Bibliography
Papers on Summarization

Dragomir Radev and Erin Doumpoulaki

October 29, 2003

This document contains a rather incomplete bibliography of research in text summarization. The list of references was compiled using materials provided by Branimir Boguraev, Gael Dias, Hongyan Jing, Mark Kantrowitz, Inderjeet Mani, Tim Ostler, Hong Qi, Horacio Saggion, Simone Teufel, and others.

References

[1] Jose Abracos and Gabriel Pereira Lopes. Statistical Methods for Retrieving Most Significant Paragraphs in Newspaper Articles. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, July 11 1997.

[2] Alfred Aho, Shih-Fu Chang, Kathleen R. McKeown, Dragomir R. Radev, John Smith, and Kazi Zaman. Columbia Digital News System: An Environment for Briefing and Search over Multimedia Information. In *Proceedings of the IEEE International Conference on Advances in Digital Libraries*, Washington, DC, 1997. http://www.cs.columbia.edu/~radev/publication/adl97.ps.

[3] Akiko Aizawa. Analysis of Source Identified Text Corpora: Exploring the Statistics of Reused Text and the Authorship. In *Proceedings of the 41th Meeting of the Association for Computational Linguistics*, Sapporo, Japan, 2003.

[4] Laura Alonso Alemany and Maria Fuentes Fort. Integrating Cohesion and Coherence for Automatic Summarization. In *Proceedings of the 11th Meeting of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary, April 12–17 2003.

[5] James Allan, Rahul Gupta, and Vikas Khandelwal. Temporal Summaries of News Topics. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 10–18, New Orleans, LA, 2001.

[6] James Allan, Ron Papka, and Victor Lavrenko. On-line New Event Detection and Tracking. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 37–45, Melbourne, Australia, 1998.

[7] Richard Alterman. Summarization in the Small. In N. Sharkey, editor, *Advances in Cognitive Science*, Chichester, England, 1986. Ellis Horwood.

[8] Richard Alterman. Text Summarization. In S. C. Shapiro, editor, *Encyclopedia of Artificial Intelligence*, volume 2, pages 1579–1587. John Wiley & Sons, Inc., 1992.

[9] Richard Alterman and L. A. Bookman. Some Computational Experiments in Summarization. *Discourse Processes*, 13:143–174, 1990.

[10] Massih-Resa Amini and Patrick Gallinari. The Use of Unlabeled Data to Improve Supervised Learning for Text Summaries. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 105–112, Tampere, Finland, 2002.

[11] Einat Amitay and Cecile Paris. Automatically Summarising Web Sites - Is There a Way Around It? In *CIKM*, pages 173–179, 2000.

[12] Rie Ando, Branimir Boguraev, Roy Byrd, and Mary Neff. Multi-Document Summarization by Visualizing Topical Content. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[13] Roxana Angheluta, Rik De Busser, and Marie-Francine Moens. The Use of Topic Segmentation for Automatic Summarization. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[14] Roxana Angheluta, Marie-Francine Moens, and Rik De Busser. K.u. leuven summarization system. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[15] American National Standard for Writing Abstracts. Technical report, American National Standards Institute, Inc., New York, NY, 1979. ANSI Z39.14.1979.

[16] Chinatsu Aone, Mary Ellen Okurowski, James Gorlinsky, and Bjornar Larsen. A Scalable Summarization System Using Robust NLP. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, pages 66–73, 1997.

[17] Chinatsu Aone, Mary Ellen Okurowski, James Gorlinsky, and Bjornar Larsen. A Trainable Summarizer with Knowledge Acquired from Robust NLP Techniques. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 71–80. The MIT Press, 1999.

[18] Maria Aretoulaki. Towards a Hybrid Abstract Generation System. In *Proceedings of the International Conference on New Methods in Language Processing*, pages 220–227, Manchester, England, 1994.

[19] Maria Aretoulaki. *COSY-MATS: A Hybrid Connectionist-Symbolic Approach to the Pragmatic Analysis of Texts for their Automatic Summarization*. PhD thesis, Centre for Computational Linguistics, Dept. of Language Engineering, University of Manchester. Institute of Science and Technology (U.M.I.S.T.), Manchester, England, 1996.

[20] Amit Bagga and Ganesh Ramesh. A Text-based Method for Detection and Filtering of Commercial Segments in Broadcast News. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain, May–June 2002.

[21] Breck Baldwin and Thomas S. Morton. Dynamic Co-Reference Based Summarization. In *Proceedings of the 3rd Conference on Empirical Methods in Natural Language Processing (EMNLP-3)*, June 1998.

[22] Breck Baldwin and Aaron Ross. Baldwin Language Technology's DUC Summarization System. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[23] Regina Barzilay and Michael Elhadad. Using Lexical Chains for Text Summarization. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 111–121. The MIT Press, 1999.

[24] Regina Barzilay, Noémie Elhadad, and Kathleen R. McKeown. Sentence Ordering in Multidocument Summarization. In *Proceedings of the Human Language Technology Conference*, 2001.

[25] Regina Barzilay, Kathleen R. McKeown, and Michael Elhadad. Information Fusion in the Context of Multi-Document Summarization. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pages 550–557, College Park, Maryland, USA, June 16–20 1999.

[26] Regina Barzilay, Kathleen R. McKeown, and Michael Elhadad. Inferring Strategies for Sentence Ordering in Multidocument News Summarization. In *Journal of Artificial Intelligence Research*, pages 35–55, July 2002.

[27] P. B. Baxendale. Man-Made Index for Technical Literature - an Experiment. *IBM Journal of Research and Development*, 2(4):354–361, 1958.

[28] Mohamed Benbrahim and Khurshid Ahmad. Computer-Aided Lexical Cohesion Analysis and Text Abridgement. Technical report, University of Surrey, 1994.

[29] Mohamed Benbrahim and Khurshid Ahmad. Text Summarization: the Role of Lexical Cohesion Analysis. *The New Review of Document & Text Management*, pages 321–335, 1995.

[30] Adam L. Berger and Vibhu O. Mittal. OCELOT: A System for Summarizing Web Pages. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 144–151, 2000.

[31] Adam L. Berger and Vibhu O. Mittal. Query-Relevant Summarization Using FAQs. In *Proceedings of the 38th Meeting of the Association for Computational Linguistics*, pages 294–301, 2000.

[32] W. J. Black and F. C. Johnson. A Practical Evaluation of Two Rule-Based Automatic Abstracting Techniques. In *Expert Systems for Information Management 1*, pages 159–177. 1988.

[33] Branimir Boguraev, Rachel Bellamy, and C. Swart. Summarization Miniaturization: Delivery of News to Hand- Helds. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 99–110, 2001.

[34] Branimir Boguraev and Chris Kennedy. Salience-Based Content Characterization of Text Documents. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, pages 2–9, 1997.

4

[35] Branimir Boguraev, Chris Kennedy, Rachel Bellamy, Sascha Brawer, Y. Wong, and Jason Swartz. Dynamic Presentation of Document Content for Rapid On-Line Skimming. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, 1998.

[36] Harold Borko, editor. *Automated Language Processing*. Wiley, New York, 1968.

[37] Harold Borko and Charles Bernier. *Abstracting Concepts and Methods*. Academic Press, New York, 1975.

[38] Harold Borko and Seymour Chatman. Criteria for Acceptable Abstracts: A Survey of Abstractors' Instructions. *American Documentation*, 14(2):149–160, 1963.

[39] Endre Boros, Paul B. Kantor, and David J. Neu. A Clustering-based Approach to Creating Multi-Document Summaries. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[40] Ron Brandow, Karl Mitze, and Lisa F. Rau. Automatic Condensation of Electronic Publications by Sentence Selection. *Information Processing and Management*, 31(5):675–685, 1995.

[41] Erik Brill, Susan Dumais, and Michele Banko. An Analysis of the AskMSR Question-Answering System. In *Proceedings of the 39th Meeting of the Association for Computational Linguistics*, July 6–13 2002.

[42] Ann L. Brown and Jeanne D. Day. Macrorules for Summarizing Text: The Developments of Expertise. *JVLVB*, 22:1–14, 1983.

[43] Meru Brunn, Yllias Chali, and Barbara Dufour. U of L Summarizer at DUC2002. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[44] Meru Brunn, Yllias Chali, and Christopher J. Pinchak. Text Summarization Using Lexical Chains. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[45] Chris Buckley and Claire Cardie. Using EMPIRE and SMART for High-Precision IR and Summarization. In *Proceedings of the TIPSTER Text Phase III 12-Month Workshop*, San Diego, CA, October 1997.

[46] Orkut Buyukkokten, Hector Garcia-Molina, and Andreas Paepcke. Seeing the Whole in Parts: Text Summarization for Web Browsing on Handheld Devices. In *Proceedings of the Tenth International World-Wide Web Conference*, 2001.

[47] James P. Callan. Passage–Level Evidence in Document Retrieval. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 301–310, Amherst, MA, 1994.

[48] Jamie P. Callan, Yi Zhang, and Thomas Minka. Filtering: Novelty and Redundancy Detection in Adaptive Filtering. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, PA, 2001.

[49] Jaime Carbonell, Yiping Geng, and Jade Goldstein. Automated Query-Relevant Summarization and Diversity-Based Reranking. In *Proceedings of the IJCAI-97 Workshop on AI in Digital Libraries*, pages 12–19, 1997.

