Bibliography

Abbreviations used in this bibliography:

ACL n Proceedings of the nth Annual Meeting of the Association for Computational Linguistics

COLING n Proceedings of the nth International Conference on Computational Linguistics (COLING-*year*)

EACL n Proceedings of the nth Conference of the European Chapter of the Association for Computational Linguistics

Steven Abney. Parsing by chunks. In Robert C. Berwick, Steven P. Abney, and Carol Tenny, editors, *Principle-Based Parsing*, pages 257–278. Kluwer Academic, Dordrecht; Boston, 1991.

Steven Abney. Part-of-speech tagging and partial parsing. In Ken Church, Steve Young, and Gerrit Bloothooft, editors, *Corpus-Based Methods in Language and Speech*. Kluwer Academic Publishers, Dordrecht, 1996a.

Steven Abney. Statistical methods and linguistics. In Judith L. Klavans and Philip Resnik, editors, *The Balancing Act: Combining symbolic and statistical approaches to language*, pages x–y. MIT Press, Cambridge, MA, 1996b.

Alfred V. Aho, Ravi Sethi, and Jeffrey D. Ullman. *Compilers: Principles, Techniques, and Tools*. Addison-Wesley, Reading, MA, 1986.

James Allen. *Natural Language Understanding*. Benjamin Cummings, Redwood City, CA, 1995.

Hiyan Alshawi, Adam L. Buchsbaum, and Fei Xia. A comparison of head transducers and transfer for a limited domain translation application. In *Proceedings of ACL 35 and EACL 8*, pages 360–365, San Francisco CA, 1997. Morgan Kaufmann Publishers.

Hiyan Alshawi and David Carter. Training and scaling preference functions for disambiguation. *Computational Linguistics*, 20(4):635–648, 1994.

Chinatsu Aone and Douglas McKee. Acquiring predicate-argument mapping information from multilingula texts. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 175–190. MIT Press, Cambridge MA, 1995.

D. E. Appelt, J. R. Hobbs, J. Bear, D. Israel, and M. Tyson. Fastus: A finite-state processor for information extraction from real-world text. In *Proc. of the 13th IJCAI*, pages 1172–1178, Chambéry, France, 1993.

Jurij D. Apresjan. Regular polysemy. *Linguistics*, 142:5–32, 1974.

Shlomo Argamon, Ido Dagan, and Yuval Krymolowski. A memory-based approach to learning shallow natural language patterns. In *Proceedings of COLING-ACL*, pages 67–73, Montreal, Canada, 1998.

Eric Atwell. Constituent-likelihood grammar. In Roger Garside, Geoffrey Leech, and Geoffrey Sampson, editors, *The Computational Analysis of English: A Corpus-Based Approach*. Longman, London; New York, 1987.

Harald Baayen and Richard Sproat. Estimating lexical priors for low-frequency morphologically ambiguous forms. *Computational Linguistics*, 22(2):155–166, 1996.

Lalit R. Bahl, Frederick Jelinek, and Robert L. Mercer. A maximum likelihood approach to continuous speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-5:179–190, 1983. Reprinted in Waibel and Lee XXXX.

Lalit R. Bahl and Robert L. Mercer. Part-of-speech assignment by a statistical decision algorithm. In *International Symposium on Information Theory*, Ronneby, Sweden, 1976.

James K. Baker. Stochastic modeling for automatic speech understanding. In A. Waibel and K.-F. Lee (eds) *Readings in Speech Recognition*, Morgan Kaufmann, 1990, 1975.

James K. Baker. Trainable grammars for speech recognition. In D. H. Klatt and J. J. Wolf, editors, *Speech Communication Papers for the 97th Meeting of the Acoustical Society of America*, pages 547–550, 1979.

Geoff Barnbrook. *Language and computers: a practical introduction to the computer analysis of language*. Edinburgh University Press, Edinburgh, 1996.

Roberto Basili, Maria Teresa Pazienza, and Paola Velardi. Integrating general-purpoose and corpus-based verb classification. *Computational Linguistics*, 22 (4):559–568, 1996.

Roberto Basili, Gianluca De Rossi, and Maria Teresa Pazienza. Inducing terminology for lexical acquisition. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, pages 125–133, Somerset NJ, 1997. Association for Computational Linguistics.

Doug Beeferman, Adam Berger, and John Lafferty. Text segmentation using exponential models. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, pages 35–46, Somerset NJ, 1997. Association for Computational Linguistics.

Timothy C. Bell, John G. Cleary, and Ian H. Witten. *Text Compression*. Prentice Hall, Englewood Cliffs, NJ, 1990.

Julian Benello, Andrew W. Mackie, and James A. Anderson. Syntactic category disambiguation with neural networks. *Computer Speech and Language*, 3:203–217, 1989.

M. Benson. The structure of the collocational dictionary. *International Journal of Lexicography*, 2(1), 1989.

Morton Benson, Evelyn Benson, and Robert Ilson. *The BBI combinatory dictionary of English.* John Benjamins Publishing Company, Amsterdam/Philadelphia, 1993.

A. P. Berber Sardinha. *Automatic Identification of Segments in Written Texts*. PhD thesis, University of Liverpool, 1997.

Adam L. Berger, Stephen A. Della Pietra, and Vincent J. Della Pietra. A maximum entropy approach to natural language processing. *Computational Linguistics*, 22 (1):39–71, 1996.

M. Berry and P. Young. Using latent semantic indexing for multilanguage information retrieval. *Computers and the Humanities*, 29(6):413–429, 1995.

Michael W. Berry. Large-scale sparse singular value computations. *The International Journal of Supercomputer Applications*, 6(1):13–49, 1992.

Michael W. Berry, Susan T. Dumais, and Gavin W. O'Brien. Using linear algebra for intelligent information retrieval. *SIAM Review*, 37(4):573–595, 1995.

Douglas Biber, Susan Conrad, and Randi Reppen. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press, Cambridge, 1998.

E. Black, S. Abney, D. Flickinger, C. Gdaniec, R. Grishman, P. Harrison, D. Hindle, R. Ingria, F. Jelinek, J. Klavans, M. Liberman, M. Marcus, S. Roukos, B. Santorini, and T. Strzalkowski. A procedure for quantitatively comparing the syntactic coverage of English grammars. In *Proceedings, Speech and Natural Language Workshop, Pacific Grove, CA*, pages 306–311. DARPA, 1991.

Ezra Black. An experiment in computational discrimination of English word senses. *IBM Journal of Research and Development*, 32:185–194, 1988.

Ezra Black, Fred Jelinek, John Lafferty, David M. Magerman, Robert Mercer, and Salim Roukos. Towards history-based grammars: Using richer models for probabilistic parsing. In *ACL 31*, pages 31–37, 1993. Also appears in the Proceedings of the DARPA Speech and Natural Language Workshop, Feb. 1992, pp. 134–139.

Rens Bod. *Enriching Linguistics with Statistics: Performance Models of Natural Language*. PhD thesis, University of Amsterdam, 1995.

Bran Boguraev and Ted Briscoe. *Computational Lexicography for Natural Language Processing*. Longman, London; New York, 1989.

Branimir Boguraev and James Pustejovsky. Issues in text-based lexicon acquisition. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 3–17. MIT Press, Cambridge MA, 1995.

Branimir K. Boguraev. The contribution of computational lexicography. In Madeleine Bates and Ralph M. Weischedel, editors, *Challenges in natural language processing*, pages 99–132. Cambridge University Press, Cambridge, 1993.

Susan Bonzi and Elizabeth D. Liddy. The use of anaphoric resolution for document description in information retrieval. In *SIGIR 88*, pages 53–66, 1988.

Abraham Bookstein and Don R. Swanson. A decision theoretic foundation for indexing. *JASIS*, 26(1):45–50, 1975.

T. L. Booth. Probabilistic representation of formal languages. In *Tenth Annual IEEE Symposium on Switching and Automata Theory*, pages xx–yy, 1969.

Andrew Borthwick, John Sterling, Eugene Agichtein, and Ralph Grishman. Exploting diverse knowledge sources via maximum entropy in named entity recognition. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 152–160, New Brunswick NJ, 1998. Association for Computational Linguistics.

Didier Bourigault. An endogeneous corpus-based method for structural noun phrase disambiguation. In *Proceedings of EACL 6*, pages 81–86. Association of Computational Linguistics, 1993.

