 93–115.

Fraser, Kathleen C.; ———; Meltzer, Jed A.; Mack, Jennifer E.; and Thompson, Cynthia K. “Using statistical parsing to detect agrammatic aphasia.” *Proceedings, BioNLP 2014 Workshop*, Baltimore, June 2014, 134–142.

Fraser, Kathleen C.; ———; Graham, Naida L.; Meltzer, Jed A.; Black, Sandra E.; and Rochon, Elizabeth. “Comparison of different feature sets for identification of variants in progressive aphasia.” *Proceedings, Workshop on Computational Linguistics and Clinical Psychology*, Baltimore, June 2014, 17–26.

Wang, Tong and ———. “Applying a naive Bayes similarity measure to word sense disambiguation.” *Proceedings, 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, June 2014, 531–537.

Feng, Vanessa Wei and ———. “A linear-time bottom-up discourse parser with constraints and post-editing.” *Proceedings, 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, June 2014, 511–521.

Kolhatkar, Varada; Zinsmeister, Heike; and ———. “Interpreting anaphoric shell nouns using cataphoric shell nouns as training data.” *Proceedings, 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, October 2013, 300–310.

Brooke, Julian and ———. “Hybrid models for lexical acquisition of correlated styles.” *Proceedings, 6th International Joint Conference on Natural Language Processing*, Nagoya, October 2013, 82–90.

Feng, Vanessa and ———. “Detecting deceptive opinions with profile compatibility.” *Proceedings, 6th International Joint Conference on Natural Language Processing*, Nagoya, October 2013, 338–346.

Mohammad, Saif; Dorr, Bonnie; ———; Turney, Peter. “Computing lexical contrast.” *Computational Linguistics*, 39(3), September 2013, 555–590.

Feng, Vanessa Wei and ———. “Authorship verification with entity coherence and other rich linguistic features.” *Proceedings, PAN 2013 Lab: Uncovering Plagiarism, Authorship and Social Software Misuse — at the CLEF 2013 Conference and Labs of the Evaluation Forum: Information Access Evaluation meets Multilinguality, Multimodality, and Visualization*, Valencia, September 2013 [refereed by extended abstract only].

Kolhatkar, Varada; Zinsmeister, Heike; and ———. “Annotating anaphoric shell nouns with their antecedents.” *Proceedings, The 7th Linguistic Annotation Workshop & Interoperability with Discourse*, Sofia, August 2013, 112–121.

Strauss, John; Martínez Peguero, Arturo; and ———. “Machine learning methods for clinical forms analysis in mental health.” *Proceedings, 14th World Congress on Medical and Health Informatics (MEDINFO 2013)*, edited by Christoph Ulrich Lehmann, Elske Ammenwerth, and Christian Nøhr, Copenhagen, August 2013, IOS Press, page 1024. [1-page refereed abstract of poster presentation.]

Cook, Paul and ———. “Automatically assessing whether a text is clichéd, with applications to literary analysis.” *Proceedings, 9th Workshop on Multiword Expressions (MWE 2013)*, Atlanta, June 2013, 52–57.

Brooke, Julian and ———. “Using other learner corpora in the 2013 NLI shared task.” *Proceedings, 8th Workshop on Innovative Use of NLP for Building Educational Applications*, Atlanta, June 2013, 188–196.

Hammond, Adam; Brooke, Julian; and ———. “A tale of two cultures: Bringing literary analysis and computational linguistics together.” *Proceedings, Second ACL Workshop on Computational Linguistics for Literature*, Atlanta, June 2013, 1–8.

Brooke, Julian; ———; and Hammond, Adam. “Clustering voices in *The Waste Land*.” *Proceedings, Second ACL Workshop on Computational Linguistics for Literature*, Atlanta, June 2013, 41–46.

Brooke, Julian and ———. “A multi-dimensional Bayesian approach to lexical style.” *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Atlanta, June 2013, 673–679.

———; Talent, Nadia; and Scharf, Sara. “Detecting semantic overlap and discovering precedents in the biodiversity research literature.” *Proceedings of the First International Workshop on Semantics for Biodiversity* (CEUR Workshop Proceedings, volume 979), Montpellier, France, 27 May 2013.

Brooke, Julian and ———. “Native language detection with ‘cheap’ learner corpora.” In: Granger, Sylviane; Gilquin, Gaëtanelle; and Meunier, Fanny (editors) *Twenty Years of Learner Corpus Research: Looking back, moving ahead*. Louvain-la-Neuve: Presses universitaires de Louvain, 2013, 37–47.

Kennedy, Alistair and ———. “Measuring semantic relatedness across languages.” *Proceedings, xLiTe: Cross-Lingual Technologies Workshop at the Neural Information Processing Systems Conference*, 2012, December, Lake Tahoe, NV.

Brooke, Julian and ———. “Robust, lexicalized native language identification.” *Proceedings, 24th International Conference on Computational Linguistics (COLING-2012)*, Mumbai, December 2012, 391–408.

Brooke, Julian and ———. “Paragraph clustering for intrinsic plagiarism detection using a stylistic vector-space model with extrinsic features.” *Proceedings, PAN 2012 Lab: Uncovering Plagiarism, Authorship and Social Software Misuse — at the CLEF 2012 Conference and Labs of the Evaluation Forum: Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics*, Rome, September 2012 [refereed by extended abstract only].

Morris, Colin and ———. “Identifying sexual predators by SVM classification with lexical and behavioral features.” *Proceedings, PAN 2012 Lab: Uncovering Plagiarism, Authorship and Social Software Misuse — at the CLEF 2012 Conference and Labs of the Evaluation Forum: Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics*, Rome, September 2012 [refereed by extended abstract only].

Kolhatkar, Varada and ———. “Resolving ‘this-issue’ anaphora.” *2012 Conference on Empirical*

*Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*, 1255–1265, July 2012, Jeju, Korea.

Feng, Vanessa Wei and ———. “Text-level discourse parsing with rich linguistic features.” *Proceedings, 50th Annual Meeting of the Association for Computational Linguistics*, Jeju, Korea, July 2012, 60–68.

Cook, Paul and ———. “Do Web-corpora from top-level domains represent national varieties of English?” *Proceedings, 11<sup>es</sup> Journées internationales d’analyse statistique des données textuelles / 11th International Conference on the Statistical Analysis of Textual Data (JADT 2012)*, Liège, June 2012, 281–293.

Brooke, Julian; Hammond, Adam; and ———. “Unsupervised stylistic segmentation of poetry with change curves and extrinsic features.” *Proceedings, Workshop on Computational Linguistics for Literature*, Montreal, June 2012, 26–35.

Brooke, Julian; Tsang, Vivian; Jacob, David; Shein, Fraser; and ———. “Building readability lexicons with unannotated corpora.” *Proceedings, Workshop on Predicting and Improving Text Readability for Target Reader Populations*, Montreal, June 2012, 33–39.

——— and Feng, Vanessa Wei. “Changes in style in authors with Alzheimer’s disease.” *English Studies* (special issue on stylometry and authorship attribution), 93(3), May 2012, 357–370.

Brooke, Julian and ———. “Measuring interlanguage: Native language identification with L1-influence metrics.” *Proceedings, 8th ELRA Conference on Language Resources and Evaluation (LREC 2012)*, Istanbul, May 2012. [Refereed by extended abstract]

Feng, Vanessa Wei and ———. “Extending the entity-based coherence model with multiple ranks.” *Proceedings, 13th Conference of the European Chapter of the Association for Computational Linguistics*, Avignon, April 2012, 315–324.

Rudzicz, Frank; ———; and van Lieshout, Pascal. “Vocal tract representation in the recognition of cerebral palsied speech.” *Journal of Speech, Language, and Hearing Research*, 55(4), August 2012, 1190–1207.

Cook, Paul and ———. “Automatic identification of words with novel but infrequent senses”. *Proceedings, 25th Pacific Asia Conference on Language, Information and Computation*, Singapore, December 2011, 265–274.

Brooke, Julian; Wang, Tong; and ———. “Predicting word clipping with latent semantic analysis”. *Proceedings, 5th International Joint Conference on Natural Language Processing*, Chiang Mai, November 2011, 1392–1396.

Wang, Tong and ———. “Refining the notions of depth and density in WordNet-based semantic similarity measures.” *2011 Conference on Empirical Methods in Natural Language Processing*, Edinburgh, July 2011, 1003–1011.

Le, Xuan; Lancashire, Ian; ———; and Jokel, Regina. “Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists.” *Literary and Linguistic Computing*, 26(4), December 2011, 435–461.

Wang, Tong and ———. “Exploring patterns in dictionary definitions for synonym extraction.” *Natural Language Engineering*, 18(3), July 2012, 313–342.

Feng, Vanessa Wei and ———. “Classifying arguments by scheme.” *Proceedings, 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oregon, June 2011, 987–996.

Wang, Tong and ———. “Near-synonym lexical choice in latent semantic space.” *Proceedings, 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, August 2010, 1182–1190.

Brooke, Julian; Wang, Tong; and ———. “Automatic acquisition of lexical formality.” *Proceedings, 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, August 2010, Poster volume pages 90–98.

Brooke, Julian; Wang, Tong; and ———. “Inducing lexicons of formality from corpora.” *Workshop on Methods for the Automatic Acquisition of Language Resources and their Evaluation Methods, 7th Lexical Resources and Evaluation Conference*, Valetta, Malta, May 2010, 17–22.

———; Riabinin, Yaroslav; and Graham, Jory. “Party status as a confound in the automatic classification of political speech by ideology.” *Proceedings, 10th International Conference on Statistical Analysis of Textual Data / 10es Journées internationales d’Analyse statistique des Données Textuelles (JADT 2010)*, Rome, June 2010, 731–742.

Marathe, Meghana and ———. “Lexical chains using distributional measures of concept distance.” *Proceedings, 11th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2010)* [Iași, Romania, March 2010], (Lecture Notes in Computer Science 6008), Berlin: Springer-Verlag, 2010, 291–302.

Wang, Tong and ———. “Extracting synonyms from dictionary definitions.” *Conference on Recent Advances in Natural Language Processing*, September 2009, Borovets, Bulgaria, 470–476.

———. “Limitations of the philosophy of language understanding implicit in computational linguistics.” *Proceedings, 7th European Conference on Computing and Philosophy*, Barcelona, July 2009, 108–109.

———. “Ontology and the lexicon.” In: Staab, Steffen and Studer, Rudi (editors), *Handbook on Ontologies* (second edition), Berlin: Springer Verlag (International Handbooks on Information Systems), 2009, 269–292. (Revision of 2004 version, invited for a new edition of the book in which the original appeared.)

Rudzicz, Frank; van Lieshout, Pascal; ———; Penn, Gerald; Shein, Fraser; and Wolff, Talya. “Towards a comparative database of dysarthric articulation.” *Proceedings of the 8th International Speech Production Seminar (ISSP 2008)*, Strasbourg, France, December 2008, 285–288.

Mohammad, Saif; Dorr, Bonnie J.; ———. “Computing word-pair antonymy.” *Proceedings, 2008 Conference on Empirical Methods in Natural Language Processing (EMNLP 2008)*, Honolulu, October 2008, 982–991.

Wilcox-O’Hearn, Amber; ———; and Budanitsky, Alexander. “Real-word spelling correction with trigrams: A reconsideration of the Mays, Damerou, and Mercer model.” In: Gelbukh, Alexander (editor), *Proceedings, 9th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2008)* [Haifa, February 2008], (Lecture Notes in Computer Science 4919), Berlin: Springer-Verlag, 2008, 605–616. **Award for best poster.**

——— and Feiguina, Ol’ga. “Bigrams of syntactic labels for authorship discrimination of short texts.” *Literary and Linguistic Computing*, 22(4), 2007, 405–417.

Niu, Yun and ———. “Identifying cores of semantic classes in unstructured text with a semi-supervised learning approach.” *Proceedings, International Conference on Recent Advances in Natural Language Processing*, September 2007, Borovets, Bulgaria, 418–424.

Feiguina, Ol’ga and ———. “Authorship attribution for small texts: Literary and forensic ex-

periments." *Proceedings, International Workshop on Plagiarism Analysis, Authorship Identification, and Near-Duplicate Detection*, 30th Annual International ACM SIGIR Conference (SIGIR '07), Amsterdam, July 2007. [Refereed by abstract only]

Mohammad, Saif; ———; and Resnik, Philip. "TOR, TORMD: Distributional profiles of concepts for unsupervised word sense disambiguation." *SemEval-2007: 4th International Workshop on Semantic Evaluations*, Prague, June 2007, 326–333.

Mohammad, Saif; Gurevych, Iryna; ———; and Zesch, Torsten. "Cross-lingual distributional profiles of concepts for measuring semantic distance." *2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2007)*, Prague, June 2007, 571–580.

———. "Views of text-meaning in computational linguistics: Past, present, and future." In: *Computation, Information, Cognition — The Nexus and the Liminal*, Dodig Crnkovic, Gordana and Stuart, Susan (editors), Cambridge Scholars Publishing, Newcastle-upon-Tyne, 2007, 270–279.

Niu, Yun; Zhu, Xiaodan; and ———. "Using outcome polarity in sentence extraction for medical question-answering." *Proceedings of the American Medical Informatics Association 2006 Annual Symposium*, Washington, D.C., November 2006, 599–603.

Mohammad, Saif and ———. "Distributional measures of concept-distance: A task-oriented evaluation." *Proceedings, 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, Sydney, Australia, July 2006, 35–43.

Inkpen, Diana and ———. "Building and using a lexical knowledge-base of near-synonym differences." *Computational Linguistics*, 32(2), June 2006, 223–262.

Mohammad, Saif and ———. "Determining word sense dominance using a thesaurus." *Proceedings of the 11th conference of the European chapter of the Association for Computational Linguistics (EACL-2006)*, April 2006, Trento, Italy, 121–128.

Budanitsky, Alexander and ———. "Evaluating WordNet-based measures of lexical semantic relatedness." *Computational Linguistics*, 32(1), March 2006, 13–47.

Graham, Neil; ———; and Marthi, Bhaskara. "Segmenting documents by stylistic character." *Natural Language Engineering*, 11(4), December 2005, 397–415.

Niu, Yun; Zhu, Xiaodan; Li, Jianhua; and ———. "Analysis of polarity information in medical text." *Proceedings of the American Medical Informatics Association 2005 Annual Symposium*, Washington, D.C., October 2005, 570–574.

Morris, Jane and ———. "The subjectivity of lexical cohesion in text." In: Shanahan, James G.; Qu, Yan; and Wiebe, Janyce (eds.) *Computing Attitude and Affect in Text: Theory and Applications*, Dordrecht: Springer, 2005, 41–48.

Inkpen, Diana; Feiguina, Ol'ga; and ———. "Generating more-positive and more-negative text." In: Shanahan, James G.; Qu, Yan; and Wiebe, Janyce (eds.) *Computing Attitude and Affect in Text: Theory and Applications*, Dordrecht: Springer, 2005, 187–198.