[50] Jaime G. Carbonell and Jade Goldstein. The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. In Alistair Moffat and Justin Zobel, editors, *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 335–336, Melbourne, Australia, 1998.

[51] Denis Carcagno and Lidija Iordanskaja. Content Determination and Text Structuring in Gossip. In *Extended Abstracts, Second European Natural Language Generation Workshop*, pages 15–22, Edinburgh, Scotland, April 6–8 1989.

[52] Jean Carletta. Assessing Agreement on Classification Tasks: The Kappa Statistic. *CL*, 22(2):249–254, 1996.

[53] Lynn Carlson, John M. Conroy, Daniel Marcu, Dianne P. O'Leary, Mary E. Okurowski, Anthony Taylor, and William Wong. An Empirical Study of the Relation between Abstracts, Extracts, and the Discourse Structure of Texts. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[54] Soumen Chakrabarti, Mukul Joshi, and Vivek Tawde. Enhanced Topic Distillation Using Text, Markup Tags, and Hyperlinks. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 208–216, New Orleans, LA, 2001.

[55] S. Chan, Tom Lai, W. Gao, and Benjamin T'sou. Mining Discourse Markers for Chinese Textual Summarization. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[56] Wesley T. Chuang and Jihoon Yang. Extracting Sentence Segments for Text Summarization: A Machine Learning Approach. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 152–160, 2000.

[57] Jonathan D. Cohen. Highlights: Language- and Domain-Independent Automatic Indexing Terms for Abstracting. *Journal of the American Society for Information Science*, 46(3):162–174, 1995.

[58] Ronald E. Cole, editor. *Survey of the State of the Art in Human Language Technology*, chapter 13, pages 475–518. Cambridge University Press, November 15 1995.

[59] James Conroy and Dianne O'Leary. Text Summarization via Hidden Markov Models and Pivoted QR Matrix Decomposition. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 406–407, New Orleans , LA, 2001.

[60] John M. Conroy, Judith D. Schlesinger, Dianne P. O'Leary, and Mary Ellen Okurowski. Using HMM and Logistic Regression to Generate Extract Summaries for DUC. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[61] Terry Copeck, Nathalie Japkowicz, and Stan Szpakowicz. Text Summarization as Controlled Search. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[62] Terry Copeck, Stan Szpakowicz, and Nathalie Japkowicz. Learning How Best to Summarize. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[63] Simon Corston-Oliver. Beyond String Matching and Cue Phrases: Improving Efficiency and Coverage in Discourse Analysis . In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on*

7

Intelligent Text Summarization, pages 34–43, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[64] Simon Corston-Oliver. Text Compaction for Display on Very Small Screens. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 89–98, 2001.

[65] Timothy C. Craven. Customized Extracts Based on Boolean Queries and Sentence Dependency Structures. *Intelligent Classification*, 16:11–14, 1989.

[66] Edward Cremmins. Valuable and Meaningful Text Summarization in Thoughts, Words, and Deeds. In Brigitte Endres-Niggemeyer, Jerry Hobbs, and Karen Sparck-Jones, editors, *Summarising Text for Intelligent Communication*. Dagstuhl, Germany, 1993.

[67] Edward T. Cremmins. *The Art of Abstracting*. Information Resources Press, Arlington, VA, 2nd edition, 1996.

[68] Maxime Crochemore and Wojciech Rytter. *Text Algorithms*. Oxford University Press, 1994.

[69] Graham Crookes. Towards a Validated Analysis of Scientific Text Structure. *Applied Linguistics*, 7(1):57–70, 1986.

[70] Naomi Daniel, Dragomir Radev, and Timothy Allison. Sub-Event-Based Multi-Document Summarization. In Dragomir Radev and Simone Teufel, editors, *HLT NAACL Workshop on Text Summarization*, pages 9–16, Edmonton, Alberta, Canada, May 2003. Association for Computational Linguistics.

[71] Gerald Francis DeJong. *Fast Skimming of News Stories: The FRUMP System*. PhD thesis, Yale University, New Haven, CT, 1978.

[72] Gerald Francis DeJong. Skimming Stories in Real Time: An Experiment in Integrated Understanding. Technical Report 158, New Haven ,CT, 1979.

[73] Gerald Francis DeJong. An Overview of the FRUMP System. In W. G. Lehnert and M. H. Ringle, editors, *Strategies for Natural Language Processing*, pages 149–176. Lawrence Erlbaum Associates, Publishers, 1982.

[74] Jean-Francois Delannoy, Ken Barker, Terry Kopeck, Martin Laplante, Stan Matwin, and Stan Szpakowicz. Flexible Summarization. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[75] J.Y. Delort, B. Bouchon-Meunier, and M. Rifqi. Enhanced Web Document Summarization Using Hyperlinks. In *Proceedings of the 14th ACM conference on Hypertext and Hypermedia*, pages 208–215. ACM Press, 2003.

[76] Robert L. Donaway, Kevin W. Drummey, and Laura A. Mather. A Comparison of Rankings Produced by Summarization Evaluation Measures. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, pages 69–78. Association for Computational Linguistics, April 30 2000.

[77] Bonnie Dorr, David Zajic, and Richard Schwartz. Hedge Trimmer: A Parse-and-Trim Approach to Headline Generation. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 1–8, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[78] Daniel M. Dunlavy, John M. Conroy, Judith D. Schlesinger, Sarah A. Goodman, Mary Ellen Okurowski, Dianne P. O'Leary, and Hans van Halteren. Performance of a Three-Stage System for Multi-Document Summarization. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[79] Lois L. Earl. Experiments in Automatic Extracting and Indexing. *Information Storage and Retrieval*, 6:313–334, 1970.

[80] H. P. Edmundson. Problems in Automatic Extracting. *Communications of the Association for Computing Machinery*, 7:259–263, 1964.

[81] H. P. Edmundson. New Methods in Automatic Extracting. *Journal of the Association for Computing Machinery*, 16(2):264–285, April 1969.

[82] Brigitte Endres-Niggemeyer. A Naturalistic Model of Abstracting. In *Preprints of Summarizing Text for Intelligent Communication. Dagstuhl Seminar Report 79*, pages 21–25, Schloss Dagstuhl, Germany, December 13–17 1993.

[83] Brigitte Endres-Niggemeyer. SimSum: Simulation of Summarizing. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, July 11 1997.

[84] Brigitte Endres-Niggemeyer and Neugebauer Elizabeth. Professional Summarizing: No Cognitive Simulation Without Observation. In *Proceedings of the International Conference in Cognitive Science*, San Sebastian, May 2–6 1995.

[85] Brigitte Endres-Niggemeyer, Jerry Hobbs, and Karen Sparck-Jones, editors. *Dagstuhl Seminar Report*. Schloss Dagstuhl, Wadern, Germany, 1993.

[86] Brigitte Endres-Niggemeyer, Jerry Hobbs, and Karen Sparck-Jones. Summarizing Text for Intelligent Communication. Schloss Dagstuhl, Wadern, Germany, 1993. Dagstuhl Seminar Report IBFI GmbH.

[87] Brigitte Endres-Niggemeyer, Elizabeth Maier, and Alexander Sigel. How to Implement a Naturalistic Model of Abstracting: Four Core Working Steps of an Expert Abstractor. *Information Processing & Management*, 31(5):631–674, 1995.

[88] Atefeh Farzindar and Guy Lapalme. Using Background Information for Multi-Document Summarization and Summaries in Response to a Question. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[89] Thérèse Firmin and Michael J. Chrzanowski. An Evaluation of Automatic Text Summarization Systems. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 325–336. MIT Press, 1999.

[90] N. M. Fontana. Summarising Strategies in L1 and L2. Ma dissertation, University College of North Wales, Bangor, 1989.

[91] Hannah Francis and Elizabeth Liddy. Structured Representation of Theoretical Abstracts: Implications for User Interface Design. In M. Dillon, editor, *Interfaces for Information Retrieval and Online Systems: The State of the Art*. Greenwood Press, 1991.

[92] Maria Fuentes, Marc Massot, Horacio Rodríguez, and Laura Alonso. Headline extraction combining statistic and symbolic techniques. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[93] Fumiyo Fukumoto and Yoshimi Suzuki. Extracting Key Paragraphs Based on Topic and Event Detection - Towards Multi-Document Summarization. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2002.

[94] Fumiyo Fukumoto, Yoshini Suzuki, and Jun'ichi Fukumoto. An Automatic Extraction of Key Paragraphs Based on Context Dependency. In *Proceedings of the 5th International on Applied Natural Language Processing*, Washington, 1997.

[95] Takahiro Fukusima and Manabu Okumura. Text Summarization Challenge: Text Summarization Evaluation in Japan. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 51–59, 2001.

[96] Danilo Fum, Giovanni Guida, and Carlo Tasso. Forward and Backward Reasoning in Automatic Abstracting. In *Proceedings of the 9th International Conference on Computational Linguistics*, pages 83–88, Prague, 1982.

[97] Danilo Fum, Giovanni Guida, and Carlo Tasso. Evaluating Importance: A Step Towards Text Summarization. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pages 840–844, Los Angeles, CA, August 18–23 1985.

[98] Robert Fung and Brendan Del Favero. Applying Bayesian Networks to Information Retrieval. *Communications of the ACM*, 38(3):42–48, March 1995.

[99] Robert Futrelle. Summarization of Documents that Include Graphics. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[100] Robert P. Futrelle. Summarization of Diagrams in Documents. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 403–421. MIT Press, Cambridge, MA, 2000.