G. Box and G. Tiao. *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Reading, MA, 1973.

Thorsten Brants. Estimating Hidden Markov Model Topologies. In Jonathan Ginzburg, Zurab Khasidashvili, Carl Vogel Jean-Jacques Lévy, and Enric Vallduví, editors, *The Tbilisi Symposium on Logic, Language and Computati on: Selected Papers*, pages 163 – 176. CSLI Publications, Stanford, California, 1998.

Thorsten Brants and Wojciech Skut. Automation of treebank annotation. In *Proceedings of NeMLaP-98*, Sydney, Australia, 1998.

L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Wadsworth International Group, Belmont CA, 1984.

Michael R. Brent. From grammar to lexicon: Unsupervised learning of lexical syntax. *Computational Linguistics*, 19(2):243–262, 1993.

Eric Brill. Automatic grammar induction and parsing free text: A transformation-based approach. In *Proceedings of ACL 31*, pages 259–265, Morristown NJ, 1993a. Association of Computational Linguistics.

Eric Brill. *A Corpus-Based Approach to Language Learning*. PhD thesis, University of Pennsylvania, 1993b.

Eric Brill. Transformation-based error-driven parsing. In *Proceedings Third International Workshop on Parsing Technologies*, pages x–y, Tilburg/Durbuy, The Netherlands/Belgium, 1993c.

Eric Brill. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational Linguistics*, 21 (4):543–565, 1995.

Eric Brill, David Magerman, Mitch Marcus, and Beatrice Santorini. Deducing linguistic structure from the statistics of large corpora. In *Proceedings of the DARPA Speech and Natural Language Workshop*, pages 275–282, San Mateo CA, 1990. Morgan Kaufmann.

Eric Brill and Philip Resnik. A transformation-based approach to prepositional phrase attachment disambiguation. In *Proceedings of COLING*, Kyoto, Japan, 1994.

Ted Briscoe and John Carroll. Generalized probabilistic LR parsing of natural language (corpora) with unification-based methods. *Computational Linguistics*, 19: 25–59, 1993.

J. L. Britton, editor. *Collected Works of A. M. Turing: Pure Mathematics*. North-Holland, Amsterdam, 1992.

Peter F. Brown, , John Cocke, Stephen A. Della Pietra, Vincent J. Della Pietra, Fredrick Jelinek, John D. Lafferty, Robert L. Mercer, and Paul S. Roossin. A statistical approach to machine translation. *Computational Linguistics*, 16(2):79–85, 1990.

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, Jennifer C. Lai, and Robert L. Mercer. An estimate of an upper bound for the entropy of English. *Computational Linguistics*, 18:31–40, 1992a.

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. Word-sense disambiguation using statistical methods. In *Proceedings of ACL 29*, pages 264–270, Morristown NJ, 1991a. Association of Computational Linguistics.

Peter F. Brown, Jennifer C. Lai, and Robert L. Mercer. Aligning sentences in parallel corpora. In *ACL 29*, pages 169–176, 1991b.

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, John D. Lafferty, and Robert L. Mercer. Analysis, statistical transfer, and synthesis in machine translation. In *Proceedings of the 4th International Conference on Theoretical and Methodological Issues in Machine Translation*, pages 83–100, 1992b.

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. A statistical approach to sense disambiguation in machine translation. In *Proceedings of the DARPA Workshop on Speech and Natural Language Workshop*, pages 146–151, 1991c.

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19(2):263–311, 1993.

Peter F. Brown, Vincent J. Della Pietra, Peter V. deSouza, Jenifer C. Lai, and Robert L. Mercer. Class-based *n*-gram models of natural language. *Computational Linguistics*, 18(4):467–479, 1992c.

Rebecca Bruce and Janyce Wiebe. Word-sense disambiguation using decomposable models. In *Proceedings of ACL 32*, pages 139–145, Morristown NJ, 1994. Association of Computational Linguistics.

Rebecca F. Bruce and Janyce M. Wiebe. Decomposable modeling in natural language processing. *Computational Linguistics*, 1999. to appear.

Jennifer Brundage, Maren Kresse, Ulrike Schwall, and Angelika Storrer. Multiword lexemes: A monolingual and contrastive typology for natural language processing and machine translation. Technical Report 232, Institut fuer Wissensbasierte Systeme, IBM Deutschland GmbH, Heidelberg, 1992.

Chris Buckley, Amit Singhal, Mandar Mitra, and Gerard Salton. New retrieval approaches using SMART: TREC 4. In *The Second Text REtrieval Conference (TREC-2)*, pages 25–48, 1996. In (Harman 1994).

Paul Buitelaar. *CoreLex: Systematic Polysemy and Underspecification*. PhD thesis, Brandeis University, 1998.

Curt Burgess and Kevin Lund. Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, 12:177–210, 1997.

Robin Burke, Kristian Hammond, Vladimir Kulyukin, Steven Lytinen, Noriko Tomuro, and Scott Schoenberg. Question answering from frequently asked question files ... *AI Magazine*, 18(2):57–66, 1997.

Claire Cardie. Empirical methods in information extraction. *AI Magazine*, 18(4): 65–79, 1997.

G. Carroll and E. Charniak. Two experiments on learning probabilistic dependency grammars from corpora. In Carl Weir, Stephen Abney, Ralph Grishman, and Ralph Weischedel, editors, *Working Notes of the Workshop Statistically-Based NLP Techniques*, pages 1–13. AAAI Press, Menlo Park, CA, 1992.

John Carroll. Relating complexity to practical performance in parsing with wide-coverage unification grammars. In *ACL 32*, pages 287–294, 1994.

Jason S. Chang and Mathis H. Chen. An alignment method for noisy parallel corpora based on image processing techniques. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, pages 297–304, 1997.

Jean-Pierre Chanod and Pasi Tapanainen. Tagging French - comparing a statistical and a constraint-based method. In *Proceedings of EACL 7*, pages 149–156, Dublin, Ireland, 1995. Association for Computational Linguistics.

Eugene Charniak. *Statistical Language Learning*. MIT Press, Cambridge, MA, 1993.

Eugene Charniak. Tree-bank grammars. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI '96)*, pages 1031–1036, 1996.

Eugene Charniak. Statistical techniques for natural language parsing. *AI Magazine*, pages 33–43, 1997.

Eugene Charniak, Curtis Hendrickson, Neil Jacobson, and Mike Perkowitz. Equations for part-of-speech tagging. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 784–789, Menlo Park CA, 1993.

Peter Cheeseman, James Kelly, Matthew Self, John Stutz, Will Taylor, and Don Freeman. AutoClass: A Bayesian classification system. In *Proceedings of the Fifth International Conference on Machine Learning*, pages 54–64, San Francisco CA, 1988. Morgan Kaufmann.

Jen Nan Chen and Jason S. Chang. Topical clustering of MRD senses based on information retrieval techniques. *Computational Linguistics*, 24(1):61–95, 1998.

Stanley F. Chen. Aligning sentences in bilingual corpora using lexical information. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pages 9–16, 1993.

Stanley F. Chen and Joshua Goodman. An empirical study of smoothing techniques for language modeling. In *ACL 34*, pages 310–318, 1996.

Stanley F. Chen and Joshua Goodman. An empirical study of smoothing techniques for language modeling. Technical Report TR-10-98, Center for Research in Computing Technology, Harvard University, 1998.

Zhiyi Chi and Stuart Geman. Estimation of probabilistic context-free grammars. *Computational linguistics*, 24:299–305, 1998.

Mahesh V. Chitrao and Ralph Grishman. Statistical parsing of messages. In *Proceedings of the DARPA Speech and Natural Language Workshop, Hidden Valley, PA*, pages 263–266. Morgan Kaufmann, 1990.

Noam Chomsky. *Syntactic Structures*. Mouton, The Hague, 1957.

Noam Chomsky. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA, 1965.

Noam Chomsky. *Rules and Representations*. Columbia University Press, New York, 1980.

Noam Chomsky. *Knowledge of Language: Its Nature, Origin, and Use*. Prager, New York, 1986.

Yaacov Choueka. Looking for needles in a haystack or locating interesting collocational expressions in large textual databases. In *Proceedings of the RIAO*, pages 43–38, 1988.

Yaacov Choueka and Serge Lusignan. Disambiguation by short contexts. *Computers and the Humanities*, 19:147–158, 1985.