Li, Jianhua and ———. "Semantic knowledge in a word completion task." *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS-2005)*, Baltimore, October 2005, 121–128.

——— and Budanitsky, Alexander. "Correcting real-word spelling errors by restoring lexical cohesion". *Natural Language Engineering*, 11(1), March 2005, 87–111.

Niu, Yun and ———. “Analysis of semantic classes in medical text for question answering.” *Workshop on Question Answering in Restricted Domains at the 42nd Annual Meeting of the Association for Computational Linguistics*, Barcelona, July 2004, 54–61.

Morris, Jane and ———. “Non-classical lexical semantic relations.” *Workshop on Computational Lexical Semantics, Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, Boston, May 2004, 46–51. Reprinted in: Hanks, Patrick (editor), *Lexicology: Critical Concepts in Linguistics*, Routledge, 2007.

Morris, Jane and ———. “The subjectivity of lexical cohesion in text.” *AAAI Spring Symposium: Exploring Attitude and Affect in Text: Theories and Applications* (published as AAAI Technical report SS-04-07), Stanford, March 2004, 102–105.

Inkpen, Diana Zaiu; Feiguina, Ol’ga; and ———. “Generating more-positive and more-negative text.” *AAAI Spring Symposium: Exploring Attitude and Affect in Text: Theories and Applications* (published as AAAI Technical report SS-04-07), Stanford, March 2004, 83–89.

———. “Ontology and the lexicon.” In: Staab, Steffen and Studer, Rudi (editors), *Handbook on Ontologies*, Berlin: Springer Verlag (International Handbooks on Information Systems), 2004, 209–229.

Inkpen, Diana Zaiu and ———. “Near-synonym choice in natural language generation.” *Fourth Conference on Recent Advances in Natural Language Processing (RANLP-2003)*, September 2003, Borovets, Bulgaria, 204–211. Reprinted, slightly abridged, in: Nicolas Nicolov, Kalina Bontcheva, Galia Angelova, and Ruslan Mitkov, editors. *Recent Advances in Natural Language Processing III: Selected Papers from RANLP 2003*, John Benjamins Publishing Company, 2004, 141–150.

Graham, Neil and ———. “Segmenting a document by stylistic character.” *Workshop on Computational Approaches to Style Analysis and Synthesis, 18th International Joint Conference on Artificial Intelligence*, Acapulco, August 2003, 47–54.

Niu, Yun; ———; McArthur, Gregory; and Rodriguez-Gianolli, Patricia. “Answering clinical questions with role identification.” *Proceedings, Workshop on Natural Language Processing in Biomedicine*, 41st annual meeting of the Association for Computational Linguistics, Sapporo, Japan, July 2003, 73–80.

———. “Natural language processing, Disambiguation in”, *Encyclopedia of Cognitive Science*, Nature Publishing Group (Macmillan), 2003, Vol. 3, 181–188.

Fazly, Afsaneh and ———. “Testing the efficacy of part-of-speech information in word completion.” *Workshop on Language Modeling for Text Entry Methods, 11th Conference of the European Chapter of the Association for Computational Linguistics*, Budapest, April 2003, 9–16.

Inkpen, Diana Zaiu and ———. “Automatic sense disambiguation of the near-synonyms in a dictionary entry.” *Proceedings of the Fourth International Conference on Computational Linguistics and Intelligent Text Processing (CICLing-2003)*, Mexico City, February 2003, (Lecture Notes in Computer Science, Berlin: Springer-Verlag), 258–267.

———. “Negotiation, compromise, and collaboration in interpersonal and human–computer conversations.” *Workshop on Meaning Negotiation, 18th National Conference on Artificial Intelligence*, Edmonton, 28 July 2002, 1–4.

Inkpen, Diana Zaiu and ———. “Acquiring collocations for lexical choice between near-synonyms.” *SIGLEX Workshop on Unsupervised Lexical Acquisition, 40th meeting of the Association for Computational Linguistics*, Philadelphia, 12 July 2002, 67–76.

Edmonds, Philip and ———. “Near-synonymy and lexical choice.” *Computational Linguistics*, 28(2), June 2002, 105–144.

Budanitsky, Alexander and ———. “Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures.” *Workshop on WordNet and Other Lexical Resources, Second meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, June 2001, 29–34.

Inkpen, Diana Zaiu and ———. “Building a lexical knowledge-base of near-synonym differences.” *Workshop on WordNet and Other Lexical Resources, Second Meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, June 2001, 47–52.

Inkpen, Diana Zaiu and ———. “Experiments on extracting knowledge from a machine-readable dictionary of synonym differences.” In: Gelbukh, Alexander (editor), *Computational Linguistics and Intelligent Text Processing*, [Proceedings of the Second International Conference, Mexico City, February 2001], (Lecture Notes in Computer Science 2004), Berlin: Springer-Verlag, 2001, 264–278.

Baljko, Melanie and ———. “The importance of subjectivity in computational stylistic assessment.” *Text Technology*, 9(1), Spring 1999 [published April 2000], 5–17.

Edmonds, Philip and ———. “Reconciling fine-grained lexical knowledge and coarse-grained ontologies in the representation of near-synonyms.” *Workshop on Semantic Approximation, Granularity, and Vagueness*, Breckenridge, Colorado, April 2000.

———. “Context as a spurious concept.” *Proceedings, Conference on Intelligent Text Processing and Computational Linguistics*, Mexico City, February 2000, 273–287.

———. “What exactly are lexical concepts?” *Behavioral and Brain Sciences*, 22(1), February 1999, 45–46.

Ansari, Daniel and ———. “Generating warning instructions by planning accidents and injuries.” *Proceedings, 9th International Workshop on Natural Language Generation*, Niagara-on-the-Lake, Ontario, August 1998, 118–127.

——— and St-Onge, David. “Lexical chains as representations of context for the detection and correction of malapropisms”. In: Christiane Fellbaum (editor), *WordNet: An Electronic Lexical Database*, Cambridge, MA: The MIT Press, 1998, 305–332.

———; DiMarco, Chrysanne; Hovy, Eduard; and Parsons, Kimberley. “Authoring and generating health-education documents that are tailored to the needs of the individual patient.” In: Jameson, Anthony; Paris, Cécile; and Tasso, Carlo (editors), *User Modeling: Proceedings of the Sixth International Conference, UIM97* (Chia Laguna, Sardinia, Italy), Vienna and New York: Springer Wien New York, June 1997, 107–118.

DiMarco, Chrysanne; ———; and Hovy, Eduard. “Generation by selection and repair as a method for adapting text for the individual reader.” *Proceedings of the Flexible Hypertext Workshop* (held in conjunction with the 8th ACM International Hypertext Conference, Southampton, April 1997); published as Report C/TR97-06, Microsoft Research Institute, Macquarie University, August 1997. 36–43.

——— and DiMarco, Chrysanne. “Automatic customization of health-education brochures for individual patients.” *Proceedings, Information Technology and Community Health Conference (ITCH-96)*, Victoria, B.C., November 1996, 222–228.

Marcu, Daniel and ———. “A formal and computational characterization of pragmatic infelic-

ities." *Proceedings of the Twelfth European Conference on Artificial Intelligence*, Budapest, August 1996, 587–591. [An earlier version of this paper appeared as: Marcu, Daniel and ———. "Detecting pragmatic infelicities." *AAAI Symposium on Computational implicature: Computational approaches to interpreting and generating conversational implicature*, Stanford University, March 1996, 64–70.]

Glover, Angela and ———. "Detecting stylistic inconsistencies in collaborative writing." In: Sharples, Mike and van der Geest, Thea (editors), *The New Writing Environment: Writers at Work in a World of Technology*. London: Springer-Verlag, 1996. 147–168.

Wiebe, Janyce M.; ———; and Horton, Diane. "Language use in context." *Communications of the ACM*, **39**(1), January 1996, 102–111.

McRoy, Susan and ———. "The repair of speech act misunderstandings by abductive inference." *Computational Linguistics*, **21**(4), December 1995, 435–478.

Marcu, Daniel and ———. "A uniform treatment of pragmatic inferences in simple and complex utterances and sequences of utterances". *Proceedings, 33rd Annual Meeting, Association for Computational Linguistics*, Cambridge, MA, June 1995, 144–150.

Heeman, Peter and ———. "Collaborating on referring expressions." *Computational Linguistics*, **21**(3), September 1995, 351–382. [Earlier version appeared as technical report 435, Department of Computer Science, University of Rochester, August 1992; revised version, April 1995.]

———. "Near-synonymy and the structure of lexical knowledge." *AAAI Symposium on Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity, and Generativity*, Stanford University, March 1995, 51–56.

Marcu, Daniel and ———. "An implemented formalism for computing linguistic presuppositions and existential commitments." *Proceedings, International Workshop on Computational Semantics*, Tilburg, The Netherlands, December 1994, 141–150.

———; McRoy, Susan; Heeman, Peter; Edmonds, Philip; and Horton, Diane. "Repairing conversational misunderstandings and non-understandings." *Speech communication*, **15**(3–4), December 1994, 213–229. [An earlier version appeared as an invited paper in *Proceedings, International Symposium on Spoken Dialogue*, Waseda University, Tokyo, November 1993, 185–196, and in *Proceedings, ATR International Workshop on Speech Translation*, ATR Interpreting Telecommunications Research Laboratories, Kyoto, November 1993 [unpaginated].]

———. "Natural language analysis by computer." *The Encyclopedia of Language and Linguistics*, (R.E. Asher, editor-in-chief), Oxford: Pergamon Press, 1994. Vol 5, 2730–2736.

Mellish, C.S. and ———. "Natural language processing." *The Encyclopedia of Language and Linguistics*, (R.E. Asher, editor-in-chief), Oxford: Pergamon Press, 1994. Vol 5, 2748.

———. "Reference and anaphor resolution in natural language processing." *The Encyclopedia of Language and Linguistics*, (R.E. Asher, editor-in-chief), Oxford: Pergamon Press, 1994. Vol 7, 3487–3489.

———. "Semantic interpretation in natural language processing." *The Encyclopedia of Language and Linguistics*, (R.E. Asher, editor-in-chief), Oxford: Pergamon Press, 1994. Vol 7, 3801–3804.

McRoy, Susan and ———. "Misunderstanding and the negotiation of meaning." *AAAI Fall Symposium on Human–Computer Collaboration*, Raleigh, NC, October 1993, 57–62.

DiMarco, Chrysanne and ———. "Usage notes as the basis for a representation of near-synonymy for lexical choice." *Proceedings, 9th Annual Conference of the University of Waterloo Centre for the*



*New Oxford English Dictionary and Text Research*, Oxford, September 1993, 33–43.

DiMarco, Chrysanne and ———. "A computational theory of goal-directed style in syntax." *Computational Linguistics*, 19(3), September 1993, 451–499.

DiMarco, Chrysanne; ———; and Makuta-Giluk, Marzena. "A goal-based grammar of rhetoric." *Association for Computational Linguistics, Workshop on Intentionality and Structure in Discourse Relations*, Ohio State University, June 1993, 15–18.

——— and Wu, Dekai. "Not all reflexive reasoning is deductive." *Behavioral and Brain Sciences*, 16(3), September 1993, 462–463.

McRoy, Susan and ———. "Abductive explanations of dialogue misunderstandings." *Proceedings, 6th Conference of the European Chapter of the Association for Computational Linguistics*, Utrecht, April 1993, 277–286.

Horton, Diane and ———. "Knowledge about planning: On the meaning and representation of plan decomposition." *AAAI Spring Symposium on Reasoning about Mental States: Formal Theories and Applications*, Stanford, CA, March 1993, 74–78.

DiMarco, Chrysanne; ———; and Stede, Manfred. "The semantic and stylistic differentiation of synonyms and near-synonyms." *AAAI Spring Symposium on Building Lexicons for Machine Translation*, Stanford, CA, March 1993, 114–121.

Payette, Julie and ———. "An intelligent computer assistant for stylistic instruction." *Computers and the Humanities*, 26(2), 1992, 87–102.

——— and Ryan, Mark. "Mixed-depth representations for natural language text." In: Jacobs, Paul S. (editor). *Text-Based Intelligent Systems*, Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, 59–82.

Dick, Judy and ———. "A case-based representation of legal text for conceptual retrieval." *Workshop on Language and Information Processing, American Society for Information Science*, Washington, DC, October 1991.

Horton, Diane and ———. "Discrepancies in discourse models and miscommunication in conversation." *AAAI Fall Symposium on Discourse Structure in Natural Language Understanding and Generation*, Pacific Grove, California, November 1991, 31–32.

McRoy, Susan and ———. "An abductive account of repair in conversation." *AAAI Fall Symposium on Discourse Structure in Natural Language Understanding and Generation*, Pacific Grove, California, November 1991, 52–57.

McRoy, Susan and ———. "Repairs in communication are abductive inferences." *AAAI Fall Symposium on Knowledge and Action at the Social and Organizational Levels*, Pacific Grove, California, November 1991, 89–91.

Regoczei, Stephen and ———. "The corporation as mind: Lessons for AI." *AAAI Fall Symposium on Knowledge and Action at the Social and Organizational Levels*, Pacific Grove, California, November 1991, 95–97.

Dick, Judy and ———. "Intelligent text retrieval." *Text retrieval: Workshop notes from the Ninth National Conference on Artificial Intelligence (AAAI-91)*, Anaheim, California, July 1991.

———. "Existence assumptions in knowledge representation." *Artificial Intelligence*, 49, May 1991, 199–242. (This issue of the journal was reprinted as: Brachman, Ronald J.; Levesque, Hector J.; and Reiter, Raymond (editors). *Knowledge Representation*. Cambridge, MA: The MIT

Press, 1992.)

Payette, Julie and ———. "Computer-assisted instruction in syntactic style", *Proceedings, ACH/ALLC '91, "Making connections" [International Conference on Computers and the Humanities]*, Tempe, Arizona, March 1991. 333–340.

Morris, Jane and ———. "Lexical cohesion computed by thesaural relations as an indicator of the structure of text." *Computational Linguistics*, 17(1), March 1991, 21–48.

Catt, Mark and ———. "An intelligent CALI system for grammatical error diagnosis." *Computer Assisted Language Learning*, 3, November 1990, 3–26.

DiMarco, Chrysanne and ———. "Accounting for style in machine translation." *Third International Conference on Theoretical Issues in Machine Translation*, Austin, June 1990.

Lyons, Dan and ———. "A frame-based semantics for focusing subjuncts." *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, Pittsburgh, June 1990, 54–61. Also published as technical report ACT-NL-109-90, Microelectronics and Computer Technology Corporation, Austin, Texas, July 1990.