[101] Robert Gaizauskas, Paul Clough, and S. L. Piao. Building and Annotating a Corpus for the Study of Journalistic Text Reuse. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain, May–June 2002.

[102] Ruth Garner. Efficient Text Summarization: Costs and Benefits. *Journal of Education Research*, 75:275–279, 1982.

[103] Philip Gladwin, Stephen Pulman, and Karen Sparck-Jones. Shallow Processing and Automatic Summarizing: A First Study. Technical Report Technical Report No. 223, University of Cambridge Computer Laboratory, May 1991.

[104] Jade Goldstein, Mark Kantrowitz, Vibhu O. Mittal, and Jamie G. Carbonell. Summarizing Text Documents: Sentence Selection and Evaluation Metrics. In *Research and Development in Information Retrieval*, pages 121–128, Berkeley, California, 1999.

[105] Jade Goldstein and Chin-Yew Lin, editors. *Proceedings of the Workshop on Automatic Summarization at the 2nd Conference of the North American Chapter of the Association for Computational Linguistics.* Pittsburgh, PA, 2001.

[106] Jade Goldstein, Vibhu O. Mittal, Jamie Carbonell, and Mark Kantrowitz. Multi-Document Summarization by Sentence Extraction. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[107] Yihong Gong and Xin Liu. Generic Text Summarization Using Relevance Measure and Latent Semantic Analysis. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[108] Stephen J. Green. Building Hypertext Links in Newspaper Articles Using Semantic Similarity. Technical report, Department of Computer Science,University of Toronto, 1997.

[109] Gregory Grefenstette, editor. *Cross-Language Information Retrieval.* Kluwer Academic Publishers, USA, 1998.

[110] Gregory Grefenstette. Producing Intelligent Telegraphic Text Reduction to Provide an Audio Scanning Service for the Blind. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 111–117, Stanford, CA, March 1998.

[111] Gregory Grefenstette. *The Problem of Cross-Language Information Retrieval*, pages 1–9. Kluwer Academic Publishers, 1998.

[112] Amardeep Grewal, Timothy Allison, Stanko Dimitrov, and Dragomir Radev. Multi-document Summarization Using Off the Shelf Compression Software. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 17–24, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[113] Joseph E. Grimes. The Thread of Discourse. *Jangua Linguarum, Series Minor*, (207), 1975.

[114] Barbara J. Grosz and Candace L. Sidner. Attention, Intention, and the Structure of Discourse. *Computational Linguistics*, 12(3), 1986.

[115] Claire Grover, Ben Hachey, and Chris Korycinski. Summarising Legal Texts: Sentential Tense and Argumentative Roles. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 33–40, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[116] Udo Hahn. Topic Parsing: Accounting for Text Macro Structures in Full-Text Analysis. *Information Processing & Management*, 26(1):135–170, 1990.

[117] Udo Hahn and Donna Harman, editors. *Proceedings of the 2nd Document Understanding Conference*. Philadelphia, PA, July 2002.

[118] Udo Hahn and Donna Harman, editors. *Proceedings of the Workshop on Text Summarization at the 40th Meeting of the Association for Computational Linguistics*. Philadelphia, PA, July 11–12 2002.

[119] Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors. *Proceedings of the Workshop on Automatic Summarization the 6th Applied Natural Language Processing Conference and at the 1st Meeting of the North American Chapter of the Association for Computational Linguistics*. Seattle, WA, April 29– May 4 2000.

[120] Udo Hahn and Ulrich Reimer. Knowledge-Based Text Summarization: Salience and Generalization Operators for Knowledge Base Abstraction. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 215–232. MIT Press, July 1999.

[121] Udo Hahn and Michael Strube. Centered Segmentation: Scaling Up the Centering Model to Global Discourse Structure. In *Proceedings of the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, 1997.

[122] Thérèse F. Hand. A Proposal for Task-Based Evaluation of Text Summarization Systems. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, pages 31–38, Madrid, Spain, July 1997.

[123] Sanda Harabagiu. From Lexical Cohesion to Textual Coherence: A Data Driven Perspective. *Journal of Pattern Recognition and Artificial Intelligence*, 13(2)(4):247–265, 1999.

[124] Sanda Harabagiu and Finley Lacatusu. Generating single and multi document summaries with GISTEXTER. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[125] Hilda Hardy, Nobuyuki Shimizu, Tomek Strzalkowski, Liu Ting, Xinyang Zhang, and Bowden G. Wise. Cross-Document Summarization by Concept Classification. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 121–128, Tampere, Finland, 2002.

[126] Chou V. Hare and Kathleen M. Borchardt. Direct Instruction of Summarization Skills. *Reading Research Quarterly*, 20:62–78, 1984.

[127] Donna Harman and Daniel Marcu, editors. *Proceedings of the 1st Document Understanding Conference*. New Orleans, LA, September 2001.

[128] Koiti Hasida, Syun Ishizaki, and Hitoshi Isahara. A Connectionist Approach to the Generation of Abstracts. In G. Kempen, editor, *Natural Language Generation: New Results in Artificial Intelligence, Psychology and Linguistics*, Dordrecht,the Netherlands, 1987. Nijhoff,Martinus NATO Advanced Science Institutes Series.

[129] Marti A. Hearst. Subtopic Structuring for Full-Length Document Access. In *Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Pittsburgh, PA, 1993.

[130] Marti A. Hearst. Multi-Paragraph Segmentation of Expository Text. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Las Cruces, NM, 1994.

[131] Tsutomu Hirao, Yutaka Sasaki, and Hideki Isozaki. An Extrinsic Evaluation for Question-Biased Text Summarization on QA Tasks. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 61–68, 2001.

[132] Tsutomu Hirao, Yutaka Sasaki, Hideki Isozaki, and Eisaku Maeda. NTT's Text Summarization System for DUC 2002 . In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document*

Understanding Conference at the 40th Meeting of the Association for Computational Linguistics, 2002.

[133] Jerry R. Hobbs. On the Coherence and Structure of Discourse. In *CSLI-85-37 Center for the Study of Language and Information*, 1985.

[134] Eduard Hovy. Parsimonious and Profligate Approaches to the Question of Discourse Structure Relations. In *Proceedings of the 5th International Workshop on Natural Language Generation*, pages 128–136, Dawson, PA, 1990.

[135] Eduard Hovy. Automated Discourse Generation Using Discourse Structure Relations. *Artificial Intelligence*, 63:341–385, 1993.

[136] Eduard Hovy. In Defense of Syntax: Informational, Intentional, and Rhetorical Structures in Discourse. pages 35–39, June 1993.

[137] Eduard Hovy and Chin Yew Lin. Automated Text Summarization in SUMMARIST. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 81–94. The MIT Press, 1999.

[138] Eduard Hovy and Chin-Yew Lin. Manual and Automatic Evaluation of Summaries. In Udo Hahn and Donna Harman, editors, *Proceedings of the Workshop on Text Summarization at the 40th Meeting of the Association for Computational Linguistics*, July 11–12 2002.

[139] Eduard Hovy and Dragomir R. Radev, editors. *Intelligent Text Summarization. Papers from the 1998 AAAI Spring Symposium*. The AAAI Press, Stanford, California, USA, March 23–25 1998.

[140] Xiaorong Huang. Planning Reference Choices for Argumentative Texts. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, pages 190–197, Madrid, Spain, 1997.

[141] John Hughes and Kathleen McCoy. Observations and Directions in Text Structure. pages 40–43, 1993.

[142] John Hutchins. Summarization: Some Problems and Methods. In K.P. Jones, editor, *Meaning: The Frontier of Informatics*, volume 9, pages 151–173. Aslib, 1987.

[143] Documentation—Abstracts for Publication and Documentation. ISO 214-1976. Technical report, International Organisation for Standardisation, 1976.

[144] Paul S. Jacobs and Lisa F. Rau. SCISOR: Extracting Information from On-line News. *Communications of the ACM*, 33(11):88–97, 1990.

[145] Hongyan Jing. Sentence Reduction for Automatic Text Summarization. In *Proceedings of the 6th Applied Natural Language Processing Conference*, pages 310–315, Seattle,WA, April 29–May 4 2000.

[146] Hongyan Jing. Using Hidden Markov Modelling to Decompose Human-Written Summaries. *Computational Linguistics*, 28(4), 2002.

[147] Hongyan Jing, Daniel Lopresti, and Chilin Shih. Summarization of Noisy Documents: A Pilot Study. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 25–32, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[148] Hongyan Jing and Kathleen R. McKeown. The Decomposition of Human-Written Summary Sentences. In M. Hearst, Gey. F., and R. Tong, editors, *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 129–136, University of California, Beekely, August 1999.

[149] Hongyan Jing and Kathleen R. McKeown. Cut and Paste-Based Text Summarization. In *Proceedings of the 6th Applied Natural Language Processing Conference and the 1st Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 178–185, Seattle, WA, April 2000.

[150] Hongyan Jing, Kathleen R. McKeown, Regina Barzilay, and Michael Elhadad. Summarization Evaluation Methods: Experiments and Analysis. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 60–68, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[151] Frances C. Johnson, Chris D. Paice, William J., and A. P. Neal. The Application of Linguistic Processing to Automatic Abstract Generation. *Journal of Document and Text Management*, 1(3):215–241, 1993.