Kenneth Church, William Gale, Patrick Hanks, and Donald Hindle. Using statistics in lexical analysis. In Uri Zernik, editor, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 115–164. Lawrence Erlbaum, Hillsdale NJ, 1991.

Kenneth Church and Ramesh Patil. Coping with syntactic ambiguity or how to put the block in the box on the table. *Computational Linguistics*, 8(3-4):139–149, 1982.

Kenneth W. Church. A stochastic parts program and noun phrase parser for unrestricted text. In *Proceedings of ANLP*, pages 136–143, Morristown NJ, 1988. Association of Computational Linguistics.

Kenneth W. Church and William A. Gale. A comparison of the enhanced Good-Turing and deleted estimation methods for estimating probabilities of English bigrams. *Computer Speech and Language*, 5:19–54, 1991a.

Kenneth W. Church and William A. Gale. Concordances for parallel text. In *Proceedings of the Seventh Annual Conference of the UW Centre for the New OED and Text Research*, pages 40–62, England: Oxford, 1991b.

Kenneth W. Church and William A. Gale. Poisson mixtures. *Natural Language Engineering*, 1:163–190, 1995.

Kenneth W. Church and Robert L. Mercer. Introduction to the special issue on computational linguistics using large corpora. *Computational Linguistics*, 19:1–24, 1993.

Kenneth Ward Church. Char_align: A program for aligning parallel texts at the character level. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pages 1–8, 1993.

Kenneth Ward Church. One term or two? In *Proceedings of SIGIR '95*, pages 310–318, New York, 1995. Association for Computing Machinery.

Kenneth Ward Church and Patrick Hanks. Word association norms, mutual information and lexicography. In *Proceedings of ACL 27*, pages 76–83, Vancouver, Canada, 1989.

Kenneth Ward Church and Mark Y. Liberman. A status report on the ACL/DCI. In *Proceedings of the 7th Annual Conference of the UW Centre for New OED and Text Research: Using Corpora*, pages 84–91, 1991.

Eve Clark and Herbert Clark. When nouns surface as verbs. *Language*, 55:767–811, 1979.

Cyril W. Cleverdon and J. Mills. The testing of index language devices. *Aslib Proceedings*, 15(4):106–130, 1963. Reprinted in (Sparck Jones and Willett 1998).

Sam Coates-Stephens. The analysis and acquisition of proper names for the understanding of free text. *Computers and the Humanities*, 26:441–456, 1993.

Michael John Collins. A new statistical parser based on bigram lexical dependencies. In *ACL 34*, pages 184–191, 1996.

Michael John Collins. Three generative, lexicalised models for statistical parsing. In *ACL 35/EACL 8*, pages 16–23, 1997.

Michael John Collins and James Brooks. Prepositional phrase attachment through a backed-off model. In *Proceedings of the Third Workshop on Very Large Corpora*, pages 27–38, 1995.

Ann Copestake and Ted Briscoe. Semi-productive polysemy and sense extension. *Journal of Semantics*, 12(1):15–68, 1995.

Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*. MIT Press, Cambridge, MA, 1990.

Garrison W. Cottrell. *A Connectionist Approach to Word Sense Disambiguation*. Pitman, London, 1989.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, 1991.

Wayne Cowart. *Experimental syntax: Applying objective methods to sentence judgments*. Sage Publications, Thousand Oaks CA, 1997.

W. B. Croft and D. J. Harper. Using probabilistic models of document retrieval without relevance information. *Journal of Documentation*, 35(4):285–295, 1979.

Terry Crowley, John Lynch, Jeff Siegel, and Julie Piau. *The Design of Language: An introduction to descriptive linguistics*. Longman Paul, Auckland, 1995.

David Crystal. *The Cambridge Encyclopedia of Language*. Cambridge University Press, Cambridge, England, 1987.

Doug Cutting, Julian Kupiec, Jan Pedersen, and Penelope Sibun. A practical part-of-speech tagger. In *The 3rd Conference on Applied Natural Language Processing*, pages 133–140, Trento, Italy, 1991.

Douglas R. Cutting, David R. Karger, and Jan O. Pedersen. Constant interaction-time scatter/gather browsing of very large document collections. In *Proceedings of SIGIR'93*, 1993.

Douglas R. Cutting, Jan O. Pedersen, David Karger, and John W. Tukey. Scatter/gather: A cluster-based approach to browsing large document collections. In *Proceedings of SIGIR '92*, pages 318–329, New York, 1992. Association of Computing Machinery.

Walter Daelemans and Antal van den Bosch. Language-independent data-oriented grapheme-to-phoneme conversion. In J. Van Santen, R. Sproat, J. Olive, and J. Hirschberg, editors, *Progress in Speech Synthesis*, pages 77–90. Springer Verlag, New York, 1996.

Walter Daelemans, Jakub Zavrel, Peter Berck, and Steven Gillis. Mbt: A memory-based part of speech tagger generator. In E. Ejerhed and I. Dagan, editors, *Proceedings of the Fourth Workshop on Very Large Corpora*, pages 14–27, Copenhagen, 1996. ACL SIGDAT.

Ido Dagan, Kenneth Church, and William Gale. Robust bilingual word alignment for machine aided translation. In *Proceedings of the Workshop on Very Large Corpora*, pages 1–8, Morristown NJ, 1993. Association for Computational Linguistics.

Ido Dagan and Alon Itai. Word sense disambiguation using a second language monolingual corpus. *Computational Linguistics*, 20(4):563–596, 1994.

Ido Dagan, Alon Itai, and Ulrike Schwall. Two languages are more informative than one. In *Proceedings of ACL 29*, pages 130–137, Morristown NJ, 1991. Association of Computational Linguistics.

Ido Dagan, Yael Karov, and Dan Roth. Mistake-driven learning in text categorization. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, pages 55–63, Somerset NJ, 1997a. Association for Computational Linguistics.

Ido Dagan, Lillian Lee, and Fernando Pereira. Similarity-based methods for word sense disambiguation. In *Proceedings of ACL 35 and EACL 8*, pages 56–63, San Francisco CA, 1997b. Morgan Kaufmann Publishers.

Ido Dagan, Fernando Pereira, and Lillian Lee. Similarity-based estimation of word cooccurrence probabilities. In *Proceedings of the ACL 32*, pages 272–278, Morristown NJ, 1994. Association of Computational Linguistics.

Fred J. Damerau. Generating and evaluating domain-oriented multi-word terms from texts. *Information Processing & Management*, 29(4):433–447, 1993.

J. N. Darroch and D. Ratcliff. Generalized iterative scaling for log-linear models. *The Annals of Mathematical Statistics*, 43(3):1470–1480, 1972.

Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407, 1990.

Morris H. DeGroot. *Probability and Statistics*. Addison-Wesley, Reading, MA, 1975.

Stephen Della Pietra, Vincent Della Pietra, and John Lafferty. Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (4), 1997.

A.J. Demers. Generalized left corner parsing. In *Proceedings of the Fourth Annual ACM Symposium on Principles of Programming Languages*, pages 170–181, 1977.

A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Society Series B*, 39:1–38, 1977.

Evangelos Dermatas and George Kokkinakis. Auotmatic stochastic tagging of natural language texts. *Computational Linguistics*, 21(2):137–164, 1995.

Steven J. DeRose. Grammatical category disambiguation by statistical optimization. *Computational Linguistics*, 14(1):31–39, 1988.

Anne-Marie Derouault and Bernard Merialdo. Natural language modeling for phoneme-to-text transcription. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):742–649, 1986.

Luca Dini, Vittorio Di Tomaso, and Frédérique Segond. Error-driven word sense disambiguation. In *Proceedings of ACL 36*, 1998.

William B. Dolan. Word sense ambiguation: Clustering related senses. In *Proceedings of COLING*, pages 712–716, Kyoto, Japan, 1994.

Pedro Domingos and Michael Pazzani. On the optimality of the simple bayesian classifier under zero-one loss. *Machine Learning*, 29:103–130, 1997.

Bonnie J. Dorr and Mari Broman Olsen. Deriving verbal and compositional lexical aspect for nlp applications. In *Proceedings of ACL 35 and EACL 8*, pages 151–158, San Francisco CA, 1997. Morgan Kaufmann Publishers.

Mark Dras and Mike Johnson. Death and lightness: Using a demographic model to find support verbs. In *Proceedings of the 5th International Conference on the Cognitive Science of Natural Language Processing*, pages x–y, Dublin, 1996.