———. "Mixed-depth representations for natural language text." *AAAI Spring Symposium on Text-Based Intelligent Systems*, Stanford, March 1990, 25–29. Reprinted in: Jacobs, Paul S. (editor). *Text-Based Intelligent Systems: Current research in text analysis, information extraction and retrieval*. Technical report 90CRD198, Corporate Research and Development, General Electric Company, September 1990, 10–14.

McRoy, Susan Weber and ———. "Race-based parsing and syntactic disambiguation." *Cognitive Science*, 14(3), July–September 1990, 313–353.

Regoczei, Stephen and ———. "Sortal analysis with SORTAL, a software assistant for knowledge acquisition." *Proceedings, Fourth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, Banff, October 1989. [Also published as technical report CSRI-232, Computer Systems Research Institute, University of Toronto, August 1989.]

———. "Ontological assumptions in knowledge representation." *Proceedings, First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, May 1989. San Mateo, CA: Morgan Kaufmann Publishers, 157–169. [A short earlier version appears in the pre-prints of the Workshop on Formal Aspects of Semantic Networks, Santa Catalina, February 1989.]

Regoczei, Stephen and ———. "On 'Extracting knowledge from text': Modelling the architecture of language users." *Proceedings, Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems*, Paris, July 1989, 196–211. [Also published as technical report CSRI-225, Computer Systems Research Institute, University of Toronto, January 1989.]

Regoczei, Stephen and ———. "The meaning triangle as a tool for the acquisition of abstract, conceptual knowledge." *International Journal of Man–Machine Studies*, 33(5), November 1990, 505–520. [Previously published in *Proceedings, 3rd Workshop on Knowledge Acquisition for Knowledge-Based Systems*, Banff, November 1988.] [Also published as technical report CSRI-211, Computer Systems Research Institute, University of Toronto, May 1988.]

Horton, Diane and ———. "Presuppositions as beliefs." *12th International Conference on Computational Linguistics (COLING-88)*, Budapest, August 1988, 255–260.

DiMarco, Chrysanne and ———. "Stylistic grammars in language translation." *12th International Conference on Computational Linguistics (COLING-88)*, Budapest, August 1988, 148–153.

———. “Resolving lexical ambiguity computationally with spreading activation and Polaroid Words.” In: Small, Steven; Cottrell, Garrison and Tanenhaus, Michael (editors). *Lexical Ambiguity Resolution*, Los Altos, CA: Morgan Kaufmann, 1988, 73–107.

———. “Semantic interpretation and ambiguity.” *Artificial Intelligence*, **34**(2), March 1988, 131–177.

Selman, Bart and ———. “Parsing as an energy minimization problem.” In: Davis, Lawrence (editor). *Genetic Algorithms and Simulated Annealing* (Research notes in artificial intelligence), Pitman, 1987, 141–154. Revised version appears in: Geert Adriaens and Udo Hahn (editors). *Parallel Natural Language Processing*. Norwood, NJ: Ablex Publishing, 1994, 238–254.

———. “Semantics.” In: Shapiro, Stuart Charles and Eckroth, David (editors). *Encyclopedia of Artificial Intelligence*. New York: Wiley-Interscience / John Wiley, 1987, 1024–1029.

———. “Why dictionaries should list case structures.” *Advances in Lexicology: Proceedings of the Second Annual Conference of the University of Waterloo Centre for the New Oxford English Dictionary*. University of Waterloo, November 1986, 147–162.

Fawcett, Brenda and ———. “The detection and representation of ambiguities of intension and description.” *Proceedings of the 24th Annual Meeting, Association for Computational Linguistics*, New York, June 1986, 192–199.

Selman, Bart and ———. “A rule-based connectionist parsing system.” *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, Irvine, August 1985, 212–221.

———. “A semantic process for syntactic disambiguation.” *Proceedings, Fourth National Conference on Artificial Intelligence (AAAI-84)*, Austin, August 1984, 148–152.

———. “Jumping to conclusions: Psychological reality and unreality in a word disambiguation program.” *Proceedings, Sixth Meeting of the Cognitive Science Society*, Boulder, June 1984, 179–182.

———. “A foundation for semantic interpretation”. *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, Mass., June 1983, 64–73. [Also published as technical report CS-83-03, Department of Computer Science, Brown University, January 1983.]

——— and Charniak, Eugene. “Word sense and case slot disambiguation”. *Proceedings of the Second National Conference on Artificial Intelligence (AAAI-82)*, Pittsburgh, August 1982, 95–98.

———. “An evaluation of evidence for innate sex differences in linguistic ability”. *Journal of Psycholinguistic Research*, **11**(2), March 1982, 95–113.

———. “Discourse-oriented anaphora resolution in natural language understanding: A review”. *American Journal of Computational Linguistics*, **7**(2), April–June 1981, 85–98.

——— and Talent, Nadia. “What should Computer Scientists read?” *Proceedings of the Eighth Australian Computer Conference*, Canberra, August 1978, 1707–1716.

———. “Discipline impact factors: A method for determining core journal lists”. *Journal of the American Society for Information Science*, **29**(4), July 1978, 171–172.

——— and Talent, Nadia. “Computer Science journals—An iterated citation analysis”. *IEEE Transactions on Professional Communication*, **PC-20**(4), December 1977, 233–238.

### Unrefereed publications

Ethayarajh, Kawin; Duvenaud, David; and ———. “Towards understanding linear word analogies.” *arXiv:1810.04882*, 11 October 2018 (v5, 24 December 2018).

Hammond, Adam; Brooke, Julian; and ———. “Modeling modernist dialogism: Close reading with big data.” *Reading Modernism with Machines: Digital humanities and modernist literature*, edited by Shawna Ross and James O’Sullivan. London: Palgrave Macmillan, 2016, 49–77.

——— and Feng, Vanessa Wei. “Automatic exploration of argument and ideology in political texts.” In: Mohammed, Dima and Lewiński, Marcin (editors), *Argumentation and Reasoned Action: Proceedings of the First European Conference on Argumentation, Lisbon, 9–12 June 2015*, volume II, College Publications (Studies in Logic and Argumentation, volume 52), 2016, 493–504.

———; Feng, Vanessa Wei; Cochrane, Christopher; and Naderi, Nona. “Argumentation, ideology, and issue framing in parliamentary discourse.” *Proceedings of the Workshop on Frontiers and Connections between Argumentation Theory and Natural Language Processing*, Forlì-Cesena, Italy, July 2014. CEUR Workshop Proceedings, volume 1341.

———. “Computational Linguistics.” In: Allan, Keith (editor), *The Oxford Handbook of the History of Linguistics*, Oxford University Press, 2013, 707–726.

———. “Overcoming linguistic barriers to the multilingual Semantic Web” [extended abstract]. In: Paul Buitelaar, Key-Sun Choi, Philipp Cimiano, and Eduard H. Hovy (editors), *The Multilingual Semantic Web (Dagstuhl Seminar 12362)*, *Dagstuhl Reports*, 2(9), 2013. 44–45.

Brooke, Julian and ———. “Lexicalizing computational stylistics for language learner feedback.” *Proceedings, Conference on Stylistics Across Disciplines*, Leiden, June 2011. [Reviewed by extended abstract.]

——— and Mohammad, Saif. “Semantic distance measures with distributional profiles of coarse-grained concepts.” In: Mehler, Alexander; Kühnberger, Kai-Uwe; Lobin, Henning; Lungen, Harald; Storrer, Angelika; and Witt, Andreas (editors), *Modeling, Learning, and Processing of Text Technological Data Structures*. Springer (Studies in Computational Intelligence series, vol. 370), 2011, 61–79 [Invited book chapter].

Niu, Yun and ———. “Analyzing the text of clinical literature for question answering.” In: Prince, Violaine and Roche, Mathieu (editors), *Information Retrieval in Biomedicine*, IGI Global, 2009, 190–220 [Invited book chapter].

———. “The future of text-meaning in computational linguistics.” In: Sojka, Petr; Horák, Aleš; Kopeček, Ivan; and Pala, Karel (editors), *Proceedings, 11th International Conference on Text, Speech and Dialogue (TSD 2008)* [Brno, Czech Republic, September 2008], (Lecture Notes in Artificial Intelligence 5246), Berlin: Springer-Verlag, 2008, 1–9. [Paper to accompany invited keynote presentation]

———. Foreword to *Word Sense Disambiguation*, edited by Eneko Agirre and Philip Edmonds, Springer, 2006. [Invited]

———. “Association for Computational Linguistics.” In: Brown, Keith (editor-in-chief); Anderson, Anne H.; Bauer, Laurie; Berns, Margie; Hirst, Graeme; and Miller, Jim (coordinating editors). *Encyclopedia of Language and Linguistics* (second edition), Elsevier Science, December 2005.

———. “Human language technology.” In: Brown, Keith (editor-in-chief); Anderson, Anne H.; Bauer, Laurie; Berns, Margie; Hirst, Graeme; and Miller, Jim (coordinating editors). *Ency-*

*lopedia of Language and Linguistics* (second edition), Elsevier Science, December 2005.

Morris, Jane; Beghtol, Clare; and ———. "Term relationships and their contribution to text semantics and information literacy through lexical cohesion." *Proceedings of the 31st Annual Conference of the Canadian Association for Information Science*, Halifax, Nova Scotia, 30 May – 1 June 2003, 153–168.

DiMarco, Chrysanne; ———; Wanner, Leo; and Wilkinson, John. "HealthDoc: Customizing patient information and health education by medical condition and personal characteristics." *Workshop on Artificial Intelligence in Patient Education*, Glasgow, August 1995.

———. "Natural-language processing." In: Bitter, Gary G. (editor). *The Macmillan Encyclopedia of Computers*. New York: Macmillan, 1992, 699–703.

Ryan, Mark; DiMarco, Chrysanne; and ———. "Focus shifts as indicators of style in paragraphs." In DiMarco, Chrysanne *et al*, *Four Papers on Computational Stylistics*. Research report CS-92-35, Department of Computer Science, University of Waterloo, June 1992.

Regoczei, Stephen and ———. "Knowledge and knowledge acquisition in the computational context." In: Hoffmann, Robert R. (editor). *The Psychology of Expertise: Cognitive research and empirical AI*. New York: Springer-Verlag, 1992, 12–25.

———. "Planning the future of natural language research (even in Canada)." *Canadian Artificial Intelligence*, number 26, February 1991, 10–13.

———. "Computational models of ambiguity resolution." In: Gorfein, David S. (editor). *Resolving Semantic Ambiguity* (Cognitive science series), Springer-Verlag, 1989, 255–275.

Regoczei, Stephen and ———. "Knowledge acquisition as knowledge explication by conceptual analysis." Technical report CSRI-205, Computer Systems Research Institute, University of Toronto, January 1988.

Levesque, Hector. *Functional Programming in Lisp* (revised and edited by Graeme Hirst). Department of Computer Science, University of Toronto, April 1987.

Horton, Diane and ———. "Presuppositions as beliefs: A new approach." In: Geller, James and Bettinger, Keith (editors). *UBGCSS-87: Proceedings of the Second UB Graduate-Conference on Computer Science*. Technical report 87-04, Department of Computer Science, State University of New York at Buffalo, March 1987, 30–38.

### Edited books

Brown, Keith (editor-in-chief); Anderson, Anne H.; Bauer, Laurie; Berns, Margie; ———; and Miller, Jim (coordinating editors). *Encyclopedia of Language and Linguistics* (second edition), (14 vols.), Elsevier Science, December 2005.

### Monographs

———. *Semantic Interpretation and the Resolution of Ambiguity* (Studies in natural language processing). Cambridge, England: Cambridge University Press, 1987. (Reprinted 1992.)

———. *Anaphora in Natural Language Understanding: A survey* (Lecture Notes in Computer Science 119). Berlin: Springer-Verlag, 1981.

### Databases

Rudzicz, Frank; ———; Van Lieshout, Pascal; Penn, Gerald; Shein, Fraser; Namasivayam, Aravind Kumar; Wolff, Talya. The TORGO database of dysarthric articulation. Catalog number LDC2012S02, Linguistic Data Consortium, 2012.

### Review articles

———. “Does Conversation Analysis have a role in computational linguistics?” *Computational linguistics*, 17(2), June 1991, 211–227. [Review of: Luff, Paul; Gilbert, Nigel; and Frohlich, David (editors). *Computers and Conversation*. London: Academic Press, 1990]. [semi-refereed]

### Book reviews, etc.

———. Review of: Clark, Alexander; Fox, Chris; and Lappin, Shalom (editors), *The Handbook of Computational Linguistics and Natural Language Processing*, Malden, MA: Wiley-Blackwell, 2010. In: *Language*, 87(4), December 2011, 897–899.

———. Review of: Scott, Mike and Thompson, Geoff (editors), *Patterns of Text: In Honour of Michael Hoey*, Amsterdam: John Benjamins, 2001. In: *Computational Linguistics*, 28(4), December 2002, 560–564.

———. Review of: Biber, Douglas; Johansson, Stig; Leech, Geoffrey; Conrad, Susan; and Finegan, Edward, *Longman Grammar of Spoken and Written English*, Harlow, Essex: Pearson Education Ltd, 1999. In: *Computational Linguistics*, 27(1), March 2001, 132–139.

———. Review of: Vossen, Piek (editor). *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*, Dordrecht: Kluwer Academic Publishers, 1998. In: *Computational Linguistics*, 25(4), December 1999, 628–630.

———. Review of: Cottrell, Garrison W. *A Connectionist Approach to Word Sense Disambiguation* (Research notes in artificial intelligence). London: Pitman and San Mateo, CA: Morgan Kaufmann, 1989. In: *Computational Linguistics*, 16(4), December 1990, 241–242.

———. Review of: Batori, István S.; Lenders, Winfried; and Putschke, Wolfgang (editors). *Computational Linguistics: An International Handbook on Computer Oriented Language Research and Applications / Computerlinguistik: Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendungen* (Handbooks of linguistics and communication science 4 / Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Walter de Gruyter, 1989. In: *Computational Linguistics*, 16(2), June 1990, 121–123.

———. Review of: Butler, Christopher. *Computers in Linguistics*. Oxford: Blackwell, 1985. *Computational Linguistics*, 13(3–4), July–September 1987, 335–336.

———. Review of: van Bakel, Jan. *Automatic Semantic Interpretation: A Computer Model of Understanding Natural Language*, Dordrecht: Foris Publications, 1984. *Computational Linguistics*, 11(2–3), April–September 1985, 185–186.

Many short reviews in *Computational Linguistics*, 1988–2014.

Many short reviews in *Computing Reviews*, 1982–2000.

Many short reviews in *Canadian Artificial Intelligence*, 1985–1996.

### Unpublished conference and workshop presentations

Leung, Yvonne; Wong, Jiahui; Gancarz, Mathew; Gratzner, David; ———; Da Silva, Daswin; Alahakoon, Daminda; Bender, Jackie; French, Leon; and Esplen, Mary Jane. “Building an artificial intelligent system to enhance online support groups in cancer.” 22nd World Congress of Psycho-Oncology, Kyoto, August 2020. [Rescheduled to 2021.]