[152] Paul A. Jones and Chris D. Paice. A 'Select and Generate' Approach to Automatic Abstracting. In A. M. McEnry and Chris D. Paice, editors, *Proceedings of the 14th British Computer Society Information Retrieval Colloquium*, pages 151–154. Springer Verlag, 1992.

[153] M. P. Jordan. The Linguistic Genre of Abstracts. In A. Della Volpe, editor, *The Seventeenth LACUS Forum. Linguistics Association of Canada and the United States*, pages 507–527, 1991.

[154] Murat Karamuftuoglu. An Approach to Summarization Based on Lexical Bonds. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[155] Walter Kintsch and Teun A. van Dijk. Comment on se rappelle et on résume des histoires. *Langages*, 40:98–116, December 1975.

[156] Walter Kintsch and Teun A. van Dijk. Toward a Model of Text Comprehension and Production. *Psychological Review*, 85(5):363–394, 1978.

[157] Kevin Knight and Daniel Marcu. Statistics-Based Summarization — Step One: Sentence Compression. In *Proceedings of the 17th National Conference of the American Association for Artificial Intelligence*, pages 703–710, 2000.

[158] Alastair Knott. Using Linguistic Phenomena to Motivate a Set of Coherence Relations. *Discourse Processes*, 18(1):35–62, 1994.

[159] Alastair Knott. *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. PhD thesis, Department of Artificial Intelligence, University of Edinburgh, 1996.

[160] Alastair Knott and Robert Dale. Choosing a Set of Coherence Relations for Text Generation: a Data-Driven Approach. 1996.

[161] Aleksander Kolcz, Vidya Prabakarmurthi, and Jugal Kalita. Summarization as Feature Selection for Text Categorization. In *Proceedings of the 10th International Conference on Information and Knowledge Management*, pages 365–370, Atlanda, GA, 2001.

[162] Wessel Kraaij, Martin Spitters, and Anette Hulth. Headline Extraction Based on a Combination of Uni- and Multi-Document Summarization Techniques. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[163] Wessel Kraaij, Martin Spitters, and Martine van der Heiden. Combining a Mixture Language Model and Naive Bayes for Multi-Document Summarisation. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[164] Klaus Krippendorff. *Content Analysis: An Introduction to its Methodology*. Sage Publications, Beverly Hills, CA, 1980.

[165] Julian Kupiec, Jan O. Pedersen, and Francine Chen. A Trainable Document Summarizer. In *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 68–73, 1995.

[166] Ka Lok Kwok, N. Grunfeld, N. Dinstl, and M. Chan. TREC-9 Cross Language, Web and Question-Answering Track Experiments using PIRCS. In *The 9th Text REtrieval Conference*, 2000.

[167] Finley Lacatusu, Paul Parker, and Sanda Harabagiu. Lite-GISTexter: Generating Short Summaries with Minimal Resources. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[168] Horald Ladas. Summarising Research: A Case Study. Review of an Issue on Empirical Studies in Discourse Interpretation and Generation, 1997.

[169] Adenike Lam-Adesina and Gareth Jones. Applying Summarization Techniques for Term Selection in Relevance Feedback. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[170] Frederick Wilfrid Lancaster. *Indexing and Abstracting in Theory and Practice*. Library Association, London, UK, 1998.

[171] Robin J. Landis and G. G. Koch. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33:159–174, 1977.

[172] Mirella Lapata. Probabilistic Text Structuring: Experiments with Sentence Ordering. In *Proceedings of the 41th Meeting of the Association for Computational Linguistics*, Sapporo, Japan, 2003.

[173] Dawn Lawrie, W. Bruce Croft, and Arnold Rosenberg. Finding Topic Words for Hierarchical Summarization. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 349–357, New Orleans, LA, 2001.

[174] Dominique Le Roux, Jean-Luc Minel, and Jawad Berri. SERAPHIN project. In *First European Conference of Cognitive Science in Industry*, Luxembourg, September 28-30 1994.

[175] Aberrafih Lehmam. *Le resume des textes techniques et scientifiques, aspects linguistiques et computationnels*. PhD thesis, Universite de Nancy 2, 1995.

[176] Wendy G. Lehnert. Plot Units and Narrative Summarization. *Cognitive Science*, 5(4):293–331, 1981.

[177] Wendy G. Lehnert and Beth Sundheim. A Performance Evaluation of Text Analysis Technology. *AI magazine*, 12(3):81–94, 1991.

[178] Hang Li and Kenji Yamanishi. Document Classification Using a Finite Mixture Model. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, 1997.

[179] Elizabeth D. Liddy. Anaphora in Natural Language Processing and Information Retrieval. *Information Processing and Management*, 26(1):39–52, 1990.

[180] Elizabeth D. Liddy. Discourse-level Structure of Empirical Abstracts: An Exploratory Study. *Information Processing and Management*, 27(1):550–81, 1991.

[181] Elizabeth D. Liddy, Susan Bonzi, Jeffrey Katzer, and E. Oddy. A Study of Discourse Anaphora in Scientific Abstracts. *Journal of the American Society for Information Science*, 38(4):255–261, 1987.

[182] Chin-Yew Lin. Assembly of Topic Extraction Modules in SUMMARIST. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 34–43, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[183] Chin-Yew Lin. Training a Selection Function for Extraction. In *Proceedings of the 18th Annual International ACM Conference on Information and Knowledge Management (CIKM)*, pages 55–62, Kansas City, KS, November 2–6 1999.

[184] Chin-Yew Lin. Summary Evaluation Environment, 2001. http://www.isi.edu/~cyl/SEE.

[185] Chin-Yew Lin and Eduard Hovy. Identifying Topics by Position. In *Proceedings of the 5th Conference on Applied Natural Language Processing*, pages 283–290. Association for Computational Linguistics, March 31 - April 3 1997.

[186] Chin-Yew Lin and Eduard Hovy. The Automated Acquisition of Topic Signatures for Text Summarization. In *Proceedings of the 18th COLING Conference*, Saarbrücken, Germany, 2000.

[187] Chin-Yew Lin and Eduard Hovy. From Single to Multi-document Summarization: A Prototype System and its Evaluation. In *Proceedings of the 2nd Document Understanding Conference at the 40th Meeting of the*

Association for Computational Linguistics, pages 457–464, Philadelphia, PA, July 2002.

[188] Chin-Yew Lin and Eduard Hovy. Manual and Automatic Evaluation of Summaries. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[189] Chin-Yew Lin and Eduard Hovy. NeATS in DUC 2002. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[190] Chin-Yew Lin and Eduard Hovy. The Potential and Limitations of Automatic Sentence Extraction for Summarization. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[191] Robert Longacre. The Paragraph as a Grammatical Unit. In T. Givon, editor, *Syntax and Semantics 12*. Academic Press, 1979.

[192] Natalia Loukachevitch. Text Summarization Based on Thematic Representation of Texts. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 34–43, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[193] H. P. Luhn. The Automatic Creation of Literature Abstracts. *IBM Journal of Research Development*, 2(2):159–165, 1958.

[194] Kavi Mahesh. Hypertext Summary Extraction for Fast Document Browsing. In *Natural Language Processing for the World Wide Web. Papers from the 1997 AAAI Spring Symposium*, pages 95–104, Stanford, CA, 1999.

[195] Robert E. Maizell, Julian F. Smith, and T.E.R. Singer. *Abstracting Scientific and Technical Literature*. Wiley-Interscience, A Division of John Wiley & Son, Inc., 1971.

[196] Inderjeeet Mani, David House, Mark Maybury, and Morgan Green. Towards Content-Based Browsing of Broadcast News Video. In Mark T. Maybury, editor, *Multimedia Information Retrieval*. AAAI/MIT Press, 1997.

[197] Inderjeet Mani. *Automatic Summarization*. John Benjamins Publishing Company, Amsterdam/Philadephia, 2001.

[198] Inderjeet Mani. Recent developments in text summarization. In *Proceedings of the 10th International Conference on Information and Knowledge Management*, pages 529–531, Atlanta, Georgia, USA, 2001.

[199] Inderjeet Mani and Eric Bloedorn. Multi-Document Summarization by Graph Search and Matching. In *Proceedings of the 14th National Conference on Artificial Intelligence*, pages 622–628, Providence, Rhode Island, 1997.

[200] Inderjeet Mani and Eric Bloedorn. Summarizing Similarities and Differences Among Related Documents. volume 1, 2000.

[201] Inderjeet Mani, Kristian Concepción, and Linda van Guilder. Using Summarization for Automatic Briefing Generation. In *Proceedings of the 6th Applied Natural Language Processing Conference and the 1st Meeting of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[202] Inderjeet Mani, Thérèse Firmin, David House, Gary Klein, Beth Sundheim, and Lynette Hirschman. The TIPSTER SUMMAC Text Summarization Evaluation. In *Natural Language Engineering (to appear)*, 2001.

[203] Inderjeet Mani, Barbara Gates, and Eric Bloedorn. Using Cohesion and Coherence Models for Text Summarization. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 69–76, Stanford, CA, March 23–25 1998. AAAI Press.

[204] Inderjeet Mani, Barbara Gates, and Eric Bloedorn. Improving Summaries by Revising Them. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pages 558–565, College Park, Maryland, USA, June 1999.

[205] Inderjeet Mani, David House, G. Klein, Lynette Hirshman, Leo Orbst, Thérèse Firmin, Michael Chrzanowski, and Beth Sundheim. The TIPSTER SUMMAC Text Summarization Evaluation. Technical Report MTR 98W0000138, The Mitre Corporation, McLean, Virginia, 1998.