Richard O. Duda and Peter E. Hart. *Pattern classification and scene analysis*. Wiley, New York, 1973.

Susan T. Dumais. Latent semantic indexing (lsi): Trec-3 report. In *The Second Text REtrieval Conference (TREC 2)*, pages 219–230, 1995.

Ted Dunning. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74, 1993.

Jason Eisner. Three new probabilistic models for dependency parsing: An exploration. In *COLING 16*, pages 340–345, 1996.

Jeffrey L. Elman. Finding structure in time. *Cognitive Science*, 14:179–211, 1990.

David Elworthy. Does baum-welch re-estimation help taggers? In *Proceedings of ANLP 4*, pages 53–58, Morristown NJ, 1994. Association of Computational Linguistics.

J. B. Estoup. *Gammes Sténographiques*. Paris, 4th edition, 1916.

David A. Evans, Kimberly Ginther-Webster, Mary Hart, Robert G. Lefferts, and Ira A. Monarch. Automatic indexing using selective nlp and first-order thesauri. In *Proceedings of the RIAO*, volume 2, pages 624–643, 1991.

David A. Evans and Chengxiang Zhai. Noun-phrase analysis in unrestricted text for information retrieval. In *Proceedings of ACL 34*, pages 17–24, 1996.

Joel L. Fagan. Automatic phrase indexing for document retrieval: An examination of syntactic and non-syntactic methods. In *SIGIR 87*, pages 91–101, 1987.

Joel L. Fagan. The effectiveness of a nonsyntactic approach to automatic phrase indexing for document retrieval. *JASIS*, 40(2):115–132, 1989.

Robert M. Fano. *Transmission of information; a statistical theory of communications*. MIT Press, New York, 1961.

C. J. Fillmore and B. T. S. Atkins. Starting where the dictionaries stop: the challenge of corpus lexicography. In B.T.S. Atkins and A. Zampolli, editors, *Computational Approaches to the Lexicon*. Oxford University Press, Oxford, England, 1994.

Steven Finch and Nick Chater. Distributional bootstrapping: From word class to proto-sentence. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pages 301–306, Hillsdale NJ, 1994. Lawrence Erlbaum.

Steven Paul Finch. *Finding Structure in Language*. PhD thesis, University of Edinburgh, 1993.

J. R. Firth. A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis*, pages 1–32. Philological Society, Oxford, 1957. Reprinted in F. R. Palmer (ed), *Selected Papers of J. R. Firth 1952–1959*, London: Longman, 1968.

Thierry Fontenelle, Walter Brüls, Luc Thomas, Tom Vanallemeersch, and Jacques Jansen. Decide, mlap-project 93-19, deliverable d-1a: survey of collocation extraction tools. Technical report, University of Liege, Liege, Belgium, 1994. Available at http://engdep1.philo.ulg.ac.be/decide/reports.

Marilyn Ford, Joan Bresnan, and Ronald M. Kaplan. A competence-based theory of syntactic closure. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 727–796. MIT Press, Cambridge, MA, 1982.

William B. Frakes and Ricardo Baeza-Yates, editors. *Information Retrieval*. Prentice Hall, Englewood Cliffs NJ, 1992.

W. Nelson Francis and Henry Kučera. *Manual of information to accompany a standard corpus of present-day edited American English, for use with digital computers*. Dept of Linguistics, Brown University, Providence, RI, 1964.

W. Nelson Francis and Henry Kučera. *Frequency Analysis of English Usage: Lexicon and Grammar*. Houghton Mifflin, Boston, MA, 1982.

Alexander Franz. *Automatic Ambiguity Resolution in Natural Language Processing*, volume 1171 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, Berlin, 1996.

Alexander Franz. Independence assumptions considered harmful. In *Proceedings of ACL 35 and EACL 8*, pages 182–189, Morristown NJ, 1997. Association of Computational Linguistics.

Lyn Frazier. *On Comprehending Sentences: Syntactic Parsing Strategies*. PhD thesis, University of Connecticut, 1978.

David Freedman, Robert Pisani, and Roger Purves. *Statistics*. W. W. Norton, New York, 1998. 3rd ed.

King-Sun Fu. *Syntactic Methods in Pattern Recognition*. Academic Press, London; New York, 1974.

Pascale Fung and Kenneth W. Church. K-vec: A new approach for aligning parallel texts. In *Proceedings of the Fifteenth International Conference on Computational Linguistics (COLING-94)*, pages 1096–1102, 1994.

Pascale Fung and Kathleen McKeown. Aligning noisy parallel corpora across language groups: Word pair feature matching by dynamic time warping. In *Proceedings of the Association for Machine Translation in the Americas (AMTA-94)*, pages 81–88, 1994.

William A. Gale and Kenneth W. Church. Estimation procedures for language context: Poor estimates of context are worse than none. In *Proceedings in Computational Statistics (COMPSTAT 9)*, pages 69–74, 1990a.

William A. Gale and Kenneth W. Church. Poor estimates of context are worse than none. In *Proceedings of the June 1990 DARPA Speech and Natural Language Workshop*, pages 283–287, Hidden Valey, PA, 1990b.

William A. Gale and Kenneth W. Church. A program for aligning sentences in bilingual corpora. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 177–184, 1991.

William A. Gale and Kenneth W. Church. A program for aligning sentences in bilingual corpora. *Computational Linguistics*, 19:75–102, 1993.

William A. Gale and Kenneth W. Church. What's wrong with adding one? In Nelleke Oostdijk and Pieter de Haan, editors, *Corpus-Based Research into Language: in honour of Jan Aarts*. Rodopi, Amsterdam, 1994.

William A. Gale, Kenneth W. Church, and David Yarowsky. Estimating upper and lower bounds on the performance of word-sense disambiguation programs. In *Proceedings of ACL 30*, pages 249–256, Morristown NJ, 1992a. Association of Computational Linguistics.

William A. Gale, Kenneth W. Church, and David Yarowsky. A method for disambiguating word senses in a large corpus. *Computers and the Humanities*, 26: 415–439, 1992b.

William A. Gale, Kenneth W. Church, and David Yarowsky. A method for disambiguating word senses in a large corpus. Technical report, AT&T Bell Laboratories, Murray Hill NJ, 1992c.

William A. Gale, Kenneth W. Church, and David Yarowsky. Using bilingual materials to develop word sense disambiguation methods. In *Proceedings of the 4th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-92)*, pages 101–112, 1992d.

William A. Gale, Kenneth W. Church, and David Yarowsky. Work on statistical methods for word sense disambiguation. In Robert Goldman, Peter Norvig, Eugene Charniak, and Bill Gale, editors, *Working Notes of the AAAI Fall Symposium on Probabilistic Approaches to Natural Language*, pages 54–60, Menlo Park CA, 1992e. AAAI Press.

William A. Gale and Geoffrey Sampson. Good-turing frequency estimation without tears. *Journal of Quantitative Linguistics*, 2(3):217–237, 1995.

Robert G. Gallager. *Information theory and reliable communication*. Wiley, New York, 1968.

Roger Garside. Grammatical tagging of the spoken part of the British National Corpus: a progress report. In Geoffrey N. Leech, Greg Myers, and Jenny Thomas, editors, *Spoken English on computer: transcription, mark-up, and application*. Longman, Harlow, Essex; New York, 1995.

Roger Garside and Geoffrey Leech. xxx. In Roger Garside, Geoffrey Leech, and Geoffrey Sampson, editors, *The Computational Analysis of English: A Corpus-Based Approach*. Longman, London; New York, 1987.

Roger Garside, Geoffrey Sampson, and Geoffrey Leech, editors. *The Computational analysis of English: a corpus-based approach*. Longman, London; New York, 1987.

Éric Gaussier. Flow network models for word alignment and terminology extraction from bilingual corpora. In *Proceedings of COLING-ACL*, pages 444–450, Montreal, Canada, 1998.

Niyu Ge, John Hale, and Eugene Charniak. A statistical approach to anaphora resolution. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 161–170, New Brunswick NJ, 1998. Association for Computational Linguistics.

Zoubin Ghahramani. Solving inverse problems using an EM approach to dnesity estimation. In Michael C. Mozer, Paul Smolensky, David S. Touretzky, and Andreas S. Weigend, editors, *Proceedings of the 1993 Connectionist Models Summer School*, Hillsdale NJ, 1994. Erlbaum Associates.