Hammond, Adam; Vishnubhotla, Krishnapriya; and ———. “The words themselves: A content-based approach to quote attribution.” Digital Humanities Conference, Ottawa, July 2020.

Naderi, Nona and ———. “Reputation management strategies in argumentative political exchanges.” Second European Conference on Argumentation, Fribourg, 20–23 June 2017.

———, “Automatic exploration of argument and ideology in political texts.” First European Conference on Argumentation, Lisbon, 9–12 June 2015.

Rheault, Ludovic; Beelen, Kaspar; Cochrane, Christopher; ———. “Measuring emotion in parliamentary debates using natural language processing”, Annual Conference of the Canadian Political Science Association, 2-4 June 2015, Ottawa. [Refereed by abstract only; presentation by Rheault.]

Brooke, Julian and ———. “Investigating the influence of multi-L1 learner corpora variables on native language identification.” Learner Corpus Research Conference, Bergen, Norway, 27–29 September 2013. [Refereed by abstract only; presentation by Brooke.]

———. “Usability — a neglected dimension.” *Beyond the PDF 2* Conference, Amsterdam, March 2013.

——— and Talent, Nadia. “NLP methods for discovering precedent research in the legacy literature.” *Beyond the PDF 2* Conference, Amsterdam, March 2013.

Brooke, Julian and ———. “Multidimensional analysis versus latent semantic analysis for constructing a register space: Are hand-coded features needed or is bags-of-words enough?” Presented at the Linguistic Society of Belgium conference on Genre- and Register-related Text and Discourse Features in Multilingual Corpora, Brussels, January 2013. [Refereed by abstract only; presentation by Brooke.]

———. “Overcoming linguistic barriers to the multilingual Semantic Web.” Dagstuhl Seminar on the Multilingual Semantic Web, Leibniz-Zentrum für Informatik, Schloss Dagstuhl, Germany, September 2012. [Invited]

Brooke, Julian and ———. “Factors of formality: A dimension of register in a sociolinguistic corpus.” Poster at the Georgetown University Round Table on Languages and Linguistics 2012: Measured Language: Quantitative Approaches to Acquisition, Assessment, Processing and Variation, Washington D.C., March 2012. [Refereed by abstract only; presentation by Brooke.]

Brooke, Julian and ———. “Native language detection with ‘cheap’ learner corpora” *LCR 2011: 20 years of learner corpus research: Looking back, moving ahead*, Louvain-la-Neuve, Belgium, September 2011. [Refereed by abstract only; presentation by Brooke.]

Wang, Tong and ———. “Associating difficulty in near-synonymy choice with types of nuance using core vocabulary.” Symposium on Re-thinking Synonymy, Helsinki, October 2010. [Reviewed by extended abstract; presentation by Hirst.]

———; Riabinin, Yaroslav; Graham, Jory; and Boizot-Roche, Magali. “Text to Ideology or Text to Party Status?” From Text to Political Positions workshop, University of Amsterdam, 8–9

April 2010. [Reviewed by full paper submission; presentation by Hirst.]

Lancashire, Ian and ———. "Vocabulary changes in Agatha Christie's mysteries as an indication of dementia: A case study." 19th Annual Rotman Research Institute Conference, Cognitive Aging: Research and Practice, 8–10 March 2009, Toronto. [Poster; reviewed by extended abstract.]

Feiguina, Ol'ga and ———. "Forensic authorship attribution for small texts." 8th Biennial Conference on Forensic Linguistics / Language and Law, Seattle, July 2007 [reviewed by abstract; presentation by Hirst].

———. "Views of text-meaning in computational linguistics." European Conference on Computing and Philosophy (E-CAP 2005), Västerås, Sweden, June 2005. [refereed by extended abstract]

Morris, Jane and ———. "Readers' perceptions of lexical cohesion and lexical semantic relations in text." *14th annual meeting of the Society for Text and Discourse*. Chicago, August 2004. [refereed by abstract; presentation by Morris]

Fazli, Afsaneh; ———; Shein, Fraser, and Nantais, Tom. "Word prediction as a writing aid." *CITO Research Review*, Kingston, May 2001.

Baljko, Melanie and ———. "Subjectivity in stylistic assessment: Helping computer-supported collaborative writers." *Writing and Computers 10*, Brighton, September 1997.

———. "The conceptual and subconceptual differentiation of near-synonyms." *Second Conference on Conceptual Structure, Discourse, and Language*, Buffalo, April 1996.

Glover, Angela and ———. "Automatically detecting stylistic inconsistencies in collaborative writing." *European Conference on Writing and Computers*, Utrecht, October 1994.

———. "Conceptual and sub-conceptual differentiation." Presented at the Workshop on cognitive and ontological foundations of knowledge engineering, First International Summer Institute in Cognitive Science, State University of New York at Buffalo, July 1994.

Ryan, Mark; DiMarco, Chrysanne; and ———. "Focus shifts as indicators of style in text." Presented at the second annual meeting of the Society for Text and Discourse, San Diego, June 1992.

DiMarco, Chrysanne and ———. "Linguistic style as constituent order." Presented at the 20th Annual University of Wisconsin-Milwaukee Linguistics Symposium, Milwaukee, Wisconsin, April 1991.

———. "Focus in reference resolution and natural language understanding." *Language and Speech Conference*, University of Melbourne, November 1977.

Collins, Anthony and ———. "Computer-assisted teaching of music." 46th Congress of the Australian and New Zealand Association for the Advancement of Science, Canberra, January 1975.



### Professional tutorials, summer schools, etc

“Lexical similarity based on lexicographic resources”, class at the ACL/HCSNet Advanced Program in Natural Language Processing, University of Melbourne, 10–14 July 2006.

“Introduction to non-statistical natural language processing”, tutorial at the combined Human Language Technology Conference and 3rd Meeting of the North American Chapter of the Association for Computational Linguistics, Edmonton, May 2003.

Lecturer, Eurolan Summer Institute, Iași, Romania, July–August 2001.

“Natural language understanding and computational linguistics”, a credit course of eight 100-minute lectures at the First International Summer Institute in Cognitive Science, State University of New York at Buffalo, July 1994.

[With Chrysanne DiMarco] Tutorial on current approaches to natural language semantics, National Conference on Artificial Intelligence (AAAI-90), Boston, July 1990.

Workshop “Introduction to computational linguistics”, sponsored by the Centre for Computing in the Humanities at the Ninth International Summer Institute for Semiotic and Structural Studies, Toronto, June 1987.

Tutorial on current approaches to natural language semantics, 24th annual meeting, Association for Computational Linguistics, New York, June 1986.

### Popular media

- Research featured in *AI in Healthcare*, “Developing regions may benefit by AI analysis of ‘verbal autopsies’”, 18 July 2019.  
 Interviewed on Global TV News, Toronto, about identifying Ontario government document leaker, 6 February 2019.  
 Quoted in article on Ontario government document leaks, *Toronto Star* in print and online and in *iPolitics*, 6 February 2019.  
 Mentioned in article on Lipad project (digitization of Canadian Parliamentary Proceedings), *University Affairs*, November 2018, page 10.  
 Interviewed on City TV News, Toronto, about online voting in Burlington, 12 October 2018.  
 Interviewed on City TV News, Toronto, about Facebook data breach, 28 September 2018.  
 Five-minute interview on CBC Radio on Google’s collection of location data; given eight times to local hosts in Toronto (*Here and Now*), New Brunswick, Edmonton, London Ontario, Charlottetown, Halifax, Whitehorse, and Yellowknife, 13 August 2018; and on radio Canada International 14 August 2018.  
 Quoted in article “Globalization! Pecarity! Robots”, *UTSC Commons*, Spring 2018.  
 Five-minute interview on CBC Radio on Facebook data misuse and privacy; given five times to local hosts in Toronto (*Here and Now*), Victoria, Whitehorse, Sydney, and Yellowknife; 5 April 2018.  
 Quoted in article “DriveHer, ride-sharing app for women, suspends service after data breach exposes personal information”, *Toronto Star*, 5 April 2018.  
 Quoted in article “AI can now identify people as gay or straight from their photo”, *The Varsity*, 24 September 2017.  
 Quoted in article “Users face consequences as Facebook struggles to filter hate speech”, *The Globe and Mail*, 27 July 2017.  
 Featured in article “‘A stupendous person’: Friends laud Julie Payette, Canada’s next GG” on CBC News website, with corresponding clip on CBC Radio “The World at Six”. 13 July 2017.  
 Featured in article on early detection of Alzheimer’s disease by linguistic means, *Nautilus*, issue 40, 29 Sept 2016 (“Alzheimer’s early tell”, by Adrienne Day).  
 Quoted in article on Internet vigilantism, *Toronto Star*, 7 October 2016.  
 Referred to in article on detection of early Alzheimer’s disease, *New York Times*, 31 March 2015 (with link to Le et al (2011) in online version).  
 Quoted extensively in Canadian Press / *The Globe and Mail* article on the Turing test, 10 June 2014; picked up by others, including CTV News website, *Times-Colonist*, and *Science Recorder*.  
 Three-minute interview on CTV News Channel on the Turing test, 9 June 2014.  
 Quoted in online *Toronto Star* article on automatic detection of sarcasm, 5 June 2014.  
 Extensively quoted in article on the role of “sentient computers” in doomsday scenarios, *UofT Magazine*, Winter 2013.  
 Quoted in online *Toronto Star* article on computer security, 4 October 2013.  
 Quoted in Global News online item on privacy issues regarding the Memoto lifelogging camera, 24 Oct 2012.  
 Quoted in CBC News web item on online protest, 16 March 2012.  
 Quoted in article on privacy on BlackBerries, *The Varsity*, 30 January 2012.  
 Quoted in article on linguistic code-breaking, *Los Angeles Times*, 26 October 2011; reprinted in the *Toronto Star*, 12 November 2011.  
 Interviewed re empirical natural language processing and our research on markers of dementia in writing on the radio program *Word of Mouth*, (episode: “Counting Word Incidences”), BBC Radio 4, 23 August 2011.  
 Quoted in article on authorship attribution software in biblical scholarship, Associated Press, 30 June 2011.  
 Interviewed re our research on markers of dementia in the writing of Agatha Christie, on *Breakfast with Red Symons*, ABC Radio, Melbourne, 24 February 2011.  
 Interviewed re privacy on CBC Radio *Here and Now*, 26 November 2010.  
 Quoted in article in privacy, *Toronto Sun*, 26 November 2010.  
 Quoted in article on Facebook privacy, *Toronto Star* online article, 2 November 2010.

Interviewed for, and appeared in short clip on, *CBC News Toronto*, 19 October 2010, on privacy issues in disk-based photocopiers. Quoted in article on security aspects of computer- and network-based photocopiers, *Toronto Star*, 18 March 2010.

Article in *UofT Magazine*, Winter 2010, on our research with Ian Lancashire on markers of dementia in Agatha Christie's writing. Short item in *The Globe and Mail*, 9 March 2010, on our research with Ian Lancashire on markers of dementia in Agatha Christie's writing.

Feature article in *The Toronto Star*, 23 January 2010, reported our research with Ian Lancashire and Xuan Le on markers of dementia in the writing of Iris Murdoch, P.D. James, and Agatha Christie.

Short article in *New York Times Magazine*, 13 December 2009 (the annual "Year in Ideas" issue), reported our research with Ian Lancashire on markers of dementia in Agatha Christie's writing.

Quoted in article on machine translation, *Globe and Mail*, 25 November 2009.

Interviewed for, and appeared in short clip on, *First Local*, Rogers TV, York Region, 8 October 2009, on privacy issues in Google StreetView.

Full-page article in *Macleans* (week ending 6 April 2009) reported our research with Ian Lancashire on markers of dementia in Agatha Christie's writing. Picked up and reported in *The Guardian* (London, 3 April 2009), *The Daily Telegraph* (London, 4 April 2009), United Press International (4 April 2009), *Calgary Herald* (6 April 2009), *Hindustan Times* (New Dehli, 4 April 2009), *Gazetepost* (Turkish online newspaper, 6 April 2009), *El Universal* (Mexico City, 5 April 2009), *Libero* (Milan, 4 April 2009), and various other websites.

Interviewed in segment on Facebook privacy on *The National*, CBC Television, 18 February 2009.

Interviewed [with William Gibson and David Annandale] by Shelagh Rogers on *Sounds Like Canada*, CBC Radio, 4 October 2007.

Quoted in feature article on software piracy, *The Toronto Star*, 30 April 2002.

Guest on "The Chatroom", CTV TalkTV, 16 April 2001, talking about the effect of computers on language (45-minute discussion).

Quoted in feature article on intelligent home appliances, *The Toronto Star*, 4 February 2001.

Extensively quoted in feature article on artificial intelligence in the film *2001: A Space Odyssey*, *The Toronto Star*, 30 December 2000.

Extensively quoted in feature article on privacy and technology, *The Toronto Star*, 12 August 2000.

Quoted in feature article on the future of technology, *The Toronto Star*, 11 June 2000.

Quoted in feature article on the mouse, *The Toronto Star*, 6 February 2000.

Quoted in feature article on robot pets, *The Toronto Star*, 15 August 1999.

Extensively quoted in feature article on artificial intelligence in the film *2001: A Space Odyssey*, *The Toronto Star*, 12 January 1997.

Appeared on "Minds, Machines, and Mystery", an episode of *Techno TV*, Discovery Channel, 3 December 1996.

Speaker, Symposium "Artificial intelligence: 2001 and beyond", Canadian Society for the Weizmann Institute of Science, January 1990.

Interview by Connie Bryson, *Canadian artificial intelligence*, number 21, October 1989, 8–11.

"Mother tongue to boolean". *Electric word*, 14, July/August 1989, 38–43.

Appeared on the TVOntario series "Artificial intelligence: Fulfilling the dream", 1987.

Series of two interviews by Peter Czowski on "Morningside", CBC Radio, 25–26 February 1986.

## **Invited lectures, talks, colloquia, etc.**

---

### **Invited or keynote conference speaker**

Responsible AI Summit, Montreal, 24 October 2019. "Ethics in Natural Language Processing".

Conference on Ethics in Artificial Intelligence, Ottawa, 21 February 2019. "Ethics in Natural Language Processing".

Workshop on Ethics in Natural Language Processing, 15th Conference of the European Chapter of the Association for Computational Linguistics, Valencia, 4 April 2017. "NLP Ethics in a Dangerous Time."

Workshop on Semantic Technologies for Research in the Humanities and Social Sciences (STRiX), University of Gothenburg, 24 November 2014. "Automatically identifying ideology, argumentation schemes, and discourse structure in parliamentary text."

World Affairs Conference, Upper Canada College, 4 February 2014. "Surveillance for security."

ExPoSe project inaugural meeting, Tweede Kamer, The Hague, 23 Oct 2013. "Finding positions in parliamentary text."