[206] Inderjeet Mani and Mark T. Maybury, editors. *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*. Madrid, Spain, July 1997.

[207] Inderjeet Mani and Mark T. Maybury, editors. *Advances in Automatic Text Summarization*. MIT Press, Cambridge, MA, 1999.

[208] Inderjeet Mani, Barry Schiffman, and Jianping Zhang. Inferring Temporal Ordering of Events in News.

[209] William Mann and Sandra Thompson. Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text*, 8(3):243–281, 1988.

[210] Daniel Marcu. Discourse Trees Are Good Indicators of Importance in Text. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 123–136, Cambridge, MA, 1995. MIT Press.

[211] Daniel Marcu. Building Up Rhetorical Structure Trees. In *Proceedings of the 13th National Conference on Artificial Intelligence*, pages 1069–1074, Portland, Oregon, 1996.

[212] Daniel Marcu. From Discourse Structures to Text Summaries. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, pages 82–88, Madrid, Spain, July 11 1997.

[213] Daniel Marcu. *The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts*. PhD thesis, University of Toronto, 1997.

[214] Daniel Marcu. To Build Text Summaries of High Quality, Nuclearity is Not Sufficient. In *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 1–8, Stanford, California, USA, March 23–25 1998.

[215] Daniel Marcu. The Automatic Construction of Large-Scale Corpora for Summarization Research. In M. Hearst, Gey. F., and R. Tong, editors, *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 137–144, University of California, Berkely, August 1999.

[216] Daniel Marcu. *The Theory and Practice of Discourse Parsing and Summarization*. MIT Press, Cambridge/London, 2000.

[217] Daniel Marcu. Discourse-based Summarization in DUC-2001. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[218] Daniel Marcu, Hal Daumé, Abdessamad Echihabi, Dragos Stefan Munteanu, and Radu Soricut. GLEANS: A Generator of Logical Extracts and Abstracts for Nice Summaries. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[219] Daniel Marcu and Laurie Gerber. An Inquiry into the Nature of Multi-document Abstracts, Extracts, and Their Evaluation. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 1–8, Pittsburgh, PA, June 2001.

[220] Mark T. Maybury. Generating Summaries from Event Data. *Information Processing and Management*, 31(5):735–751, 1995.

[221] Mark T. Maybury and Andrew E. Merlino. An Empirical Study of the Optimal Presentation of Multimedia Summaries of Broadcast News. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 392–401. MIT Press, 1999.

[222] Diana Maynard, Kalina Bontcheva, Horacio Saggion, Hamish Cunningham, and Oana Hamza. Using a Text Engineering Framework to Build an Extendable and Portable IE-based Summarisation System. In *Proceedings of the 39th Meeting of the Association for Computational Linguistics*, July 6–13 2002.

[223] Daniel McDonald and Hsinchun Chen. Using Sentence Selection Heuristics to Rank Text Segments in TXTRACTOR. In *Proceedings of the 2nd ACM/IEEE Joint Conference on Digital Libraries*, pages 25–38, Portland, Oregon, 2002.

[224] Clinton J. McGirr. Guidelines for Abstracting. *Technical Communication*, 25(2):2–5, 1973.

[225] Kathleen McKeown, Regina Barzilay, Sasha Blair-Goldensohn, David Evans, Vasileios Hatzivassiloglou, Judith Klavans, Ani Nenkova, Barry Schiffman, and Sergey Sigelman. The Columbia Multi-Document Summarizer. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[226] Kathleen R. McKeown. Generating the Complex Sentences of Summaries Using Syntactic and Lexical Constraints: Two Applications. In Brigitte Endres-Niggemeyer, Jerry Hobbs, and Karen Sparck-Jones, editors, *Preprints of Summarizing Text for Intelligent Communication*, number 79. Schloss Dagstuhl, Germany, December 13–17 1993.

[227] Kathleen R. McKeown, Regina Barzilay, David Evans, Vasileios Hatzivassiloglou, Simone Teufel, Yen M. Kan, and Barry Schiffman. Columbia

Multi-Document Summarization: Approach and Evaluation. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[228] Kathleen R. McKeown, Shih-Fu Chang, James Cimino, Steven Feiner, Carol Friedman, Luis Gravano, and Vasileios Hatzivassiloglou. PERSIVAL: A System for Personalized Search and Summarization Over Multimedia Healthcare Information. In *Proceedings of the 1st ACM IEEE-CS Joint Conference on Digital Libraries*, pages 331–340, Roanoke, VA, January 2001.

[229] Kathleen R. McKeown, Vasileios Hatzivassiloglou, Judith L. Klavans, Holcombe Melissa L., Regina Barzilay, and Min-Yen Kan. SIMFinder: A Flexible Clustering Tool for Summarization. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages pages 41–49, 2001.

[230] Kathleen R. McKeown, Desmond Jordan, and Vasileios Hatzivassiloglou. Generating Patient-Specific Summaries of On-Line Literature. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 34–43, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[231] Kathleen R. McKeown, M-Y Kan, and Judith Klavans. Domain-Specific Informative and Indicative Summarization for Information Retrieval. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[232] Kathleen R. McKeown, Judith Klavans, Vasileios Hatzivassiloglou, Regina Barzilay, and Eleazar Eskin. Towards Multidocument Summarization by Reformulation: Progress and Prospects. In *Proceedings of the 16th National Conference on Artificial Intelligence*, pages 453–460, July 18–22 1999.

[233] Kathleen R. McKeown and Dragomir R. Radev. Generating Summaries of Multiple News Articles. In *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 74–82, Seattle, Washington, July 1995.

[234] Kathleen R. McKeown, Jacques Robin, and Karen Kukich. Generating Concise Natural Language Summaries. *Information Processing & Management*, 31(5):702–733, 1995.

[235] Michael A. K. Halliday and Ruqaiya Hasan . *Cohesion in English*. Longmans, London, 1996.

[236] Herbert B. Michaelson. *How to Write and Publish Engineering Papers and Reports*. Oryx Press, Phoenix, AZ, 1980.

[237] Seiji Miike, Etsuo Itoh, Kenji Ono, and Kazuo Sumita. A Full-Text Retrieval System With A Dynamic Abstract Generation Function. In W. Bruce Croft and C. J. van Rijsbergen, editors, *Proceedings of the 17th International Conference on Research and Development in Information Retrieval*, pages 152–161, Dublin, Ireland, July 3–6 1994.

[238] Jean Luc Minel, Sylvaine Nugier, and Gerald Piat. How to appreciate the quality of automatic text summarization? examples of fan and mluce protocols and their results on seraphin. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, pages 25–30, Madrid, Spain, 1997.

[239] Ruslan Mitkov, Dominique Le Roux, and Jean Pierre Desclés. Knowledge-Based Automatic Abstracting: Experiments in the Sublanguage of Elementary Geometry. In C. Martin-Vide, editor, *Current Issues in Mathematical Linguistics*. North-Holland, The Netherlands, 1994.

[240] Mandar Mitra, Amit Singhal, and Chris Buckley. Automatic Text Summarization by Paragraph Extraction. In *Proceedings of the Workshop on Intelligent Scalable Text Summarization*, pages 39–46, Madrid, Spain, July 1997. Association for Computational Linguistics.

[241] A. Morris, G. Kasper, and D. Adams. The Effects and Limitations of Automated Text Condensing on Reading Comprehension Performance. *Information Systems Research*, 3(1):17–35, 1992.

[242] James Morris and Graeme Hirst. Lexical Cohesion Computed by Thesaural Relations as an Indicator of the Structure of Text. *Computational Linguistics*, 17(1):21–43, 1991.

[243] Sumiko Mushakoji. Constructing "Identity" and "Differences" in Original Scientific Texts and Their Summaries: Its Problems and Solutions. In Brigitte Endres-Niggemeyer, Jerry J. Hobbs, and Karen Sparck-Jones, editors, *Workshop on Summarising Text for Intelligent Communication*. Dagstuhl, Germany, 1993.

[244] Sumiko Mushakoji and Atsutake Nozoe. Toward Qualified Medical Abstracts: Rethinking the Process of Producing Author Abstracts. In K.C. Lun et al., editor, *Elsevier*. Medinfo, 1992.

[245] Yoshio Nakao. An Algorithm for One-Page Summarization of a Long Text Based on Thematic Hierarchy Detection. In *Proceedings of the 38th*

Meeting of the Association for Computational Linguistics, pages 302–309, 2000.

[246] Yoshio Nakao. How small a distinction among summaries can an ir-based evaluation method identify? In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 69–78, 2001.

[247] Hidetsugu Nanba and Manabu Okumura. Producing More Readable Extracts by Revising Them. In *Proceedings of the 18th International Conference on Computational Linguistics (COLING-2000)*, pages 1071–1075, 2000.

[248] Masumi Narita, Kazuya Kurokawa, and Takehito Utsuro. A Web-based English Abstract Writing Tool Using A Tagged E-J Parallel Corpus. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain, May–June 2002.

[249] Ani Nenkova, Barry Schiffman, Andrew Schlaiker, Sasha Blair-Goldensohn, Regina Barzilay, Sergey Sigelman, Vasileios Hatzivassiloglou, and Kathleen McKeown. Columbia at the DUC 2003. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[250] Tadashi Nomoto. ModDBS-X: A Diversity-based Summarizer for DUC2001. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[251] Tadashi Nomoto and Yuji Matsumoto. A New Approach to Unsupervised Text Summarization. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, New Orleans, LA, 2001.