Ted Gibson and N. Pearlmutter. A corpus-based analysis of psycholinguistics constraints on pp attachment. In Jr. Charles Clifton, Lyn Frazier, and Keith Rayner, editors, *Perspectives on Sentence Processing*, pages x–y. Lawrence Erlbaum, Hillsdale, NJ, 1994.

E. Mark Gold. Language identification in the limit. *Information and Control*, 10: 447–474, 1967.

Moises Goldszmidt and Mehran Sahami. A probabilistic approach to full-text document clustering. Technical Report SIDL-WP-1998-0091, Stanford Digital Library Project, Stanford CA, 1998.

Gene H. Golub and Charles F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, 1989.

I. J. Good. The population frequencies of species and the estimation of population parameters. *Biometrika*, 40:237–264, 1953.

I. J. Good. Studies in the history of probability and statistics. xxxvii: A. m. turing's statistical work in world war ii. *Biometrika*, 66:393–396, 1979.

Sidney Greenbaum. The tagset for the international corpus of english. In Eric Atwell and Clive Souter, editors, *Corpus-based Computational Linguistics*, pages 11–24. Rodopi, Amsterdam, 1993.

Barbara B. Greene and Gerald M. Rubin. Automatic grammatical tagging of english. Technical report, Brown University, Providence RI, 1971.

Gregory Grefenstette. Finding semantic similarity in raw text: the deese antonyms. In Robert Goldman, Peter Norvig, Eugene Charniak, and Bill Gale, editors, *Working Notes of the AAAI Fall Symposium on Probabilistic Approaches to Natural Language*, pages 61–65, Menlo Park CA, 1992a. AAAI Press.

Gregory Grefenstette. Use of syntactic context to produce term association lists for text retrieval. In *Proceedings of SIGIR '92*, pages 89–97, 1992b.

Gregory Grefenstette. *Explorations in Automatic Thesaurus Discovery*. Kluwer Academic Press, Boston, 1994.

Gregory Grefenstette. Evaluation techniques for automatic semantic extraction: Comparing syntactic and window-based approaches. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, chapter 11, pages 205–216. MIT Press, Cambridge, MA, 1996.

Gregory Grefenstette, editor. *Cross-language information retrieval*. Kluwer Academic Publishers, Boston MA, 1998.

Gregory Grefenstette and Pasi Tapanainen. What is a word, what is a sentence? Problems of tokenization. In *Proceedings of the Third International Conference on Computational Lexicography (COMPLEX '94)*, pages 79–87, Budapest, 1994. Available as Rank Xerox Research Centre technical report MLTT-004.

R. Günter, L. B. Levitin, B. Shapiro, and P. Wagner. Zipf's law and the effect of ranking on probability distributions. *International Journal of Theoretical Physics*, 35(2):395–417, 1996.

Joe A. Guthrie, Louise Guthrie, Yorick Wilks, and Homa Aidinejad. Subject-dependent co-occurrence and word sense disambiguation. In *Proceedings of ACL 29*, pages 146–152, Morristown NJ, 1991. Association of Computational Linguistics.

Louise Guthrie, James Pustejovsky, Yorick Wilks, and Brian M. Slator. The role of lexicons in natural language processing. *Communications of the ACM*, 39(1): 63–72, 1996.

Michael Halliday. Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, and R. H. Robins, editors, *In memory of J. R. Firth*, pages 148–162. Longmans, London, 1966.

D.K. Harman, editor. *The Second Text REtrieval Conference (TREC-2)*. U.S. Department of Commerce, Washington DC, 1994. NIST Special Publication 500-215.

D.K. Harman, editor. *The Third Text REtrieval Conference (TREC-4)*. U.S. Department of Commerce, Washington DC, 1996.

Stevan Harnad, editor. *Categorical perception: the groundwork of cognition*. Cambridge University Press, Cambridge, 1987.

B. Harris. Bi-text, a new concept in translation theory. *Language Monthly*, 54, 1988.

Zellig Harris. *Methods in Structural Linguistics*. University of Chicago Press, Chicago, 1951.

Philip Harrison, Steven Abney, Ezra Black, Dan Flickinger, Ralph Grishman Claudia Gdaniec, Donald Hindle, Robert Ingria, Mitch Marcus, Beatrice Santorini, and Tomek Strzalkowski. Natural Language Processing Systems Evaluation Workshop, Technical Report rl-tr-91-362. In Jeannette G. Neal and Sharon M. Walter, editors, *Evaluating Syntax Performance of Parser/Grammars of English*, 1991.

Steve Harter. A probabilistic approach to automatic keyword indexing: Part ii. an algorithm for probabilistic indexing. *JASIS*, 26(4):280–289, 1975.

Masahiko Haruno and Takefumi Yamazaki. High-performance bilingual text alignment using statistical and dictionary information. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, pages 131–138, 1996.

Vasileios Hatzivassiloglou and Kathleen R. McKeown. Towards the automatic identification of adjectival scales: clustering adjectives according to meaning. In *Proceedings of ACL 31*, pages 172–182, Morristown NJ, 1993. Association of Computational Linguistics.

Mark Hawthorne. The computer in literary analysis: Using TACT with students. *Computers and the Humanities*, 28(1):19–27, 1994.

Marti Hearst and Christian Plaunt. Subtopic structuring for full-length document access. In *Proceedings of SIGIR '93*, pages 59–68, 1993.

Marti A. Hearst. Noun homograph disambiguation using local context in large text corpora. In *Seventh Annual Conference of the UW Centre for the New OED and Text Research*, pages 1–22, England: Oxford, 1991.

Marti A. Hearst. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of Coling-92*, pages 539–545. Association of Computational Linguistics, 1992.

Marti A. Hearst. Texttiling: A quantitative approach to discourse segmentation. Technical Report Sequoia 93/24, Center for Research on Language, University of California at Berkeley, 1993.

Marti A. Hearst. *Context and Structure in Automated Full-Text Information Access*. PhD thesis, University of California at Berkeley, 1994.

Marti A. Hearst. Texttiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics*, 23(1):33–64, 1997.

Marti A. Hearst and Hinrich Schütze. Customizing a lexicon to better suit a computational task. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 77–96. MIT Press, Cambridge MA, 1995.

James Henderson and Peter Lane. A connectionist architecture for learning to parse. In *Proceedings of COLING-ACL*, pages 531–537, Montreal, Canada, 1998.

Ulf Hermjakob and Raymond J. Mooney. Learning parse and translation decisions from examples with rich context. In *ACL 35/EACL 8*, pages 482–489, 1997.

John A. Hertz, Richard G. Palmer, and Anders S. Krogh. *Introduction to the theory of neural computation*. Addison-Wesley, Redwood City CA, 1991.

Eric van Herwijnen. *Practical SGML*. Kluwer Academic, Boston/Dordrecht/London, 2nd edition, 1994.

Donald Hindle. Noun classification from predicate argument structures. In *Proceedings of the ACL 28*, pages 268–275, Morristown NJ, 1990. Association of Computational Linguistics.

Donald Hindle. A parser for text corpora. In B.T.S. Atkins and A. Zampolli, editors, *Computational Approaches to the Lexicon*, pages 103–151. Oxford University Press, Oxford, England, 1994.

Donald Hindle and Mats Rooth. Structural ambiguity and lexical relations. *Computational Linguistics*, 19(1):103–120, 1993.

Graeme Hirst. *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press, Cambridge, 1987.

Julia Hodges, Shiyun Yie, Ray Reighart, and Lois Boggess. An automated system that assists in the generation of document indexes. *Natural Language Engineering*, 2(2):137–160, 1996.

John E. Hopcroft and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley, Reading MA, 1979.

A. S. Hornby. *Oxford Advanced Learner's Dictionary of Current English*. Oxford University Press, Oxford, 1974. Third Edition.

James Jay Horning. *A study of grammatical inference*. PhD thesis, Stanford, 1969.

Rodney Huddleston. *Introduction to the Grammar of English*. Cambridge University Press, Cambridge, 1984.

David Hull. Stemming algorithms - A case study for detailed evaluation. *Journal of the American Society for Information Science*, 47(1):70–84, 1996.

David Hull. A practical approach to terminology alignment. In Didier Bourigault, Christian Jacquemin, and Marie-Claude L'Homme, editors, *Proceedings of Computerm '98*, pages 1–7, Montreal, Canada, 1998.

David Hull and Doug Oard, editors. *AAAI Symposium on Cross-Language Text and Speech Retrieval*. AAAI Press, Stanford CA, 1997.