Workshop on Lexical Semantic Resources for NLP, NoDaLiDa 2013 (19th Nordic Conference on Computational Linguistics, Northern European Association for Language Technology), Oslo, 22 May 2013. "Lexicons versus ontologies."

7th International Conference on Natural Language Processing (ICON 2009), University of Hyderabad, Hyderabad, India, 16 December 2009. "Who decides what a text means?"

2009 Conference of die Gesellschaft für Sprachtechnologie und Computerlinguistik (German Society for Computational Linguistics), Potsdam, 30 September 2009. "Who decides what a text means?"

International Conference on Text, Speech and Dialogue (TSD 2008), Brno, Czech Republic, 12 September 2008. "The future of text-meaning in computational linguistics."

Conference on Processing Text-Technological Resources, University of Bielefeld, 14 March 2008. "Semantic distance measures with distributional profiles of coarse-grained concepts." [co-authored with Saif Mohammad]

Second International Workshop on Paraphrasing, Sapporo, Japan, July 2003. "Paraphrasing paraphrased."

Inauguration Symposium, Human Language Technology Research Institute, University of Texas at Dallas, March 2002. "Context as a spurious concept."

Third conference on Recent Advances in Natural Language Processing, Tzgov Chark, Bulgaria, September 2001. "Measures of semantic relatedness and the detection and correction of real-word spelling errors".

Second International Conference on Computational Linguistics and Intelligent Text Processing, Mexico City, February 2001, "Measures of semantic relatedness: Problems in semantic distance and intelligent spelling checking".

Japan Electronic Industry Development Association, Natural Language Processing Seminar, Kyoto, November 1998, "Near-synonymy and semantic distance: Problems in lexical choice and intelligent spelling checking".

AAAI Fall Symposium on Context in Knowledge Representation and Natural Language, Cambridge, MA, November 1997. "Context as a spurious concept".

Conference on Ontological and Linguistic Categories, University of San Marino, June 1997. "Near-synonymy and the structure of lexical knowledge".

The tenth biennial conference on AI and Cognitive Science, Society for the Study of Artificial Intelligence and the Simulation of Behaviour, Sheffield, April 1995. "Repairing conversational misunderstandings and non-understandings".

International Symposium on Spoken Dialogue, Waseda University, Tokyo, November 1993. "Repairing conversational misunderstandings and non-understandings".

ATR International Workshop on Speech Translation, Kyoto, November 1993. "Repairing conversational misunderstandings and non-understandings".

Symposium on Symbol Manipulation, Connectionism, and the Semantics of Natural Language, Helsinki, November 1991. "Ontological assumptions in knowledge representation" and "Repairing conversational misunderstandings and non-understandings".

Conference of the Canadian Society for Computational Studies of Intelligence, Ottawa, May 1990. "Planning the future of natural language research (even in Canada)".

### Other invited conference participation

Discussant, 7th Conference on New Directions in Analyzing Text as Data, Boston, October 2016.

“Competitor”, NLP Idol panel [discussing old research papers that should be “plucked from obscurity”] at the 2012 Conference of the North American Chapter of the Association for Computational Linguistics, June 2012.

Panelist, Workshop on Semantic Search, at Human Language Technologies: The 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics, Los Angeles, June 2010.

Panelist, Translation Technology Showcase, Multilingual Translation Technology Centre, University of Toronto, January 2001.

Panelist, AAAI Symposium on Intelligent Text Summarization, Stanford, March 1998.

Invited respondent, Workshop on “Views on pragmatics: From the conventional to the computational”, University of Western Ontario, October 1991.

Discussant, Workshop on Formal Aspects of Semantic Networks, Santa Catalina, February 1989.

### Invited seminars and colloquia

“Classifying verbal autopsy records by cause of death using neural networks and temporal reasoning.”

Österreichisches Forschungsinstitut für Artificial Intelligence (Austrian Research Institute for Artificial Intelligence), 3 July 2019.  
Cluster of Excellence Cognitive Interactive Technology, Universität Bielefeld, 18 March 2019.  
Fachbereich Informatik, Technische Universität Darmstadt, 14 March 2019.  
Web and Data Science Group, Universität Mannheim, 12 March 2019.  
Department of Linguistics and Philology, Uppsala University, 19 January 2018.

“Digitization of the Canadian Parliamentary Proceedings.”

Big Data Workshop, Department of Comparative Politics, University of Bergen, 24 February 2017.

“Is the singularity near?”

Ideas for the World, Victoria University, 22 March 2016.

“Who decides what a text means? (and what the answer implies for computational linguistics)”

Institut für Computerlinguistik, Universität Heidelberg, 13 March 2019.  
Department of Computing and Information Systems, University of Melbourne, May 2017.  
Language Technologies Institute, Carnegie Mellon University, February 2016.  
Centre for Linguistic Theory and Studies in Probability, University of Gothenburg, March 2016.

“Automatically identifying ideology, argumentation schemes, and discourse structure in parliamentary text.”

Department of Linguistics, University of Uppsala, November 2014.

“Natural language processing methods for the detection of symptoms of Alzheimer’s disease in writing.”

Språkbanken, University of Gothenburg, May 2013.  
Department of Computer Science, University of Illinois, November 2012.  
Deutsche Forschungszentrum für Künstliche Intelligenz, Berlin, September 2012.  
Department of Linguistics, University of Potsdam, September 2012.  
Linguistic Computing Laboratory, Sapienza University of Rome, September 2012.  
College of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, July 2012.  
Faculty of Information Technology, Monash University, August 2011.  
Department of Computer Science and Software Engineering, University of Melbourne, August 2011.  
Google Research, July 2010.

“Who decides what a text means?”

Språkbanken, University of Gothenburg, May 2013.  
Faculty of Computer Science and Information Technology, University of Malaysia, Sarawak, August 2009.

“Understanding text from both the user’s and the writer’s perspective”

KMDI@13 lecture series, Knowledge Media Design Institute, University of Toronto, April 2009.

“Semantic distance measures with distributional profiles of coarse-grained concepts” [co-authored with Saif Mohammad].

Faculty of Computer Science and Information Technology, University of Malaysia, Sarawak, August 2009.  
Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, May 2009.  
Centre for Research in Intelligent Systems, Faculty of Information Technology, Monash University, April 2009.  
Center for Cognitive Science, University at Buffalo, November 2008.

“Lexical chains and semantic distance, within and across languages” [co-authored with Saif Mohammad].

Division of Chinese, Nanyang Technological University, April 2008.

“Fine-grained differences and similarities in meanings”.

Ubiquitous Knowledge Processing Lab, Department of Computer Science, Technical University of Darmstadt, May 2009.  
Institut für Maschinelle Sprachverarbeitung, University of Stuttgart, May 2009.  
European Media Lab, Heidelberg, March 2008.  
Computer Laboratory, University of Cambridge, January 2008.  
Department of Computer Science, University of Maryland, College Park, April 2007.  
Department of Computer Science, National University of Singapore, February 2007.  
Division of Information Studies, Nanyang Technological University, February 2007.  
Department of Computer Science, University of Colorado at Boulder, March 2006.  
Information Sciences Institute, University of Southern California, January 2006.  
Department of Computing, Macquarie University, November 2005.  
School of Computer Science and Software Engineering, Monash University, November 2005.  
Department of Computer Science and Software Engineering, University of Melbourne, November 2005.  
Simon Fraser University, March 2005.

“Academic publishing in computer science”.

University of Toronto, Department of Computer Science, Research skills seminar series, March 2007.

“Analyzing medical text to answer clinicians’ questions”.

Computational Pharmacology Group, University of Colorado Health Sciences Center, February 2006.

“Views of text-meaning in computational linguistics: Past, present, and future”.

Department of Linguistics, University of Colorado at Boulder, October 2005.

“Measures of semantic relatedness and the detection and correction of real-word spelling errors”.

School of Informatics, University of Edinburgh, August 2003.  
Sharp Laboratories Europe, Oxford, August 2003.  
Computational Linguistics Laboratory, Nara Institute of Science and Technology, July 2003.  
Department of Computer Science, Concordia University, October 2002.  
Department of General Linguistics, University of Helsinki, August 2002.  
Department of Computer Science, University of Minnesota, Duluth, April 2002.  
Department of Computer Science, University of Pittsburgh, February 2002.  
Department of Computing Science, University of Alberta, November 2001.

“Semantic things to do to text”.

Artificial Intelligence in Medicine, Inc., Symposium, Toronto, November 2002.

“Near-synonymy and the structure of lexical knowledge”.

IBM T.J. Watson Research Center, Hawthorne, NY, May 2000.  
Department of Computer Science, University of Colorado at Boulder, April 2000.  
Center for Language and Speech Processing, Johns Hopkins University, April 1999.  
Department of Computer Science, University of Rochester, March 1999.  
Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, February 1999.  
Natural Language Group, Department of Computer Science, Columbia University, February 1999.  
Center for Cognitive Science, State University of New York at Buffalo, January 1999.  
Microsoft Research Institute, Macquarie University, April 1998.  
Department of Computer Science, Monash University, April 1998.  
Cognitive Science “Distinguished Lecturer” Series, Carleton University, February 1998.

“Repairing conversational misunderstandings and non-understandings”

Department of Computer Science, University of Manitoba, January 1997.  
Department of Computer Science, Monash University, May 1996.  
Microsoft Research Institute, Macquarie University, May 1996.  
Department of Linguistics, The Ohio State University, March 1996.  
Center for Cognitive Science, State University of New York at Buffalo, February 1994.  
Computing Research Laboratory and Department of Computer Science, New Mexico State University, March 1993.

“Research in computational models of style” (with Chrysanne DiMarco)

Gesellschaft für Mathematik und Datenverarbeitung, Institut für integrierte Publikations- und Informationssysteme, Darmstadt, November 1994.  
 Forschungsinstitut für anwendungsorientierte Wissensverarbeitung, Ulm, November 1994.  
 Microsoft Corporation, Redmond, Washington, July 1993

“Writing and publishing in computer science”.

University of Waterloo, Department of Computer Science, Research skills seminar series, November 1992; repeated November 1993, November 1994, November 1995, November 1997.

“Ontological assumptions in knowledge representation”.

Cognitive Science colloquium, University of Waterloo, April 1995.  
 Department of Computer Science, Monash University, May 1993.  
 Department of Computer Science, Carleton University, March 1992.  
 Cognitive Science Group, Scarborough College, University of Toronto, February 1992.  
 Department of Computer Science, University of Chicago, February 1992.  
 Computing Laboratory, University of Cambridge, November 1991.  
 Corporate Research and Development Center, General Electric Company, Schenectady, NY, March 1991.  
 Department of Computer Science, State University of New York at Buffalo, October 1990.  
 Department of Computer Science, Virginia Polytechnic Institute and State University, April 1990.  
 Department of Computer Science, University of Rochester, January 1990.  
 Department of Computer Science, Carnegie Mellon University, December 1989.  
 Department of Computer Science, Brown University, October 1989.

“Knowledge representation problems for natural language understanding”.

IBM T.J. Watson Research Center, Hawthorne, NY, September 1989.  
 Institut für Wissensbasierte Systeme, IBM Deutschland, Stuttgart, July 1989.  
 MCC (Microelectronics and Computer Technology Corporation), Austin, Texas, March 1989.  
 Information Sciences Institute, University of Southern California, February 1989.  
 Department of Computer Science, New York University, November 1988.  
 Wiener Workshop Wissensbasierte Sprachverarbeitung, Österreichische Gesellschaft für Artificial Intelligence, August 1988.

“Presuppositions as beliefs”.

School of Cognitive Studies, University of Sussex, May 1988.

“Computational models of ambiguity resolution”.

Department of Computer Science, University of Glasgow, June 1988.  
 School of Cognitive Studies, University of Sussex, May 1988.  
 Department of Computer Science, University of Essex, May 1988.  
 Computing Laboratory, University of Cambridge, May 1988  
 Centre for Cognitive Science, University of Edinburgh, April 1988  
 Conference on semantic ambiguity, Adelphi University, April 1988.  
 Division of Life Sciences, Scarborough College, University of Toronto, March 1988.

“The resolution of natural language ambiguity”.

Department of Artificial Intelligence, University of Edinburgh, April 1988.  
 Department of Computer Science, State University of New York at Buffalo, February 1988.

“Where parsing preferences come from”.

Annual meeting of the Artificial Intelligence and Robotics Programme, Canadian Institute for Advanced Research, Banff, September 1986.

“Computational ambiguity resolution”.

Computing Research Laboratory, New Mexico State University, September 1985.  
 Department of Psychology, University of Rochester, March 1985.  
 Department of Computer Science, University of Toronto, March 1985.

“Semantic interpretation against ambiguity”.

Department of Electrical Engineering and Computer Science, Northwestern University, February 1987.  
 Department of Computer Science, Australian National University, June 1985.  
 Department of Psychology, Monash University, June 1985.  
 Department of Computer Science, Monash University, June 1985.  
 IBM T.J. Watson Research Center, Yorktown Heights, April 1985.  
 Bell Communications Research, April 1985.  
 AT&T Bell Laboratories, April 1985.  
 Department of Computing and Information Science, Queen’s University, February 1985.  
 Linguistics Programme and Department of Computer Science, University of Manitoba, November 1984.  
 Department of Computer Science, Indiana University, November 1984.  
 Artificial Intelligence Symposium, Department of Computer Science, Columbia University, November 1984.  
 Department of Computer Science, University of Waterloo, June 1984.  
 McLuhan Centre for Culture and Technology, University of Toronto, April 1984.

“Word sense and case slot disambiguation in formal semantics for natural language understanding”.

Department of Computer Science, Carnegie-Mellon University, April 1983.

Department of Computer Science, University of Toronto, January 1983.

Department of Computer Science, University of Rochester, December 1982.

## **Other professional activities**

---

### **Industry involvement**

Adviser, Sonic Boomerang Inc., Toronto, 2000–04.

Commercialized technology for determining sentiment towards stocks in chat and message boards.

Co-founder, Inkpot Software Inc., Toronto and Waterloo, 1997–99.

Commercialized technology that used natural-language generation techniques for personalizing health information for individual consumers and patients.

### **Editorships**

Series Editor, Synthesis Lectures on Human Language Technologies, Morgan and Claypool Publishers, 2007– .

Book review editor and editorial board member, *Computational Linguistics*, 1985–2014.

Editor, *Canadian Artificial Intelligence*, 1984–1987. Editor emeritus and book review editor, 1988–1999.

### **Associate editor or editorial board member**

*Computational Intelligence*, 2003–10.

*IEEE Expert*, 1992–96.

*Machine Translation*, 1989–2007.

### **Offices in professional organizations, etc**

Co-chair, Professional Conduct Committee, Association for Computational Linguistics, 2018–2022.

Treasurer [elected, international level], Association for Computational Linguistics, 2008–2017 [two five-year terms]. Past Treasurer, 2018.