[252] Tadashi Nomoto and Yuji Matsumoto. Modeling (In)variability of Human Judgements for Text Summarization. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tampere, Finland, 2002.

[253] Tadashi Nomoto and Yuji Matsumoto. The Diversity-based Approach to Open-domain Text Summarization. *Information Processing and Management*, 39(3):363–389, 2003.

[254] Tadashi Nomoto and Yoshihiko Nitta. A Grammatico–Statistical Approach to Discourse Partitioning. In *Proceedings of the 32th Meeting of the Association for Computational Linguistics*, 1994.

26

[255] Ryo Ochitani, Yoshio Nakao, and Fumihito Nishino. Goal Directed Approach for Text Summarization. In *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, July 11 1997.

[256] Mamiko Oka and Yoshihiro Ueda. Evaluation of Phrase-Representation Summarization Based on Information Retrieval Task. In *Proceedings of the 6th Applied Natural Language Processing Conference and the 1st Meeting of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[257] Manabu Okumura, Takahiro Fukusima, and Hidetsugu Nanba. Text summarization challenge 2 - text summarization evaluation at ntcir workshop 3. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 49–56, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[258] Manabu Okumura, Hajime Mochizuki, and Hidetsugu Nanba. Query-Biased Summarization Based on Lexical Chaining. In *Proceedings of the Pacific Association for Computational Linguistics*, pages 324–334, 1999.

[259] Mary Ellen Okurowski, Harold Wilson, Joacquin Urbina, Tony Taylor, Ruth Colvin Clark, and Frank Krapcho. A Text Summarizer in Use: Lessons Learned from Real World Deployment and Evaluation. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[260] Kenji Ono, Kazuo Sumita, and Seiji Miike. Abstract Generation Based on Rhetorical Structure Extraction. In *Proceedings of the International Conference on Computational Linguistics*, pages 344–348, Kyoto, Japan, 1994.

[261] Constantin Orasan. Building Annotated Resources for Automatic Text Summarisation. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain, May–June 2002.

[262] Constantin Orasan, Ruslan Mitkov, and Laura Hasler. Cast: a computer-aided summarization tool. In *Proceedings of the 11th Meeting of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary, April 12–17 2003.

[263] Miles Osborne. Using Maximum Entropy for Sentence Extraction. In Udo Hahn and Donna Harman, editors, *Proceedings of the Workshop on Text Summarization at the 40th Meeting of the Association for Computational Linguistics*, July 12–13 2002.

[264] V. A Oswald. Automatic Indexing and Abstracting of the Contents of Documents. *Planning Research Corporation*, 31, 1959.

[265] Jahna Otterbacher, Dragomir R. Radev, and Airong Luo. Revisions that Improve Cohesion in Multi-Document Summaries: a Preliminary Study. In Udo Hahn and Donna Harman, editors, *Proceedings of the Workshop on Text Summarization at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 11–12 2002.

[266] Chris Paice and P. A. Jones. The Identification of Important Concepts in Highly Structured Technical Papers. In R. Korfhage, E. Rasmussen, and P. Willett, editors, *Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 69–78, 1993.

[267] Chris D. Paice. The Automatic Generation of Literary Abstracts: An Approach Based on Identification of Self-Indicating Phrases. In O. R. Norman, S. E. Robertson, C. J. van Rijsbergen, and P. W. Williams, editors, *Information Retrieval Research*, London: Butterworth, 1981.

[268] Chris D. Paice. Automatic Generation and Evaluation of Back-of-Book Indexes. In *Prospects for Intelligent Retrieval*, 1989.

[269] Chris D. Paice. Constructing Literature Abstracts by Computer: Techniques and Prospects. *Information Processing and Management*, 26(1):171–186, 1990.

[270] Chris D. Paice. The Automatic Generation and Evaluation of Back–of–Books Indexes. In *Proceedings of the IO conference "Prospects for Intelligent Retrieval"*, 1990.

[271] Chris D. Paice. The Rhetorical Structure of Expository Text. In *Proceedings of Informatics 11 Conference*, 1991.

[272] Chris D. Paice and Michael P. Oakes. A Concept-Based Method for Automatic Abstracting. Technical Report Research Report 27, Library and Information Commission, 1999.

[273] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. BLEU: A Method for Automatic Evaluation of Machine Translation. Research Report RC22176, IBM, 2001.

[274] Justin Picard. Modeling and Combining Evidence Provided by Document Relationships Using Probabilistic Argumentation Systems. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 182–189, Melbourne, Australia, 1998.

[275] Livia Polanyi. Linguistic Dimensions of Text Summarization. In *Workshop on Summarising Text for Intelligent Communication, volume 9350*. Dagstuhl, Germany, 1993.

[276] J. Pollock and Antonio Zamora. Automatic Abstracting Research at Chemical Abstracts Service. *Journal of Chemical Information and Computer Sciences*, 15(4), 1975.

[277] Keith Preston and Sandra Williams. Managing the Information Overload. *Physics in Business*, June 1994.

[278] Dragomir Radev, Simone Teufel, Horacio Saggion, Wai Lam, John Blitzer, Arda Celebi, Hong Qi, Daniu Liu, Elliott Drabek. Evaluation Challenges in Large-Scale Multi-Document Summarization. In *ACL03*, Sapporo, Japan, July 7-12 2003. Association for Computational Linguistics.

[279] Dragomir Radev, Jahna Otterbacher, Hong Qi, and Daniel Tam. MEAD ReDUCs: Michigan at DUC 2003. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[280] Dragomir R. Radev. *Language Reuse and Regeneration: Generating Natural Language Summaries from Multiple On-Line Sources*. PhD thesis, Department of Computer Science, Columbia University, New York, April 1999.

[281] Dragomir R. Radev. A Common Theory of Information Fusion from Multiple Text Sources, Step One: Cross-document Structure. In *Proceedings of the 1st Workshop on Discourse and Dialogue of the Association for Computational Linguistics*, Hong Kong, October 2000.

[282] Dragomir R. Radev, Sasha Blair-Goldensohn, Zhu Zhang, and Revathi Sundara Raghavan. Interactive, Domain-Independent Identification and Summarization of Topically Related News Articles. In *5th European Conference on Research and Advanced Technology for Digital Libraries*, Darmstadt, Germany, 2001.

[283] Dragomir R. Radev, Sasha Blair-Goldensohn, Zhu Zhang, and Revathi Sundara Raghavan. NewsInEssence: A System for Domain-Independent, Real-Time News Clustering and Multi-Document Summarization. In *Proceedings of the Human Language Technology Conference*, San Diego, CA, 2001.

[284] Dragomir R. Radev and Weiguo Fan. Automatic Summarization of Search Engine Hit Lists. In *Proceedings of the Workshop on Recent Advances in NLP and IR at the 38th Meeting of the Association for Computational Linguistics*, Hong Kong, October 2000.

[285] Dragomir R. Radev, Weiguo Fan, and Zhu Zhang. WebInEssence: A Personalized Web-Based Multi-Document Summarization and Recommendation System. In *Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, Pittsburgh, PA, 2001.

[286] Dragomir R. Radev, Hongyan Jing, and Malgorzata Budzikowska. Centroid-Based Summarization of Multiple Documents: Sentence Extraction, Utility-Based Evaluation, and User Studies. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, April 2000.

[287] Dragomir R. Radev and Kathleen R. McKeown. Building a Generation Knowledge Source Using Internet-Accessible Newswire. In *Proceedings of the 5th Conference on Applied Natural Language Processing*, pages 221–228, Washington, DC, April 1997.

[288] Dragomir R. Radev and Kathleen R. McKeown. Generating Natural Language Summaries from Multiple On-Line Sources. *Computational Linguistics*, 4:469–500, September 1998.

[289] Dragomir R. Radev, Hong Qi, Jahna Otterbacher, and Adam Winkel. The University of Michigan at TREC2002: Question Answering and Novelty tracks. In *The 11th Text REtrieval Conference*, Gaithersburg, MD, November 2002.

[290] Dragomir R. Radev, Simone Teufel, Horacio Saggion, Wai Lam, John Blitzer, Arda Çelebi, Hong Qi, Elliott Drabek, and Danyu Liu. Evaluation of Text Summarization in a Cross-lingual Information Retrieval Framework. Technical report, Center for Language and Speech Processing, Johns Hopkins University, Baltimore, MD, June 2002.

[291] Dragomir R. Radev, Harris Wu, and Weiguo Fan. Towards Answer-Focused Summarization. In *Proceedings of the 1st International Conference on Information Technology and Applications*, Bathurst, Australia, November 25–28 2002.

[292] G. Rath, A. Resnick, and R. Savage. The Formation of Abstracts by the Selection of Sentences: Part 1: Sentence Selection by Man and Machines. *American Documentation*, 12(2):139–141, 1961.

[293] Lisa F. Rau and Ron Brandow. Domain-Independent Summarization of News. In *Dagstuhl Seminar, Summarizing Text for Intelligent Communication*. December 1993.

[294] Lisa F. Rau and Paul Jacobs. Creating Segmented Databases from Free Text for Text Retrieval. In *Proceedings of the 14th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 337–346, New York, NY, 1991.

[295] Lisa F. Rau, Paul S. Jacobs, and Udi Zernik. Information Extraction and Text Summarization Using Linguistic Knowledge Acquisition. *Information Processing & Management*, 25(4):419–428, 1989.