David A. Hull and Gregory Grefenstette. Querying across languages: A dictionary-based approach to multilingual information retrieval. In Karen Sparck Jones and Peter Willett, editors, *Readings in Information Retrieval*. Morgan Kaufmann, San Francisco, 1998.

David A. Hull, Jan O. Pedersen, and Hinrich Schütze. Method combination for document filtering. In *Proceedings of SIGIR '96*, pages 279–287, Zurich, 1996.

Nancy Ide and Jean Véronis, editors. *The Text Encoding Initiative: Background and Context*. Kluwer Academic, Dordrecht/Boston/London, 1995. Reprinted from *Computers and the Humanities* 29(1–3), 1995.

Nancy Ide and Jean Véronis. Introduction to thes pecial issue on word sense disambiguation: The state of the art. *Computational Linguistics*, 24(1):1–40, 1998.

Nancy Ide and Donald Walker. Introduction: Common methodologies in humanities computing and computational linguistics. *Computers and the Humanities*, 26:327–330, 1992.

K. Inui, V. Sornlertlamvanich, H. Tanaka, and T. Tokunaga. A new formalization of probabilistic glr parsing. In *Proceedings of the Fifth International Workshop on Parsing Technologies (IWPT-97)*, pages 123–134, MIT, 1997.

Pierre Isabelle. Machine translation at the TAUM group. In Margaret King, editor, *Machine Translation Today: Tthe State of the Art*, pages 247–277. Edinburgh University Press, Edinburgh, 1987.

Christian Jacquemin. Fastr : A unification-based front-end to automatic indexing. In *Proceedings of RIAO*, pages 34–47, Rockefeller University, New York, 1994.

Christian Jacquemin, Judith L. Klavans, and Evelyne Tzoukermann. Expansion of multi-word terms for indexing and retrieval using morphology and syntax. In *Proceedings of ACL 35 and EACL 8*, pages 24–31, Morristown NJ, 1997. Association of Computational Linguistics.

Anil K. Jain and Richard C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs NJ, 1988.

Harold Jeffreys. *Theory of Probability*. Clarendon Press, Oxford, 1948.

F. Jelinek, J. Lafferty, D. Magerman, R. Mercer, A. Ratnaparkhi, and S. Roukos. Decision tree parsing using a hidden derivation model. In *Proceedings of the 1994 Human Language Technology Workshop*, pages 272–277. DARPA, 1994.

F. Jelinek, J. D. Lafferty, and R. L. Mercer. Basic methods of probabilistic context free grammars. Technical Report RC 16374 (#72684), IBM T. J. Watson Research Center, 1990.

F. Jelinek, J. D. Lafferty, and R. L. Mercer. Basic methods of probabilistic context free grammars. In P. Laface and R. De Mori, editors, *Speech Recognition and Understanding: Recent Advances, Trends, and Applications*, volume 75 of *F: Computer and Systems Sciences*. Springer Verlag, 1992a.

Fred Jelinek. Self-organized language modeling for speech recognition. MS, IBM T.J. Watson research center. Reprinted in Waibel and Lee XXXX, 1990.

Fred Jelinek and John D. Lafferty. Computation of the probability of initial substring generation by stochastic context-free grammars. *Computational Linguistics*, 17:315–324, 1991.

Fred Jelinek and Robert Mercer. Probability distribution estimation from sparse data. *IBM Technical Disclosure Bulletin*, 28:2591–2594, 1985.

Frederick Jelinek. Fast sequential decoding algorithm using a stack. *IBM Journal of Research and Development*, pages 675–685, 1969.

Frederick Jelinek. Continuous speech recognition by statistical methods. *IEEE*, 64: 532–556, 1976.

Frederick Jelinek. Markov source modeling of text generation. In J. K. Skwirzynski, editor, *The Impact of Processing Techniques on Communications*, volume E91 of *NATO ASI series*, pages 569–598. M. Nijhoff, Dordrecht; Boston, 1985. (Jelinek et al. 1992b) covers the same material and is more easily accessible.

Frederick Jelinek. *Statistical Methods for Speech Recognition*. MIT Press, Cambridge, MA, 1997.

Frederick Jelinek, Lalit R. Bahl, and Robert L. Mercer. Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE Transactions on Information Theory*, 21:250–256, 1975.

Frederick Jelinek, Robert L. Mercer, and Salim Roukos. Principles of lexical language modeling for speech recognition. In Sadaoki Furui and M. Mohan Sondhi, editors, *Advances in Speech Signal Processing*, pages 651–699. Marcel Dekker, New York, 1992b.

Karen Jensen, George E. Heidorn, and Stephen D. Richardson, editors. *Natural language processing: The PLNLP approach*. Kluwer Academic Publishers, Boston, 1993.

Stig Johansson, G. N. Leech, and H. Goodluck. *Manual of information to accompany the Lancaster-Oslo/Bergen Corpus of British English, for use with digital computers*. Dept of English, University of Oslo, Oslo, 1978.

Mark Johnson. The effect of alternative tree representations on tree bank grammars. In *Proceedings of Joint Conference on New Methods in Language Processing and Computational Natural Language Learning (NeMLaP3/CoNLL98)*, pages 39–48, Macquarie University, 1998.

W. E. Johnson. Probability: deductive and inductive problems. *Mind*, 41:421–423, 1932.

Martin Joos. Review of *the psycho-biology of language*. *Language*, 12:196–210, 1936.

Julia Jorgensen. The psychological reality of word senses. *Journal of Psycholinguistic Research*, 19:167–190, 1990.

Aravind K. Joshi. Tree-adjoining grammars. In R. E. Asher, editor, *The Encyclopedia of Language and Linguistics*. Pergamon Press, Oxford, 1993.

John S. Justeson and Slava M. Katz. Co-occurrences of antonymous adjectives and their contexts. *Computational Linguistics*, 17(1):1–19, 1991.

John S. Justeson and Slava M. Katz. Principled disambiguation: Discriminating adjective senses with modified nouns. *Computational Linguistics*, 24(1):1–28, 1995a.

John S. Justeson and Slava M. Katz. Technical terminology: some linguistic properties and an algorithm for identification in text. *Natural Language Engineering*, 1(1):9–27, 1995b.

Daniel Kahneman, Paul Slovic, and Amos Tversky, editors. *Judgment under uncertainty : heuristics and biases*. Cambridge University Press, Cambridge; New York, 1982.

Min-Yen Kan, Judith L. Klavans, and Kathleen R. McKeown. Linear segmentation and segment significance. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 197–205, New Brunswick NJ, 1998. Association for Computational Linguistics.

F. Karlsson, A. Voutilainen, J. Heikkilä, and A. Anttila. *A Language-Independent System for Parsing Unrestricted Text*. Mouton de Gruyter, Berlin; New York, 1995.

Yael Karov and Shimon Edelman. Similarity-based word sense disambiguation. *Computational Linguistics*, 24(1):41–59, 1998.

Lauri Karttunen. Radical lexicalism. Technical Report 86–68, Center for the Study of Language and Information, Stanford CA, 1986.

Slava M. Katz. Estimation of probabilities from sparse data for the language model component of a speech recognizer. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35:400–401, 1987.

Slava M. Katz. Distribution of content words and phrases in text and language modelling. *Natural Language Engineering*, 2:15–59, 1996.

Leonard Kaufman and Peter J. Rousseeuw. *Finding groups in data*. Wiley, New York, 1990.

Martin Kay and Martin Röscheisen. Text-translation alignment. *Computational Linguistics*, 19:121–142, 1993.

Andrew Kehler. Probabilistic coreference in information extraction. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, pages 163–173, Somerset NJ, 1997. Association for Computational Linguistics.

Edward Kelly and Phillip Stone. *Computer Recognition of English Word Senses*. North-Holland, Amsterdam, 1975.

André Kempe. Finite state transducers approximating hidden markov models. In *Proceedings of ACL 35 and EACL 8*, pages 460–467, Morristown NJ, 1997. Association of Computational Linguistics.

Graeme Kennedy. *An Introduction to Corpus Linguistics*. Longman, x, 1998.

Roland G. Kent. Review of *relative frequency as a determinant of phonetic change*. *Language*, 6:86–88, 1930.

Adam Kilgarriff. Dictionary word sense distinctions: An enquiry into their nature. *Computers and the Humanities*, 26:365–387, 1993.