Past Chair, Executive Board, North American Chapter of the Association for Computational Linguistics (NAACL), 2006–2009. Previously, elected chair, 2004–2005; elected member 2001–2002, re-elected 2003–2004.

Chair, Human Language Technology Advisory Board [oversees the Human Language Technology conference series], 2002–2009.

Information Officer, Association for Computational Linguistics Special Interest Group on the Lexicon (ACL SIGLEX), 2000–2003.

### Conference or workshop organizer or organizing committee member

Co-organizer, Dagstuhl seminar “Natural Language Argumentation: Mining, Processing, and Reasoning over Textual Arguments”, 18–22 April 2016.

Co-organizer, 2nd Workshop on Argumentation Mining, Denver, June 2015.

Area Chair, \*SEM 2015: The 4th Joint Conference on Lexical and Computational Semantics, Denver, June 2015 [coordinated reviews by 11 reviewers of 13 papers, and made final decisions].

Organizer, Workshop on Computational Approaches to Synonymy, at the Symposium on Rethinking Synonymy (organized by SKY, the Linguistic Association of Finland), Helsinki, October 2010.

Organizer, Second Workshop on Text Meaning and Interpretation, at the 42nd Meeting of the Association for Computational Linguistics, Barcelona, July 2004 (with Sergei Nirenburg).

Organizer, Workshop on Text Meaning, at the combined Human Language Technology Conference and 3rd Meeting of the North American Chapter of the Association for Computational Linguistics, Edmonton, May 2003 (with Sergei Nirenburg).

Program committee member and area chair, First meeting of the North American Chapter of the Association for Computational Linguistics, Seattle, April 2000.

Organizer, 9th International Workshop on Natural Language Generation, Niagara-on-the-Lake, Ontario, August 1998 (with Chrysanne DiMarco and Eduard Hovy).

Committee member, Canadian Society for Computational Studies of Intelligence, Conference, Toronto, May 1998.

Sub-chair, 15th International Conference on Computational Linguistics, Kyoto, August 1994.

Co-chair, Natural language sub-committee, American Association for Artificial Intelligence, National Conference, 1991.

Program committee member, Sixth annual conference, University of Waterloo Centre for the *New Oxford English Dictionary*, 1990

Member, Natural language sub-committee, American Association for Artificial Intelligence, National Conference, 1990.

Program committee, International Conference on Computers and the Humanities and conference of the Association for Literary and Linguistic Computing, Toronto, June 1989.

Program committee, Computers in the Humanities: Today’s Research, Tomorrow’s Teaching, Toronto, April 1986.

Program committee, Canadian Society for Computational Studies of Intelligence, Conference, 1986.

Program committee, Workshop on Theoretical Approaches to Natural Language, Halifax, May 1985.



## Other major activities

Member, Discovery Grant Evaluation Group 1507 (Computer Science), Natural Sciences and Engineering Research Council of Canada, 2014–2017.

Project advisor, *Data Mining the 1918 Influenza Pandemic*, Virginia Polytechnic Institute and State University, 2012–14.

## Refereeing, reviewing, etc

### 2021

Panelist, Deutsche Forschungsgemeinschaft (DFG), review paper on Methodologies of Artificial Intelligence, May 2021 [5 research proposals, with in-person discussion panel].  
Reviewer, 59th Annual Meeting of the Association for Computational Linguistics, July 2021 [3 papers].

### 2020

Reviewer, 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2020) [3 papers].  
Reviewer, 8th International Conference on Computational Models of Argument (COMMA 2020), September 2020 [3 papers].  
Reviewer, 58th Annual Meeting of the Association for Computational Linguistics, July 2020 [5 papers].  
Panelist, Deutsche Forschungsgemeinschaft (DFG), review paper on Methodologies of Artificial Intelligence, April 2020 [11 research proposals, with in-person discussion panel].  
Reviewer, Qatar National Research Fund, April 2020 [1 research proposal].  
Reviewer, Second Workshop on Trolling, Aggression and Cyberbullying (TRAC-2), Marseille, May 2020 [2 papers].  
Reviewer, Workshop on Resources and Processing of Linguistic, Para-Linguistic and Extra-Linguistic Data from People with Various Forms of Cognitive / Psychiatric / Developmental Impairments (RaPID-2020), Marseille, May 2020 [1 paper].  
Reviewer, Discovery Grant proposals, NSERC [4 proposals].

### 2019

Member, Committee to decide on the best resource paper award, Conference on Empirical Methods in Natural Language Processing, Hong Kong, November 2019 [5 papers].  
Reviewer, Qatar National Research Fund, May 2019 [2 research proposals].  
Reviewer, 6th Workshop on Argument Mining, Florence, August 2019 [3 papers].  
Panelist, Deutsche Forschungsgemeinschaft (DFG), review paper on Computational Literary Studies, May 2019 [2-day in-person panel reviewing 33 grant proposals].  
Reviewer, *BMC Bioinformatics*, May 2019 [1 paper].  
Reviewer, *IEEE/ACM Transactions of Audio, Speech, and Language Processing*, May 2019 [1 paper].  
Reviewer, Workshop on Speech and Language Processing for Assistive Technologies, Minneapolis, June 2019 [1 paper].  
Reviewer, 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, Minneapolis, June 2019 [2 papers].  
Reviewer, Israel Science Foundation [1 proposal].  
Reviewer, Discovery Grant proposals, NSERC [2 proposals].  
Reviewer, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Minneapolis, June 2019 [4 papers].  
Reviewer, Australian Research Council [1 grant proposal]

### 2018

Reviewer, *IEEE/ACM Transactions of Audio, Speech, and Language Processing*, December 2018 [1 paper].  
Reviewer, Deutscher Akademischer Austauschdienst [1 grant proposal].  
Reviewer, 2nd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2018), Santa Fe, August 2018 [2 papers].  
External reviewer for faculty promotion file.  
Reviewer, 7th International Conference on Computational Models of Argument, Warsaw, September 2018 [2 papers].  
Reviewer, 6th International Workshop on Natural Language Processing for Social Media, Melbourne, July 2018 [1 paper].  
Reviewer, Deutsche Forschungsgemeinschaft, April 2018 [1 grant proposal].  
Reviewer, 27th International Conference on Computational Linguistics (COLING 2018), Santa Fe, August 2018 [2 papers].  
Reviewer, Workshop on Computational Linguistics and Psychology, Melbourne, July 2018 [2 papers].  
Reviewer, *Language Resources and Evaluation*, April 2018 [1 paper].  
Reviewer, grant proposal, Israel Science Foundation.  
Reviewer, 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, July 2018 [3 papers].  
External reviewer for faculty promotion file.  
Reviewer, *Computational Linguistics* [1 paper].  
Reviewer, 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, New Orleans, June 2018 [3 papers].

### 2017

External reviewer for faculty tenure and promotion file.  
Reviewer, Discovery Grant proposal, NSERC, December 2017 [1 proposal].  
Reviewer, 5th International Workshop on Natural Language Processing for Social Media, Boston, December 2017 [1 paper].  
Reviewer, Swiss National Science Foundation [1 proposal].  
Reviewer, *Frontiers in Digital Humanities* [1 paper].  
Reviewer, 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, September 2017 [5 papers].

Reviewer, Workshop on Computational Linguistics and Psychology, Vancouver, August 2017 [1 paper, 1 abstract].  
 Reviewer, Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities, Vancouver, August 2017 [2 papers].  
 Reviewer, 2nd Workshop on Automated Semantic Analysis of Information in Legal Texts, London, June 2017 [1 paper].  
 Reviewer, Einstein Foundation, Berlin [1 grant proposal].  
 Reviewer, *Argument and Computation* [1 paper].  
 Reviewer, *Language Resources and Evaluation* [2 papers].  
 Reviewer, papers for four workshops at the 15th Conference of the European Chapter of the Association for Computational Linguistics, January–February 2017 [4 papers].

## 2016

Panel reviewer, National Science Foundation, December 2016 [6 proposals, 1.5-day panel].  
 Reviewer, Workshop on Natural Language Processing and Computational Social Science, Austin, November 2016 [2 abstracts].  
 Reviewer, Workshop on Uphill Battles in Language Processing: Scalina Early Achievements to Robust Methods, Austin, November 2016 [2 abstracts].  
 Reviewer, 4th International Workshop on Natural Language Processing for Social Media, Austin, November 2016 [2 papers].  
 Reviewer, 26th International Conference on Computational Linguistics, Osaka, December 2016 [3 papers].  
 Reviewer, Conference on Empirical Methods in Natural Language Processing, Austin, November 2016 [5 papers].  
 Reviewer, CRD grant proposal, NSERC, July 2016.  
 Reviewer, 4th International Conference on Statistical Language and Speech Processing, Pilsen, October 2016 [3 papers].  
 Reviewer, 3rd Workshop on Argument Mining, Berlin, August 2016 [3 papers].  
 Reviewer, 6th International Conference on Computational Models of Argument, Potsdam, September 2016 [2 papers].  
 Reviewer, Einstein Foundation, Berlin [1 grant proposal].  
 Reviewer, The 5th Joint Conference on Lexical and Computational Semantics (“Sem”), Berlin, August 2016 [1 paper].  
 Reviewer, 54th Annual Meeting of the Association for Computational Linguistics, Berlin, August 2016 [6 papers].  
 Reviewer, 5th Workshop on Computational Linguistics for Literature, San Diego, June 2016 [2 papers].  
 Reviewer, Language Resources and Evaluation Conference, Portorož, May 2016 [4 1500-word abstracts].  
 Reviewer, *Journal of Cognitive Science* [1 paper].

## 2015

Reviewer, *Nature* [1 paper].  
 Reviewer, *INFOCOMP Journal of Computer Science* [1 paper].  
 Reviewer, *Journal of Computational Research Replication* [1 paper].  
 Reviewer, *ACM Computing Surveys* [1 long paper].  
 External reviewer for two faculty promotion files.  
 Reviewer, 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, September 2015 [3 long papers, 3 short papers].  
 Assessor, Nazarbayev University (Kazakhstan) Research Review [3 research proposals].  
 Reviewer, *Computational Linguistics* [1 very long paper].  
 Reviewer, 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (ACL–IJCNLP 2015), Beijing, July 2015 [3 long papers, 3 short papers].  
 Reviewer, Second Computational Linguistics and Clinical Psychology Workshop, Denver, June 2015 [3 papers].  
 Reviewer, 2015 Conference of the North American Chapter of the Association for Computational Linguistics, Denver [5 papers].

## 2014

Reviewer, *Language Resources and Evaluation*, December 2014 [1 paper].  
 Reviewer, *Literary and Linguistic Computing*, December 2014 [1 paper].  
 Reviewer, *Artificial Intelligence*, November 2014 [1 paper].  
 Reviewer, Australasian Language Technology Workshop, Melbourne, November 2014 [3 papers].  
 Reviewer, *Computer Speech and Language*, July 2014 [1 paper].  
 Reviewer, 2014 Conference on Empirical Methods in Natural Language Processing, Doha, October 2014 [4 papers].  
 Reviewer, Workshop on Cognitive Aspects of the Lexicon, Dublin, August 2014 [2 papers].  
 Assessor, Australian Research Council [1 grant proposal].  
 Reviewer, 25th International Conference on Computational Linguistics, Dublin, August 2014 [4 papers].  
 Reviewer, *Interaction Studies*, special issue on Mental Model Ascription by Intelligent Agents [2 papers], March 2014.  
 Reviewer, 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, June 2014 [3 long and 2 short papers].  
 Reviewer, *Computer Speech and Language*, February 2014 [1 paper].  
 Reviewer, EACL Workshop on Computational Approaches to Causality in Language, Gothenburg, April 2014 [1 paper].

## 2013

Reviewer, *Language Resources and Evaluation*, December 2013 [1 paper].  
 Reviewer, Fellowship proposal, Fernand Braudel–IFER programme, Fondation Maison de Sciences de l’Homme, November 2013.  
 Reviewer, 9th Language Resources and Evaluation Conference, Reykjavik, May 2014 [4 short papers].  
 Reviewer, 14th Conference of the European Chapter of the Association for Computational Linguistics, Gothenburg, April 2014 [3 long and 2 short papers].  
 Reviewer, 11th Annual Workshop of the Australasian Language Technology Association, Brisbane, December 2013 [1 paper].  
 Reviewer, Conference on Recent Advances in Natural Language Processing, Hissar, Bulgaria, September 2013 [3 papers].  
 Reviewer, *Computational Linguistics* [1 short squib].  
 Reviewer, The 7th Linguistic Annotation Workshop and Interoperability with Discourse, Sofia, August 2013 [2 papers].  
 Sub-reviewer, *Scientia Horticulturae* [1 paper].  
 Reviewer, International Conference of the German Society for Computational Linguistics and Language Technology, Darmstadt, Germany, September 2013 [2 papers].  
 Reviewer, 51st Annual Meeting of the Association for Computational Linguistics, Sofia, August 2013 [2 long papers, 1 short paper].  
 Reviewer, 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, June 2013 [2 papers].

Assessor, Australian Research Council [1 grant proposal].  
 Reviewer, U.S. National Science Foundation [1 grant proposal].

## 2012

External reviewer for two faculty appointment, tenure, and/or promotion files.  
 Reviewer, Discovery Grant proposals, NSERC, December 2012 [2 proposals].  
 Reviewer, *Computational Linguistics* [1 paper].  
 Reviewer, *Dialogue and Discourse* [1 paper].  
 Reviewer, 3rd Workshop on Cognitive Aspects of the Lexicon, 24th International Conference on Computational Linguistics, Mumbai, December 2012 [2 papers].  
 Reviewer, 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Jeju, Korea, July 2012 [4 papers].  
 Assessor, Australian Research Council [1 grant proposal].  
 Reviewer, First Joint Conference on Lexical and Computational Semantics, Montreal, June 2012 [1 paper].  
 Reviewer, 2012 Conference on Empirical Methods in Natural Language Understanding and Conference on Natural Language Learning (EMNLP-CoNLL-2012), Jeju, Korea, July 2012 [3 long papers, 1 short paper].  
 Reviewer, 9th Extended Semantic Web Conference, Crete, May 2012 [2 papers].  
 Reviewer, 2012 Informatics for Phylogenetics, Evolution, and Biodiversity (iEvoBio 2012), Ottawa, July 2012 [1 abstract].  
 Reviewer, 50th Annual Meeting of the Association for Computational Linguistics, Jeju, Korea, July 2012 [4 long and 4 short papers].  
 Reviewer, Workshop on Speech and Language Processing for Assistive Technologies, Montreal, June 2012 [2 papers].  
 Reviewer, 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Montreal, June 2012 [2 papers].  
 Reviewer, Student Research Workshop, 13th Conference of the European Chapter of the Association for Computational Linguistics, Avignon, April 2012 [3 papers].  
 Reviewer, *Computational Linguistics* [1 paper].