[296] Gisela Redeker. Ideational and Pragmatic Markers of Discourse Structure. *Journal of Pragmatics*, 14:367–381, 1990.

[297] Lynne M. Reder and John R. Anderson. A Comparison of Texts and Their Summaries: Memorial Consequences. *Journal of Verbal Learning and Verbal Behavior*, 19:121–134, 1980.

[298] Ulrich Reimer and Udo Hahn. Text Condensation as Knowledge-based Abstraction. In *Proceedings of the 4th Conference on Artificial Intelligence Applications*, pages 338–344, March 1988.

[299] Ehud Reiter and Robert Dale. *Building Natural Language Generation Systems*. Cambridge University Press, Cambridge, U.K., 2000.

[300] Ellen Riloff. A Corpus-Based Approach to Domain-Specific Text Summarisation: A Proposal. In Brigitte Endres-Niggemeyer, Jerry Hobbs, and Karen Sparck-Jones, editors, *Workshop on Summarising Text for Intelligent Communication*. Dagstuhl, Germany, 1993.

[301] Lucia H. M. Rino and Donia Scott. Automatic Generation of Draft Summaries: Heuristics for Content Selection. Technical Report ITRI-94-8, Information Technology Research Institute, 1994.

[302] Lucia H. M. Rino and Donia Scott. Content Selection in Summary Generation. Technical report, Dublin City University, Ireland, July 1994.

[303] Jacques Robin. Revision-Based Generation of Natural Language Summaries Providing Historical Background: Corpus Analysis, Design, Implementation and Evaluation. Technical report cucs-034-94, Columbia University, December 1994.

[304] Jacques Robin and Kathleen R. McKeown. Empirically Designing and Evaluating a New Revision-based Model for Summary Generation. *Artificial Intelligence*, 1995.

[305] Jennifer Rowley. *Abstracting and Indexing*. Bingley, London, UK, 1982.

[306] David E. Rumelhart. Understanding and Summarising Brief Stories. In D. Laberge and S.J. Samuels, editors, *Basic Processes in Reading: Perception and Comprehension*, pages 265–303. Lawrence Erlbaum Associates, 1977.

[307] James E. Rush, Antonio Zamora, and R. Salvador. Automatic Abstracting and Indexing. II, Production of Abstracts by Application of Contextual Inference and Syntactic Coherence Criteria. *Journal of the American Society for Information Science*, 22(4):260–274, 1971.

[308] Pamela Russell. Investigating Summary Typology: Considerations for Classification. *Technostyle*, 11 3/4 Spring/Fall Issue:37–47, 1994.

[309] Bogdan Sacaleanu, Paul Buitelaar, and Martin Volk. A cross language document retrieval system based on semantic annotation. In *Proceedings of the 11th Meeting of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary, April 12–17 2003.

[310] Horacio Saggion. Using Linguistic Knowledge in Automatic Abstracting. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pages 596–601, Maryland, USA, June 1999.

[311] Horacio Saggion. *Génération automatique de résumés par analyse sélective*. PhD thesis, Département d'informatique et de recherche opérationnelle. Faculté des arts et des sciences. Université de Montréal, August 2000.

[312] Horacio Saggion, Kalina Bontcheva, and Hamish Cunningham. Robust Generic and Query-Based Summarization. In *Proceedings of the 11th Meeting of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary, April 12–17 2003.

[313] Horacio Saggion and Guy Lapalme. Concept Identification and Presentation in the Context of Technical Text Summarization. In Udo Hahn, Chin-Yew Lin, Inderjeet Mani, and Dragomir R. Radev, editors, *Proceedings of the Workshop on Automatic Summarization at the 6th Applied Natural Language Processing Conference and the 1st Conference of the North American Chapter of the Association for Computational Linguistics*, Seattle, WA, USA, April 30 2000. Association for Computational Linguistics.

[314] Horacio Saggion and Guy Lapalme. Selective Analysis for Automatic Abstracting: Evaluating Indicativeness and Acceptability. In *Proceedings of the Computer-Assisted Information Searching on Internet Conference. RIAO'2000*, Paris, France, April 12–14 2000.

[315] Horacio Saggion and Guy Lapalme. Generating Indicative-Informative Summaries with SumUM. *Computational Linguistics*, 28(4), 2002.

[316] Gerard Salton. *Automatic Text Processing*. Addison-Wesley Publishing Company, 1988.

[317] Gerard Salton, James Allan, Chris Buckley, and Amit Singhal. Automatic Analysis, Theme Generation, and Summarization of Machine-Readable Texts. *Science*, 264:1421–1426, 1994.

[318] Gerard Salton, James Allan, and Amit Singhal. Automatic Text Decomposition and Structuring. *Information Processing & Management*, 32(2):127–138, 1996.

[319] Gerard Salton, Amit Singhal, Chris Buckley, and Mandar Mitra. Automatic Text Decomposition Using Text Segments and Text Themes. Technical Report Technical Report TR-95-1555, Department of Computer Science, Cornell University, 1995.

[320] Gerard Salton, Amit Singhal, Mandar Mitra, and Chris Buckley. Automatic Text Structuring and Summarization. *Information Processing & Management*, 33(2):193–207, 1997.

[321] Antonio Sanfilippo. Conditions on Consistency of Probabilistic Tree Adjoining Grammars. In *Proceedings of the 17th International Conference on Computational Linguistics*, Montreal, Canada, August 10–14 1998.

[322] Tefko Saracevic. Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science. *Journal of the American Society for Information Science*, 26(6):321–343, 1975.

[323] Satoshi Sato and Madoka Sato. Rewriting Saves Extracted Summaries. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[324] Linda Schamber, Michael B. Eisenberg, and Michael S. Nilan. A Re-Examination of Relevance: Toward a Dynamic, Situational Definition. *Information Processing and Management*, 26:755–776, 1990.

[325] Robert Schank and Robert Abelson. *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Associates, Publishers, 1977.

[326] Barry Schiffman. Building a Resource for Evaluating the Importance of Sentences. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Spain, May–June 2002.

[327] Judith D. Schlesinger and Deborah J. Baker. Using Document Features and Statistical Modeling to Improve Query-based Summarization. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[328] Judith D. Schlesinger, Mary Ellen Okurowski, John M. Conroy, Dianne P. O'Leary, Anthony Taylor, Jean Hobbs, and Harold T. Wilson. Understanding Machine Performance in the Context of Human Performance for Multi- Document Summarization. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[329] Satoshi Sekine and Chikashi Nobata. Sentence Extraction with Information Extraction Technique. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[330] Satoshi Sekine and Chikashi Nobata. A Survey for Multi-Document Summarization. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 65–72, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[331] Carol Sherrard. The Psychology of Summary Writing. *JTWC*, 15(3):247–258, 1985.

[332] Gregory H. Silber and Kathleen McCoy. Efficient Text Summarization Using Lexical Chains. In *Proceedings of the ACM Conference on Intelligent User Interfaces (IUI'2000)*, January 9–12 2000.

[333] Gregory H. Silber and Kathleen McCoy. Efficiently Computed Lexical Chains As An Intermediate Representation in Automatic Text Summarization. *Computational Linguistics*, 28(4), 2002.

[334] Eduard F. Skorochod'ko. Adaptive Method of Automatic Abstracting and Indexing. In C. Freiman, editor, *Information Processing 71: Proceedings of the IFIP Congress 71*, pages 1179–1182. North-Holland Publishing Company, 1972.

[335] Harold Somers, Bill Black, Jeremy Ellman, Luca Gilardoni, Torbjoern Lager, Annarosa Multari, Joakim Nivre, and Alex Rogers. Multilingual Generation and Summarization of Job Adverts: The TREE Project. In *Proceedings of the 5th Conference on Applied Natural Language Processing*, pages 269–276, 1997.

[336] Karen Sparck-Jones. Discourse Modelling for Automatic Summarising. Technical Report Technical Report No. 290, University of Cambridge, Computer Laboratory, 1993.

[337] Karen Sparck-Jones. What Might Be In A Summary. Information Retrieval 93: Von der Modellierung zur Anwendung, 9–26, 1993.

[338] Karen Sparck-Jones. Summarising: Where are we now? where should we go? In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, July 1997.

[339] Karen Sparck-Jones. Automatic Summarizing: Factors and Directions. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 1–13. The MIT Press, 1999.

[340] Karen Sparck-Jones. Factorial Summary Evaluation. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[341] Karen Sparck-Jones and Tetsuya Sakai. Generic Summaries for Indexing in IR. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 190–198, New Orleans, LA, September 2001.

[342] Karen Spark-Jones and Julia R. Galliers. *Evaluating Natural Language Processing Systems: An Analysis and Review*. Number 1083 in Lecture Notes in Artificial Intelligence. Springer, 1995.

[343] Gees C. Stein, Amit Bagga, and G. Bowden Wise. Evaluating Summaries for Multiple Documents in an Interactive Environment. In *Proceedings of the 1st International Conference on Language Resources and Evaluation*, pages 1651–1657, May 2000.

[344] Gees C. Stein, Amit Bagga, and G. Bowden Wise. Multi-Document Summarization: Methodologies and Evaluations. In *Proceedings of the 7th Conference on Automatic Natural Language Processing TALN*, pages 337–346, Lausanne, Switzerland, October 2000.