Adam Kilgarriff. I don't believe in word senses. *Computers and the Humanities*, 32 (2), 1998.

Adam Kilgarriff and Tony Rose. Metrics for corpus similarity and homogeneity. Manuscript, ITRI, University of Brighton, 1998.

Judith Klavans and Min-Yen Kan. Role of verbs in document analysis. In *Proceedings of COLING-ACL*, pages 680–686, Montreal, Canada, 1998.

Judith L. Klavans and Evelyne Tzoukermann. Dictionaries and corpora: Combining corpus and machine-readable dictionary da ta for building bilingual lexicons. *Journal of Machine Translation*, 10(3–4), 1995.

Sheldon Klein and Robert F. Simmons. A computational approach to grammatical coding of english words. *Journal of the Association for Computing Machinery*, 10: 334–347, 1963.

Reinhard Kneser and Hermann Ney. Improved backing-off for m-gram language modeling. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 181–184, 1995.

Kevin Knight. Automating knowledge acquisition for machine translation. *AI Magazine*, 18(4):81–96, 1997.

Kevin Knight, Ishwar Chander, Matthew Haines, Vasileios Hatzivassiloglou, Eduard Hovy, Masayo Iida, Steve Luk, Richard Whitney, and Kenji Yamada. Filling knowledge gaps in a broad-coverage mt system. In *Proceedings of IJCAI-95*, 1995.

Kevin Knight and Jonathan Graehl. Machine transliteration. In *Proceedings of ACL 35 and EACL 8*, pages 128–135, San Francisco CA, 1997. Morgan Kaufmann Publishers.

Kevin Knight and Vasileios Hatzivassiloglou. Two-level, many-paths generation. In *Proceedings of ACL 33*, Boston MA, 1995.

Teuvo Kohonen. *Self-Organizing Maps*. Springer Verlag, Berlin, Heidelberg, New York, 1997. Second Extended Edition.

Robert R. Korfhage. *Information Storage and Retrieval*. John Wiley, Berlin, 1997.

Brigitte Krenn and Christer Samuelsson. The linguist's guide to statistics. MS, University of Saarbrucken, available at: http://www.coli.uni-sb.de/ christer/, 1997.

Robert Krovetz. Lexical acquisition and information retrieval. In Uri Zernik, editor, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 45–64. Lawrence Erlbaum, Hillsdale NJ, 1991.

J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964a.

J. B. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29:115–129, 1964b.

Henry Kučera and W. Nelson Francis. *Computational Analysis of Present-Day American English.* Brown University Press, Providence, RI, 1967.

Julian Kupiec. A trellis-based algorithm for estimating the parameters of a hidden stochastic context-free grammar. In *Proceedings of the Speech and Natural Language Workshop*, pages 241–246. DARPA, 1991.

Julian Kupiec. An algorithm for estimating the parameters of unrestricted hidden stochastic context-free grammars. In *Coling 14*, pages 387–393, 1992a.

Julian Kupiec. Robust part-of-speech tagging using a Hidden Markov Model. *Computer Speech and Language*, 6:225–242, 1992b.

Julian Kupiec. An algorithm for finding noun phrase correspondences in bilingual corpora. In *Proceedings of ACL 31*, pages 17–22, Morristown NJ, 1993a. Association of Computational Linguistics.

Julian Kupiec. Murax: A robust linguistic approach for question answering using an on-line encyclopedia. In *Proceedings of SIGIR '93*, pages 181–190, 1993b.

Julian Kupiec, Jan Pedersen, and Francine Chen. A trainable document summarizer. In *Proceedings of SIGIR 95*, pages 68–73, 1995.

K. L. Kwok and M. Chan. Improving two-stage ad-hoc retrieval for short queries. In *Proceedings of SIGIR '98*, pages 250–256, Melbourne, 1998.

John Lafferty, Daniel Sleator, and Davy Temperley. Grammatical trigrams: A probabilistic model of link grammar. In *Proceedings of the 1992 AAAI Fall Symposium on Probabilistic Approaches to Natural Language*, 1992.

George Lakoff. *Women, fire, and dangerous things*. University of Chicago Press, Chicago, IL, 1987.

Pierre Simon marquis de Laplace. *Essai philosophique sur les probabilites*. Mme. Ve. Courcier, Paris, 1814.

Pierre Simon marquis de Laplace. *PHILOSOPHICAL ESSAY ON PROBABILITIES*. Springer-Verlag, New York, 1995.

K. Lari and S. J. Young. The estimation of stochastic context-free grammars using the inside-outside algorithm. *Computer Speech and Language*, 4:35–56, 1990.

K. Lari and S. J. Young. Application of stochastic context free grammar using the inside-outside algorithm. *Computer Speech and Language*, 5:237–257, 1991.

Ray Lau, Ronald Rosenfeld, and Salim Roukos. Adaptive language modeling using the maximum entropy principle. In *Proceedings of the Human Language Technology Workshop*, pages 108–113. ARPA, 1993.

Raymond Lau. Adaptive statistical language modelling. Master's thesis, Massachusetts Institute of Technology, 1994.

Mark Lauer. Corpus statistics meet the noun compound: Some empirical results. In *Proceedings of ACL 33*, Boston MA, 1995a.

Mark Lauer. *Designing Statistical Language Learners: Experiments on Noun Compounds*. PhD thesis, Macquarie University, Sydney, Australia, 1995b.

Claudia Leacock, Martin Chodorow, and George A. Miller. Using corpus statistics and wordnet relations for sense identification. *Computational Linguistics*, 24(1): 147–165, 1998.

M. E. Lesk. Word-word association in document retrieval systems. *American Documentation*, 20(1):27–38, 1969.

Michael Lesk. Automatic sense disambiguation: How to tell a pine cone from an ice cream cone. In *Proceedings of the 1986 SIGDOC Conference*, pages 24–26, New York, 1986. Association for Computing Machinery.

Beth Levin. *English Verb Classes and Alternations*. The University of Chicago Press, Chicago, 1993.

John R. Levine, Tony Mason, and Doug Brown. *Lex & Yacc*. O'Reilly & Associates, Sebastopol CA, 1992. 2nd edition.

S. E. Levinson, L. R. Rabiner, and M M. Sondhi. An introduction to the application of the theory of probabilistic functions of a markov process to automatic speech recongition. *Bell System Technical Journal*, 62:1035–1074, 1983.

David D. Lewis. An evaluation of phrasal and clustered representations on a text categorization task. In *Fifteenth Annual International Conference on Research & Development in Information Retrieval*, pages 37–50, 1992.

David D. Lewis and Karen Sparck Jones. Natural language processing for information retrieval. *Communications of the ACM*, 39(1):92–101, 1996.

David D. Lewis and Marc Ringuette. A comparison of two learning algorithms for text categorization. In *Proc. SDAIR 94*, pages 81–93, Las Vegas, NV, 1994.

David D. Lewis, Robert E. Schapire, James P. Callan, and Ron Papka. Training algorithms for linear text classifiers. In *Proceedings of SIGIR '96*, pages 298–306, Zurich, 1996.

Hang Li and Naoki Abe. Generalizing case frames using a thesaurus and the mdl principle. In *Proceedings of Recent Advances in Natural Language Processing*, pages 239–248, Tzigov Chark, Bulgaria, 1995.

Hang Li and Naoki Abe. Learning dependencies between case frame slots. In *Proceedings of COLING 96*, pages 10–15, Copenhagen, Denmark, 1996.

Hang Li and Naoki Abe. Word clustering and disambiguation based on co-occurrence data. In *Proceedings of COLING-ACL*, pages 749–755, Montreal, Canada, 1998.

Wentian Li. Random texts exhibit zipf's-law-like word frequency distribution. *IEEE Transactions on Information Theory*, 38(6):1842–1845, 1992.

G. J. Lidstone. Note on the general case of the bayes-laplace formula for inductive or *a priori* probabilities. *Transactions of the Faculty of Actuaries*, 8:182–192, 1920.

Marc Light. Morphological cues for lexical semantics. In *Proceedings of ACL 34*, pages 25–31, 1996.

Nick Littlestone. Comparing several linear-threshold learning algorithms on tasks involving superfluous attributes. In A. Prieditis, editor, *Proceedings of the 12th International Conference on Machine Learning*, pages 353–361, San Francisco CA, 1995. Morgan Kaufmann.