## 2011

External reviewer for three faculty appointment, tenure, and/or promotion files.  
 Reviewer, Discovery Grant proposals, NSERC, December 2011 [2 proposals].  
 Reviewer, *IEEE Internet Computing* [1 paper].  
 Reviewer, *Information Fusion* [1 paper].  
 Reviewer, 13th International Conference on Intelligent Text Processing and Computational Linguistics, New Delhi, March 2012 [4 papers, 3 collaboratively with students].  
 Reviewer, 8th Discourse Anaphora and Anaphora Resolution Conference, Faro, October 2011 [2 abstracts].  
 Assessor, Australian Research Council [8 grant proposals].  
 Assessor, Canada Council for the Arts, Killam Research Fellowship [1 application].  
 Reviewer, TextInfer Workshop on Textual Entailment, Edinburgh, July 2011 [1 paper].  
 Reviewer, 2011 Conference on Recent Advances in Natural Language Processing (RANLP), September 2011, Hissar, Bulgaria [5 papers].  
 Reviewer, 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, July 2011 [4 papers].  
 Reviewer, 12th Conference of the Pacific Association for Computational Linguistics, Kuala Lumpur, July 2011 [3 papers].  
 Reviewer, 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon, June 2011 [6 papers].  
 Reviewer, Discovery Grant proposals, NSERC, January 2011 [2 proposals].

[Truncated at 2010]

## Courses taught

---

### Undergraduate courses

Computer Science D03H — *Social Impact of Information Technology* [known as Physical Sciences D03H — *Computers in Contemporary Society* until Fall 2003]: Spring 1997, Spring 1999, Winter 2000, Fall 2000, Fall 2001, Winter 2002, Fall 2002, Fall 2003 (two sections), Fall 2004, Fall 2006, Winter 2007, Fall 2007, Fall 2008, Fall 2009, Fall 2010, Winter 2012, Fall 2013, Winter 2014, Fall 2014, Fall 2016, Fall 2017, Winter 2018, Fall 2018.

Computer Science 401H (cross-listed as 2511) — *Natural Language Computing*: Winter 2000, Winter 2001.

Computer Science C24H — *Principles of Programming Languages*: Spring 1987, Spring 1988, Spring 1990, Spring 1992, Spring 1993, Spring 1994.

Computer Science 158H or A68H — *Computer Applications*: Spring 1985, Spring 1986.

Computer Science 148H or A58H — *Introduction to Computer Science* (formerly *Introduction to Computing*): Fall 1984, Fall 1986, Fall 1987, Fall 1988, Fall 1989, Fall 1991, Fall 1992, Fall 1993,

Fall 1994, Spring 1995, Spring 1996.

Computer Science A06H — *Introduction to Computer Programming*: Fall 1996, Fall 1998.

### Graduate courses

Computer Science 2501H (cross-listed as 485 from 1986) — *Computational Linguistics* (formerly *Introduction to Computational Linguistics*): Fall 1984, Fall 1985, Fall 1986, Fall 1987, Fall 1988, Fall 1989 (with Janyce Wiebe), Fall 1991, Fall 1992, Fall 1993 (with Jeffrey Siskind), Fall 1994 (with Jeffrey Siskind), Fall 1995 (with Jeffrey Siskind) (simultaneously offered at the University of Waterloo as CS 684 by videoconference), Fall 1996, Fall 1998 (as 2501 only), Fall 1999, Fall 2007, Fall 2008, Fall 2009, Fall 2010, Fall 2011, Fall 2013, Fall 2014.

Computer Science 2528H — *Advanced Computational Linguistics* (formerly *Topics in Computational Linguistics*): Spring 1984, Spring 1985, Spring 1986, Spring 1987, Spring 1988, Spring 1989, Spring 1990 (with Janyce Wiebe), Spring 1992 (with Dekai Wu), Spring 1993, Spring 1994 (with Jeffrey Siskind), Spring 1996 (also offered as University of Waterloo course CS 786), Spring 1997 (also offered as University of Waterloo course CS 786), Spring 1999, Winter 2000, Winter 2001, Winter 2002, Fall 2002, Winter 2004, Fall 2004, Winter 2007, Fall 2017, Fall 2018.

### Supervision

---

Career student numbers		
	<i>In progress</i>	<i>Completed</i>
Master's	3	53
PhD	3	23
Postdoctoral fellows	1	8
Guest students	0	5
Others	0	1

### Master's students

Zihang Fu, *Question-to-question semantic similarity for question answering* [MScAC internship project], May–December 2019.

Paul Briggs, *Sentiment analysis in dialogue systems*, [MScAC internship project] May–December 2019.

Chantal Shaib, *Exploring Communication in Medical Contexts* [MScAC internship project], May–December 2019.

Krishnapriya Vishnubhotla, *A stylometric investigation of character voices in literary fiction*, September 2017–April 2019.

Jindong Liu, *Entity resolution on financial documents* [MScAC internship project], May–December 2018.

Navid Kaihanirad, *Question answering: Query on demand with natural language* [MScAC internship project], May–December 2018.

Kawin Ethayarajh, *Essays on Understanding Analogies and Associations in Word Embedding Spaces*, September 2017–March 2019.

Misha Schwartz, *Word segmentation using grouped semantic networks*, September 2016–December 2018.

- Jeff Pinto, *A privacy-preserving natural language clinical information extraction pipeline*, September 2017–December 2018.
- Mohamed Abdalla, *Lowering the cost of improved cross-lingual sentiment analysis*, September 2016–April 2018.
- Muuu Wambua, *Identifying opposing views in online discourse*, September 2016–April 2018. [Co-supervised with Frank Rudzicz.]
- Elías Jónsson, *Evaluating legal question answering* [MScAC internship project], May–December 2016.
- Vanessa Queiroz Marinho, *Complex networks for authorship attribution*, Guest student from University of São Paulo, São Carlos, March–August 2016.
- Krish Perumal, *Semi-supervised and unsupervised methods for categorizing posts in web discussion forums*, September 2014 – January 2016.
- Wenjie (Richie) Zi, *Automatic spam alert system for online social media and Linking aspects to mentions via shortest path dependencies for sentiment analysis* [MScAC internship project], May–December 2015.
- David Fernig, *Computational analysis of amateur fiction*, September 2014 – August 2015 [*withdrew*].
- L. Amber Wilcox-O’Hearn, September 2002–August 2005, reinstated April 2014. *Trigram-based noisy channel models of real-word spelling and grammatical correction*.
- Yunxing (Henry) Zhang, *Analyzing user reviews of mobile applications using natural language processing*, MScAC internship (at Electronic Arts, Kitchener, ON), May–December 2013.
- Colin Morris, *Identifying online sexual predators by SVM classification with lexical and behavioral features*. January 2013.
- Vanessa Wei Feng, *Classifying arguments by scheme*. November 2010.
- Xuan Le, *Longitudinal detection of dementia through lexical and syntactic changes in writing*. January 2010.
- Meghana Marathe, *Lexical chains using distributional measures of concept distance*. May 2009.
- Tong Wang, *Extracting synonyms from dictionary definitions*. January 2009.
- Yaroslav Riabinin, *Computational identification of ideology in text: A study of Canadian parliamentary debates*. January 2009.
- Faye Baron, *Identifying non-compositional idioms in text using WordNet synsets*, October 2007. [Co-supervised with Suzanne Stevenson from January 2007.]
- Jianhua (Jane) Li, *Modelling semantic knowledge for a word prediction task*, October 2006.
- Benjamin Bartlett. *Finding paraphrases using PNRule*, September 2006.
- Tristan Miller. *Generating coherent extracts of single documents using latent semantic analysis*, January 2003.
- Afsaneh Fazly [Fazli]. *The use of syntax in word completion utilities*, January 2002.
- Eric Joanis, 1998–September 2001 [transferred to Suzanne Stevenson].

- Vivian Tsang, 1999–July 2000 [transferred to Suzanne Stevenson].
- Neil Graham. *Automatic detection of authorship changes within single documents*, January 2000.
- Mary Ellen Foster. *Automatically generating text to accompany information graphics.*, August 1999.
- Alexander Budanitsky, 1996–1999. *Lexical semantic relatedness and its application in natural language processing*. [PhD candidacy, equivalent to MSc.]
- Vlado Kešelj [Department of Computer Science, University of Waterloo], *Multi-agent systems for Internet information retrieval using natural language processing*, September 1997.
- Kimberley Parsons [Department of Computer Science, University of Waterloo; co-supervised with Chrysanne DiMarco], *An authoring tool for customizable documents*, May 1997.
- Melanie Baljko, *Ensuring stylistic congruity in collaboratively written text: Requirements analysis and design issues*, May 1997.
- Angela Glover, *Automatically detecting stylistic inconsistencies in computer-supported collaborative writing*, January 1996.
- Daniel Ansari, *Deriving procedural and warning instructions from device and environment models*, June 1995.
- David St-Onge, *Detecting and correcting malapropisms with lexical chains*, March 1995.
- Robert Lizée, *A study of natural language quantification and anaphora through families of sets and binary relations*, February 1995.
- Daniel Marcu, *A formalism and an algorithm for computing pragmatic inferences and detecting infelicities*, September 1994.
- Linda Peto [co-supervised by Dekai Wu], *A comparison of two smoothing methods for word bigram models*, April 1994.
- Philip Edmonds, *A computational model of collaboration on reference in direction-giving dialogues*, October 1993.
- Peter Heeman, *A computational model of collaboration on referring expressions*, October 1991.
- Carson Schütze [Department of Linguistics, co-supervised by Peter Reich], *Grammaticality judgments and linguistic methodology*, September 1991.
- Julie Payette [Department of Electrical Engineering], *Intelligent computer-assisted instruction in syntactic style*, October 1990.
- Nathalie Japkowicz [co-supervised with Janyce Wiebe], *The translation of basic topological prepositions from English into French*, October 1990.
- Daniel Lyons, *A compositional semantics for focusing subjuncts*, September 1989.
- Mark Ryan [co-supervised with Chrysanne DiMarco], *The computational codification of the semantic aspects of style*, August 1989.
- Toby Walker, 1988–1990 [discontinued studies].
- Jane Morris, *Lexical cohesion, the thesaurus, and the structure of text*, October 1988.
- Mara Miezitis, *Generating lexical options by matching in a knowledge base*, October 1988.

Mark Catt, *Intelligent diagnosis of ungrammaticality in computer-assisted language instruction*, October 1988.

Murray Watt, *The realization of natural language with pragmatic effects*, August 1988.

Susan McRoy, *The influence of time and memory constraints on the resolution of structural ambiguity*, October 1987.

Diane Horton, *Incorporating agents' beliefs in a model of presupposition*, October 1986.

Barbara Brunson [Linguistics forum, co-supervised with Elizabeth Cowper, Department of Linguistics], *A processing model for Warlpiri syntax and implications for linguistic theory*, September 1986.

Brenda Fawcett, *The representation of ambiguity in opaque contexts*, October 1985.

Bart Selman, *Rule-based processing in a connectionist system for natural language understanding*, January 1985.

Yawar Ali, *Understanding adjectives*, January 1985.

### Doctoral students

Krishnapriya Vishnubhotla, September 2019– (in progress).

Mohamed Abdalla, Topic TBD, April 2018– (in progress) [co-supervised with Frank Rudzicz].

Serena Jeblee, *Automating disease diagnosis and cause-of-death classification from medical narratives using event extraction and temporal ordering*, September 2016–April 2021.

Nona Naderi, *Computational analysis of arguments and persuasive strategies in political discourse*, September 2013–December 2019.

Tong Wang, *Exploiting linguistic knowledge in lexical and compositional semantic models*, January 2009–September 2016.

Kathleen Fraser, *Automatic text and speech processing for the detection of dementia*, September 2011–September 2016 [co-supervised with Jed Meltzer, Dept of Psychology and Baycrest].

Muyu Zhang, Discourse parsing in Chinese, Guest student from the Harbin Institute of Technology, September 2014 – August 2015.

Erick Galani Maziero, Discourse parsing in Portuguese, Guest student from the University of São Paulo, São Carlos, Brazil, September 2014–August 2015.

Vanessa Wei Feng, *RST-Style discourse parsing and its applications in discourse analysis*, January 2011 – December 2014.

Varada Kolhatkar, *Resolving shell nouns*, September 2009 – December 2014.

Julian Brooke, *Computational approaches to style and the lexicon*, September 2009–December 2013.

Ulrich Germann, September 2003 – March 2013. *Syntax-aware phrase-based statistical machine translation from German to English*. Terminated in PhD program, left with MSc, April 2013.

Frank Rudzicz, September 2006 – April 2011. *Production Knowledge in the Recognition of Dysarthric Speech*.

Eric J.M. Smith, January 2007 – April 2010 [co-supervised with Elizabeth Cowper, Department

of Linguistics]. *Query-Based Annotation and the Sumerian Verbal Prefixes*.

Libby Barak, September 2008 – April 2010 [transferred to Suzanne Stevenson].

Hani Safadi, September – December 2008 [withdrew from program].

Saif Mohammad, September 2003 – February 2008. *Measuring semantic distance using distributional profiles of concepts*.

Xiaodan Zhu, 2003–04 [transferred to Gerald Penn].

Yun Niu, July 2001 – March 2007. *Analysis of semantic classes: Toward non-factoid question answering*.

Alexander Budanitsky, 1999 – January 2004. *Lexical semantic relatedness and its application in natural language processing*. [Withdrew before completion.]

Jane Morris [Faculty of Information Studies; co-supervised with Clare Beghtol], September 2000 – March 2006. *Readers' perceptions of lexical cohesion and lexical semantic relations in text*, February 2007.

Melanie Baljko, 1997 – July 2004. *Computational simulations of mediated face-to-face multimodal communication*. [Co-supervised with Fraser Shein.]

Natalia Modjeska [guest student from Division of Informatics, University of Edinburgh; co-supervised with Bonnie Webber], 2001–2003. *Resolving other-anaphora*.

Diana Zaiu Inkpen, 1998–2003. *Building a lexical knowledge base of near-synonym differences*.

Grzegorz (Greg) Kondrak, 1997–2002. *Algorithms for language reconstruction*.

Philip Edmonds, *Semantic representations of near-synonyms for automatic lexical choice*, September 1999.

Daniel Marcu, *The rhetorical parsing, summarization, and generation of natural language texts*, January 1998.

Stephen Green, *Automatically generating hypertext by computing semantic similarity*, August 1997.

Manfred Stede, *Lexical semantics and knowledge representation in multilingual sentence generation*, May 1996.

Susan McRoy, *Abductive interpretation and reinterpretation of natural language utterances*, July 1993.

Suzanne Stevenson, guest PhD student from the University of Maryland, 1991–94. *A competitive attachment model for resolving syntactic ambiguities in natural language parsing*.

Diane Horton, *Repair of plan mis-recognition*, 1986–1993 [discontinued studies].