[345] Tomek Strzalkowski. Robust Natural Language Processing and User-Guided Concept Discovery for Information Retrieval, Extraction, and Summarization: Tipster Phase III. In *In TIPSTER Text Phase III Kickoff Workshop*. Columbia, Maryland, October 1996.

[346] Tomek Strzalkowski, Gees Stein, Jing Wang, and Bowden Wise. A Robust Practical Text Summarizer. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 137–154. The MIT Press, 1999.

[347] Kazuo Sumita, Kenji Ono, and Seiji Miike. Document Structure Extraction for Interactive Document Retrieval Systems. In *Proceedings of the 11th Annual International ACM Conference on Systems Documentation*, pages 301–310, Waterloo, Ontario, Canada, 1993.

[348] Stan Szpakowicz and Terry Copeck. Coherence in Summaries. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[349] John I. Tait. Automatic Summarising of English Texts. Technical Report 47, University of Cambridge, Computer Laboratory, 1982.

[350] John I. Tait. Generating Summaries Using a Script Based Language Analyzer. In *Progress in Artificial Intelligence*, 1985.

[351] Naoyuki Tamura. Formalization and Implementation of Summary Generation. *Journal of the Japanese Society for Artificial Intelligence*, 4 (2):196–206, 1989.

[352] Simone Teufel. Meta-Discourse Markers and Problem-Structuring in Scientific Texts. In M. Stede, L. Wanner, and Eduard Hovy, editors, *Proceedings of the Workshop on Discourse Relations and Discourse Markers at the 17th International Conference on Computational Linguistics*, pages 43–49, August 15 1998.

[353] Simone Teufel and Marc Moens. Sentence Extraction as a Classification Task. In Inderjeet Mani and Mark T. Maybury, editors, *Proceedings of the Workshop on Intelligent Scalable Text Summarization at the 35th Meeting of the Association for Computational Linguistics, and the 8th Conference of the European Chapter of the Assocation for Computational Linguistics*, Madrid, Spain, 1997.

[354] Simone Teufel and Marc Moens. Sentence Extraction and Rhetorical Classification for Flexible Abstracts. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 16–25, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[355] Simone Teufel and Marc Moens. Argumentative Classification of Extracted Sentences as a First Step Towards Flexible Abstracting. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 155–171. The MIT Press, 1999.

[356] Simone Teufel and Marc Moens. Summarising Scientific Articles - Experiments with Relevance and Rhetorical Status. *Computational Linguistics*, 28(4), 2002.

[357] Anastasios Tombros, Mark Sanderson, and Phil Gray. Advantages of Query Biased Summaries in Information Retrieval. In Eduard Hovy and Dragomir R. Radev, editors, *Proceedings of the AAAI Symposium on Intelligent Text Summarization*, pages 34–43, Stanford, California, USA, March 23–25 1998. The AAAI Press.

[358] Thomas Trabasso and Linda Sperry. *Causal Relatedness and the Importance of Narrative Events*, volume 24, pages 595–611. 1985.

[359] Robin Valenza, Tony Robinson, Marianne Hickey, and Roger Tucker. Summarization of Spoken Audio Through Information Extraction. In *Proceedings of the ESCA Workshop: Accessing Information in Spoken Audio*, pages 111–116, 1999.

[360] P. van den Broek and Thomas Trabasso. Causal Networks Versus Goal Hierarchies in Summarising Text. *Discourse Processes*, 9:1–15, 1986.

[361] Teun A van Dijk. Recalling and Summarizing Complex Discourse. In W. Burchart and K. Hulker, editors, *Text Processing*, 1979.

[362] Teun A van Dijk. *News as Discourse*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1988.

[363] Hans van Halteren and Simone Teufel. Examining the Consensus between Human Summaries: Initial Experiments with Factoid Analysis. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 57–64, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[364] Alex Waibel, Michael Bett, and Michael Finke. Meeting Browser: Tracking and Summarising Meetings. In *Proceedings of the DARPA Broadcast News Workshop*, 1998.

[365] Takahiro Wakao, Terumasa Ehara, and Katsuhiko Shirai. Text Summarization for Production of Closed-Caption TV Programs in Japanese. *Computer Processing of Oriental Languages*, 12(1):87–97, 1998.

[366] Ke Wang and Huiquing Liu. Discovering Typical Structures of Documents: A Road Map Approach. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 146–154, Melbourne, Australia, 1998.

[367] Wen Wang and Mary P. Harper. The SuperARV Language Model: Investigating the Effectiveness of Tighty Integrated Multiple Knowledge Sources. In *Proceedings of the 40th Meeting of the Association for Computational Linguistics*, July 6–13 2002.

[368] Mark Wasson. Using Summaries in Document Retrieval. In Udo Hahn and Donna Harman, editors, *Proceedings of the Workshop on Text Summarization at the 40th Meeting of the Association for Computational Linguistics*, July 12–13 2002.

[369] Michael White and Claire Cardie. Selecting Sentences for Multidocument Summaries Using Randomized Local Search. In Udo Hahn and Donna Harman, editors, *Proceedings of the Workshop on Text Summarization at the 40th Meeting of the Association for Computational Linguistics*, pages 9–18, Philadelphia, July 11–12 2002.

[370] Michael White, Claire Cardie, Vincent Ng, Kiri Wagstaff, and Daryl Mc-Cullough. Detecting Discrepancies and Improving Intelligibility: Two Preliminary Evaluations of RIPTIDES. In *Proceedings of the 1st Document Understanding Conference*, New Orleans, LA, 2001.

[371] Ryen White, Joemon M. Jose, and Ian Ruthven. Query-biased Web Page Summarization: A Task- Oriented Evaluation. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 412–413, New Orleans, LA, 2001.

[372] Ryen W. White, Joemon M. Jose, and Ian Ruthven. A Task-Oriented Study on the Influencing Effects of Query-Biased Summarisation in Web-Searching. *Information Processing and Management*, 39:707–733, 2003.

[373] Peter N. Winograd. Strategic Difficulties in Summarizing Texts. *Reading Research Quarterly*, 19(4):404–425, 1984.

[374] Michael Witbrock and Vibhu O. Mittal. Ultra-Summarization: A Statistical Approach to Generating Highly Condensed Non-Extractive Summaries. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 315–316, Berkeley, CA, 1999.

[375] Yaakov Yaari. Segmentation of Expository Texts by Hierarchical Agglomerative Clustering. Technical report also available as cmp-lg/9709015, Bar-Ilan University, Israel, 1997.

[376] Yiming Yang, Tom Ault, Thomas Pierce, and Charles W. Lattimer. Improving Text Categorization Methods for Event Tracking. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 65–72, Athens, Greece, 2000.

[377] Sheryl R. Young and Philip J. Hayes. Automatic Classification and Summarization of Banking Telexes. In *Proceedings of the 2nd Conference on Artificial Intelligence Applications (CAIA)*, pages 402–408, Miami Beach, FL, December 1985.

[378] David Zajic and Bonnie Dorr. Automatic Headline Generation for News-paper Stories. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[379] Klaus Zechner. Automatic Text Abstracting by Selecting Relevant Passages. Master's thesis, Centre for Cognitive Science, University of Edinburgh, 1995.

[380] Klaus Zechner. *Automatic Summarization of Spoken Dialogues in Unrestricted Domains.* PhD thesis, Carnegie Mellon University, School of Computer Science,Language Technologies Institute, November 2001.

[381] Klaus Zechner. Automatic Summarization of Open Domain Multi-Party Dialogues in Diverse Genres. *Computational Linguistics*, 28(4), 2002.

[382] Klaus Zechner and Alon Lavie. Increasing the Coherence of Spoken Dialogue Summaries by Cross-Speaker Information Linking. In Jade Goldstein and Chin-Yew Lin, editors, *Proceedings of the Workshop on Automatic Summarization at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 22–31, 2001.

[383] Klaus Zechner and Alex Waibel. Minimizing Word Error Rate in Textual Summaries of Spoken Language. In *Proceedings of the 6th Applied Natural Language Processing Conference and the 1st Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 186–193, 2000.

[384] Dmitry Zelenko, Chinatsu Aone, and Anthony Richardella. Kernel Methods for Relation Extraction. In *Proceedings of the 39th Meeting of the Association for Computational Linguistics*, July 6–13 2002.

[385] Hongyuan Zha. Generic Summarization and Key Phrase Extraction Using Mutual Reinforcement Principle and Sentence Clustering. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Tampere, Finland, 2002.

[386] Hongyuan Zha and Xiang Ji. Summaries with SumUM: a Text Summarization System and its Expansion for Document Understanding Conference. In *Proceedings of the Workshop on Multi-Document Summarization Evaluation of the 2nd Document Understanding Conference at the 40th Meeting of the Association for Computational Linguistics*, Philadelphia, PA, July 2002.

[387] Haiqin Zhang, Zheng Chen, Wei-ying Ma, and Qingsheng Cai. A Study for Document Summarization Based on Personal Annotation. In Dragomir

Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 41–48, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[388] Zhu Zhang, Sasha Blair-Goldensohn, and Dragomir R. Radev. Towards CST-Enhanced Summarization. In *Proceedings of the 18th National Conference on Artificial Intelligence*, Edmonton, Alberta, August 2002.

[389] Liang Zhou and Eduard Hovy. Headline Summarization at ISI. In *DUC03*, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

[390] Liang Zhou and Eduard Hovy. A Web-Trained Extraction Summarization System. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 284–290, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.