Mark Ryan, 1991–92 [discontinued studies].

Judith Dick [co-supervised with Ann Schabas, Faculty of Library and Information Science], *A conceptual, case-relation representation of text for intelligent retrieval*, April 1991.

Jean-Pierre Corriveau, *Time-constrained memory for reader-based text comprehension*, January 1991.

Yves Lespérance [co-supervised with Hector Levesque as primary advisor from 1988], *A formal theory of indexical knowledge and action*, January 1991.



Kem Luther, 1990 [discontinued studies].

Chrysanne DiMarco, *Computational stylistics for natural language translation*, February 1990.

Charles Pilkington, *Connectionist models of natural language processing*, 1986–1987 [transferred to Geoff Hinton after completion of depth oral].

Edwin Plantinga, *Computational metaphor comprehension*, 1985–87. [discontinued studies].

### Undergraduates

Tanuj Ash Kumar, Natural language processing for cause-of-death prediction in Hindi, Summer 2019.

Zhaodong Yan, Character embeddings for automated cause-of-death classification from verbal autopsy, Engineering Science undergraduate thesis, September 2018–April 2019.

Parthkumar Parmar, Natural language processing for cause of death prediction in Hindi, Summer 2018.

Yoona Park, Symptom normalization for cause of death prediction using machine learning, Summer 2018.

Byung Gwan (Yohan) Lee, Computational analysis of Canadian Members of Parliament's rhetoric of public religiosity (CSC 494), Fall 2018.

Joshua Zung, Natural language processing for digital transcription of medieval Latin handwriting, Summer 2017.

Timothy Alberdingk Thijm, Using natural language processing to understand political arguments, Summer 2016.

Kawin Ethayarajh, A vector space model of narrative structure in novels (CSC 494), Winter 2016.

Gabrielle Singh Cadieux, Text classification, independent study project (CSC D94), Fall 2015.

Mohamed Abdalla, Automatic categorization of forum posts, Summer 2015.

Timothy Alberdingk Thijm, Digging into linked parliamentary data, Summer 2015.

Roman Polyovsky, Digging into linked parliamentary data, Summer 2015.

Michael Kimmins, Improving OCR for Canadian parliamentary proceedings, Summer 2014.

Nishanthan Velauthampillai, Unsupervised extraction of narrative chains and schemas, independent study project (CSC 494), Winter 2012.

Arturo Martínez Peguero, Clinical form classification, independent study project (CSC 494), Winter 2012.

Michelle Tran, Verifying author's style in Alzheimer's disease, independent study project (CSC 494), Winter 2011.

Colin Morris, Determining political beliefs from what MPs say in Parliament, NSERC Undergraduate Student Research Award, Summer 2010.

Magali Boizot-Roche, Determining political beliefs from what MPs say in Parliament, informal R.A., Winter 2010.

Jory Graham, Determining political beliefs from what MPs say in Parliament, NSERC Undergraduate Student Research Award, Summer 2009.

Youngchan Kim, Detecting markers of Alzheimer's disease in literary texts, independent study project (CSC 494), Winter 2009.

Michael Reimer, Semantics in word prediction for disabled users, NSERC Undergraduate Student Research Award, Summer 2008. [Co-supervised by Frank Rudzicz]

Matthew MacPherson, Semantics in word prediction for disabled users, NSERC Undergraduate Student Research Award, Summer 2007.

Yaroslav Riabinin, Text classification with concept distance, independent study project (CSC 494H), Winter 2007.

Ol'ga Feiguina, Computational stylometry, independent study project (CSC 494H), Winter 2004, and NSERC Undergraduate Student Research Award, Summer 2004.

James Bergstra, Named entity recognition in biomedical text, independent study project (CSC 494H), Fall 2003 – Winter 2004.

James Bergstra, Extracting information on protein interactions from biomedical text, NSERC Undergraduate Student Research Award, Summer 2003.

Ol'ga Feiguina, Extracting information from machine-readable dictionaries, research assistant and NSERC Undergraduate Student Research Award, 2002–03.

Faye Baron, Sentiment mining in text, independent study project (CSC 494H), Summer 2001.

Bhaskara Marthi, Bigrams in stylometrics, independent study project (CSC 494H), Fall 1999 – Winter 2000.

Andreas Berndt, guest student from the University of Bielefeld, Summarizer's workbench, February–April 1999.

Adrian Corduneanu, A pylonic decision-tree language model, Winter 1998–99. [Primary advisor: Frederick Jelinek, The Johns Hopkins University].

Melanie Baljko (NSERC Summer Fellow), Computational analysis of writing style, Summer 1995.

Dan Melamed, Independent study project, Spring 1992.

Noel Harding, Independent study project, Spring 1992.

Oliver Schulte (NSERC Summer Fellow), Porting Penman and Loom, Summer 1990.

Dawna Martin, Independent study projects, 1987–1989.

Steve Harris (NSERC Summer Fellow), Development of an ATN parser, Summer 1987.

Anton Geshelin (NSERC Summer Fellow), Implementing a Warlpiri parser and the Femto-Frail frame system, Summer 1986.

Barbara Brunson and Michael Stairs (NSERC Summer Fellows), Debugging and installation of ATN and Marcus parser on Lisp machines and VAX, Summer 1985.

Barbara Brunson, supervision of Linguistics 1551H (reading course) "Thematic argument structure and prepositional phrase attachment", Spring 1986.

Barbara Brunson, “Experiments on prepositional phrase attachment”, Spring 1985.

### Postdoctoral Fellows

Gelila Tilahun, July 2018 – June 2020 [co-supervised with Michael Gervers, Dept of Historical and Cultural Studies].

Kaspar Beelen, June 2014 – May 2016 [co-supervised with Christopher Cochrane, Dept of Political Science].

Julian Brooke, January 2014 – August 2015.

Alistair Kennedy, June 2012 – February 2013.

Frank Rudzicz, January–April 2011.

Paul Cook, September 2010 – June 2011.

Kinfe T. Mengistu, February 2010 – April 2011.

Dekai Wu, 1992.

Janyce Wiebe, 1989–90.

### Others

Ellen Korcovelos, Fulbright Scholar, August 2016 – April 2017.

### External dissertation or thesis examiner or committee member

External committee member, Yong Y. Auh, *Computational Analysis of Social Media texts for the Study of Patterns of Patients’ Decision Making*, Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, April 2019– .

Faculty examiner (“opponent”), Marie Dubremetz, *Detecting Rhetorical Figures Based on Repetition of Words: Chiasmus, Epanaphora, Epiphora*, Department of Linguistics and Philology, Uppsala University, January 2018.

External PhD committee member, Alla Rozovskaya, *Automated Methods for Text Correction*, Department of Linguistics, University of Illinois, November 2012–October 2013.

MASc committee member, Colleen Denyse Desaulniers Smith, *A Concept-based P300 Communication System*, Institute of Biomaterials and Bioengineering, University of Toronto, August 2012.

External PhD examiner, William Hollingsworth, *Using lexical chains to characterise scientific text*, University of Cambridge, January 2008.

External thesis proposal assessor, Jing Lin, University of Aberdeen, December 2006.

External MSc thesis examiner, Michael G. Mulcahy, *Seeking answers with lexically cohesive term expansions*, University of Limerick, April 2006.

External dissertation examiner, Ou Shiyan, *Automatic multi-document summarization using a variable-based framework*, Nanyang Technological University, Singapore, October 2005.

External doctoral committee member, Smaranda Muresan, *Learning constraint-based grammars from representative examples: Theory and applications*, Columbia University, September 2003–June 2006.

External dissertation examiner and re-examiner, Ian Edward Thomas, *An information-theoretic approach to speech understanding*, Department of Computer Science and Software Engineering, Monash University, August 2003 and June 2004.

External dissertation examiner, Hongyan Jing, *Cut-and-paste text summarization*, Department of Computer Science, Columbia University, 2001.

External dissertation examiner, Maria Milosavljevic, *Maximising the coherence of descriptions via comparison*, Department of Computing, Macquarie University, March 1999.

External doctoral committee member, Hongyan Jing, Department of Computer Science, Columbia University, 1999.

External dissertation examiner, Jian Jiang, *Lexical semantic similarity and its application to business catalogue retrieval*, Department of Management Sciences, University of Waterloo, September 1998.

External dissertation examiner, Mohsen Rais-Ghasem, *An exemplar-based account of contextual effects*, School of Computer Science, Carleton University, June 1998.

External dissertation examiner, Alpha Kamchiu Luk, *A conceptual cohesion based approach to word sense disambiguation*, Department of Computing, Macquarie University, February 1997.

External dissertation examiner, David Traum, *A computational theory of grounding in conversation*, Department of Computer Science, University of Rochester, July 1994.

External dissertation examiner, Tomasz Strzalkowski, *A theory of stratified meaning representation for natural language*, School of Computing Science, Simon Fraser University, July 1986.

## University and Community Service

---

2021–22

Associate Chair, Graduate Studies, Department of Computer Science.

2020–21

Associate Chair, Graduate Studies, Department of Computer Science.

2019–20

Associate Chair, Graduate Studies, Department of Computer Science.

Guest talk, "Computers, Language, and Artificial Intelligence", Cawthra Park Secondary School 20 February 2020.

Guest lecture, WDW 336 ("Cultural Literacy in the Digital Age"), 16 October 2019.

2018–19 [*On sabbatical January–June 2019*]

Member, Teaching Evaluation Committee for promotion of Professor Jennifer Campbell, Department of Computer Science.

Member Graduate Affairs Committee, Department of Computer Science, responsible for scholarships and awards rankings.

2017–18

Member, Research Advisory Board, UTSC.

Member Graduate Affairs Committee, Department of Computer Science, responsible for scholarships and awards rankings.

Member, Continuing Status Review Committee, Department of Computer and Mathematical Sciences.

Member, Senior Promotions Committee, Department of Computer and Mathematical Sciences.

2016–17

Member, Research Advisory Board, UTSC.

Member, Graduate Awards Review Committee, School of Graduate Studies.

Member, SGS Orals Committee, Department of Computer Science.

Member, Awards Committee, Department of Computer Science.

Member, University of Toronto Tribunal.

2015–16 [*On sabbatical July–December 2015*]

Member, SGS Orals Committee, Department of Computer Science.

Member, Awards Committee, Department of Computer Science.

Member, University of Toronto Tribunal.

## 2014–15

Member, Graduate Awards Review Committee, School of Graduate Studies.  
 Member, Awards Committee, Department of Computer Science.  
 Advisor, Admissions Committee, MScAC program, Department of Computer Science.  
 Member, University of Toronto Tribunal.

## 2013–14

Chair, Admissions Committee, MScAC program, Department of Computer Science.  
 Member, Awards Committee, Department of Computer Science.

2012–13 [*On sabbatical*]

## 2011–12

Member, Chair Search Committee, Department of Computer and Mathematical Sciences, UTSC.  
 Chair, Graduate Admissions Committee, Department of Computer Science [overseeing graduate admissions process].  
 Departmental representative, Library Advisory Committee, UTSC.  
 Member, University of Toronto Tribunal [several hearings per year].  
 Guest lecture (2 hours) on natural language processing, MHI 2001 (Master of Health Informatics program, Department of Health Policy, Management, and Evaluation), November 2011.

## 2010–11

Chair, Graduate Admissions and Graduate Visit Day Committee, Department of Computer Science [overseeing graduate admissions process and the organization of Grad Visit Day].  
 Guest lecture (2 hours) on summarization, CSC 401 (Natural Language Computing), Department of Computer Science, April 2011.  
 Guest lecture (2 hours) on artificial intelligence and the coming singularity, PSC D02 (Current Questions in Mathematics and Science, Department of Physical and Environmental Sciences, UTSC), March 2011.  
 Member, University of Toronto Tribunal [several hearings per year].  
 Panel chair, Ontario Graduate Scholarship Program, Province of Ontario.  
 Guest lecture (2 hours) on natural language processing, MHI 2001 (Master of Health Informatics program, Department of Health Policy, Management, and Evaluation), November 2010.

## 2009–10

Chair, Graduate Admissions and Graduate Visit Day Committee, Department of Computer Science [overseeing graduate admissions process and the organization of Grad Visit Day].  
 Member, University of Toronto Tribunal [several hearings per year].  
 Panel chair, Ontario Graduate Scholarship Program, Province of Ontario.  
 Guest lecture (2 hours) on natural language processing, MHI 2001 (Master of Health Informatics program, Department of Health Policy, Management, and Evaluation), November 2009.  
 Guest lecture (2 hours) on artificial intelligence and the coming singularity, PSC D01 (The Physical Sciences in Contemporary Society, Department of Physical and Environmental Sciences, UTSC), November 2009.

## 2008–09

Chair, Graduate Admissions and Graduate Visit Day Committee, Department of Computer Science [overseeing graduate admissions process and the organization of Grad Visit Day].  
 Member, University of Toronto Tribunal [several hearings per year].  
 Panel chair, Ontario Graduate Scholarship Program, Province of Ontario.  
 Local (Toronto) interviewer of undergraduate applicants to Brown University (Brown Alumni Schools Committee).  
 Member, University of Toronto Graduate Education Council [elected].  
 Member, Undergraduate Committee, Department of Computer Science, as representative of Department of Computing and Mathematical Sciences, UTSC.  
 Guest lecture (2 hours) on natural language processing, MHI 2001 (Master of Health Informatics program, Department of Health Policy, Management, and Evaluation), October 2008.  
 Guest lecture (2 hours) on artificial intelligence and the coming singularity, PSC D01 (The Physical Sciences in Contemporary Society, Department of Physical and Environmental Sciences, UTSC), November 2008.

## 2007–08

Local (Toronto) interviewer of undergraduate applicants to Brown University (Brown Alumni Schools Committee).  
 Panel chair, Ontario Graduate Scholarship Program, Province of Ontario.  
 Member, PTR / merit committee, Department of Computer and Mathematical Sciences, UTSC.  
 Member, Graduate admissions committee, Department of Computer Science.  
 Member, University of Toronto Graduate Education Council [elected].

## 2006–07

Member, Principal's Research Award Committee, UTSC.  
 Member, Sponsorship Committee, North American Computational Linguistics Olympiad.  
 Local (Toronto) interviewer of undergraduate applicants to Brown University (Brown Alumni Schools Committee).  
 Panel member (reviewer), Ontario Graduate Scholarship applications, Province of Ontario.  
 Member, tenure committee for Sheila McIlraith, Department of Computer Science.  
 Member, editorial board for '@DCS', Department of Computer Science.  
 Member, University of Toronto Graduate Education Council [elected].  
 Member, Graduate affairs committee, Department of Computer Science.  
 Ad hoc member, Graduate admissions committee, Department of Computer Science.  
 Member, Recruiting committee, Department of Computer and Mathematical Sciences, UTSC.

[*Truncated at 2005–06*]

2021 July 2