Michael L. Littman, Susan T. Dumais, and Thomas K. Landauer. Automatic cross-language information retrieval using latent semantic indexing. In Gregory Grefenstette, editor, *Cross Language Information Retrieval*. Kluwer, 1998a.

Michael L. Littman, Fan Jiang, and Greg A. Keim. Learning a language-independent representation for terms from a partially aligned corpus. In Jude Shavlik, editor, *Proceedings of the Fifteenth International Conference on Machine Learning*, pages 314–322. Morgan Kaufmann, 1998b.

Robert M. Losee, editor. *Text Retrieval and Filtering*. Kluwer Academic Publishers, Boston MA, 1998.

Julie Beth Lovins. Development of a stemming algorithm. *Translation and Computational Linguistics*, 11(1):22–31, 1968.

H. P. Luhn. Keyword-in-context index for technical literature (KWIC index). *American Documentation*, 11:288–295, 1960.

John Lyons. *Introduction to Theoretical Linguistics*. Cambridge University Press, Cambridge, 1968.

David J. C. MacKay and Linda C. Peto. Speech recognition using hidden markov models. *The Lincoln Laboratory Journal*, 3(1):41–62, 1990.

David M. Magerman. *Natural language parsing as statistical pattern recognition*. PhD thesis, Stanford University, 1994.

David M. Magerman. Statistical decision-tree models for parsing. In *ACL 33*, pages 276–283, 1995.

David M. Magerman and Mitchell P. Marcus. Pearl: A probabilistic chart parser. In *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*, 1991. Also published in the Proceedings of the 2nd International Workshop for Parsing Technologies.

David M. Magerman and Carl Weir. Efficiency, robustness, and accuracy in picky chart parsing. In *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, 1992.

Benoit Mandelbrot. Structure formelle des textes et communcation. *Word*, 10:1–27, 1954.

Benoit B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman and Company, New York, 1983.

Inderjeet Mani and T. Richard MacMillan. Identifying unknown proper names in newswire text. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 41–59. MIT Press, Cambridge MA, 1995.

Christopher D. Manning. Automatic acquisition of a large subcategorization dictionary from corpora. In *Proceedings of the ACL 31*, pages 235–242, Morristown NJ, 1993. Association of Computational Linguistics.

Christopher D. Manning and Bob Carpenter. Probabilistic parsing using left corner language models. In *Proceedings of the Fifth International Workshop on Parsing Technologies (IWPT-97)*, pages 147–158, MIT, 1997.

Mitchell Marcus, Grace Kim, Mary Ann Marcinkiewicz, Robert MacIntyre, Ann Bies, Mark Ferguson, Karen Katz, and Britta Schasberger. The Penn Treebank: Annotating predicate argument structure. In *ARPA Human Language Technology Workshop*, pages 110–115, 1994.

Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. Building a large annotated corpus of English: The Penn treebank. *Computational Linguistics*, 19:313–330, 1993.

Andrei A. Markov. An example of statistical investigation in the text of 'Eugene Onyegin' illustrating coupling of 'tests' in chains]. In *Proceedings of the Academy of Sciences, St. Petersburg*, volume 7 of *VI*, pages 153–162, 1913.

David Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, New York, 1982.

Ian Marshall. Tag selection using probabilistic methods. In Roger Garside, Geoffrey Sampson, and Geoffrey Leech, editors, *The Computational analysis of English : a corpus-based approach*, pages 42–65. Longman, London; New York, 1987.

James Martin. Representing and acquiring metaphor-based polysemy. In Uri Zernik, editor, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 389–415. Lawrence Erlbaum, Hillsdale NJ, 1991.

W. A. Martin, K. W. Church, and R. S. Patil. Preliminary analysis of a breadth-first parson algorithm: Theoretical and experimental results. In Leonard Bolc, editor, *Natural Language Parsing Systems*. Springer Verlag, Berlin, 1987. Also MIT LCS technical report TR-261.

Brij Masand, Gordon Linoff, and David Waltz. Classifying news stories using memory based reasoning. In *Proceedings of SIGIR '92*, pages 59–65, New York, 1992. Association of Computing Machinery.

John T. Maxwell III. The problem with mutual information. Manuscript, Xerox Palo Alto Research Center, September 15, 1992, 1992.

James L. McClelland, David E. Rumelhart, and the PDP Research Group, editors. *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. The MIT Press, Cambridge MA, 1986.

Peter McCullagh and John A. Nelder. *Generalized Linear Models*, chapter 4, pages 101–123. Chapman and Hall, 2nd edition, 1989.

David D. McDonald. Internal and external evidence in the identification and semantic categorization of proper names. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 21–39. MIT Press, Cambridge MA, 1995.

Tony McEnery and Andrew Wilson. *Corpus Linguistics*. Edinburgh University Press, Edinburgh, 1996.

Sean McGrath. *PARSEME.1ST: SGML for Software Developers*. Prentice Hall PTR, Upper Saddle River, NJ, 1997.

John G. McMahon and Francis J. Smith. Improving statistical language model performance with automatically generated word hierarchies. *Computational Linguistics*, 22(2):217–247, 1996.

C.M. Sperberg McQueen and Lou Burnard, editors. *Guidelines for Electronic Text Encoding and Interchange (TEI P3)*. ACH/ACL/ALLC (Association for Computers and the Humanities, Association for Computational Linguistics, Association for Literary and Linguistic Computing), Chicago, IL, 1994.

Susan W. McRoy. Using multiple knowledge sources for word sense disambiguation. *Computational Linguistics*, 18(1):1–30, 1992.

I. Dan Melamed. A portable algorithm for mapping bitext correspondence. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, pages 305–312, 1997.

I. Dan Melamed. A word-to-word model of translational equivalence. In *Proceedings of ACL 35 and EACL 8*, pages 490–497, San Francisco CA, 1998. Morgan Kaufmann Publishers.

Robert L. Mercer. Inflectional morphology needs to be authenticated by hand. In *Working Notes of the AAAI Spring Syposium on Building Lexicons for Machine Translation*, pages 99–99, Stanford CA, 1993. AAAI Press.

Bernard Merialdo. Tagging english text with a probabilistic model. *Computational Linguistics*, 20(2):155–171, 1994.

Risto Miikkulainen, editor. *Subsymbolic Natural Language Processing*. MIT Press, Cambridge MA, 1993.

Andrei Mikheev. Feature lattices for maximum entropy modelling. In *ACL 36*, pages 848–854, 1998.

George A. Miller and Walter G. Charles. Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1):1–28, 1991.

Marvin Lee Minsky and Seymour Papert, editors. *Perceptrons: an introduction to computational geometry*. MIT Press, Cambridge MA, 1969. Partly reprinted in (Shavlik and Dietterich 1990).

Marvin Lee Minsky and Seymour Papert, editors. *Perceptrons: an introduction to computational geometry*. MIT Press, Cambridge MA, 1988. Expanded edition.

Tom M. Mitchell. The need for biases in learning generalizations. Technical Report Department of Computer Science. CBM-TR-117, Rutgers University, 1980. Reprinted in Jude W. Shavlik and Thomas G. Diettrich (eds), *Readings in Machine Learning*, San Mateo, CA: Morgan Kaufmann, 1990.

Tom M. Mitchell, editor. *Machine Learning*. McGraw-Hill, New York, 1997.

Mandar Mitra, Chris Buckley, Amit Singhal, and Claire Cardie. An analysis of statistical and syntactic phrases. In *Proceedings of RIAO*, 1997.

Alistair Moffat and Justin Zobel. Exploring the similarity space. *ACM SIGIR Forum*, 32(1), 1998.

Alexander M. Mood, Franklin A. Graybill, and Duane C. Boes. *Introduction to the theory of statistics*. McGraw-Hill, New York, 1974. 3rd edition.

Raymond J. Mooney. Comparative experiments on disambiguating word senses: An illustration of the role of bias in machine learning. In *XXXX*, pages x–y, 1996. cmp-lg/9612001.

David S. Moore and George P. McCabe. *Introduction to the practice of statistics*. Freeman, New York, 1989.

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

Frederick Mosteller and David L. Wallace. *Applied Bayesian and Classical Inference - The Case of The Federalist Papers*. Springer Series in Satistics, Springer-Verlag, 1984.

Makoto Nagao. A framework of a mechanical translation between japanese and english by analogy principle. In Alick Elithorn and Ranan B. Banerji, editors, *Artificial and Human Intelligence*, pages 173–180. North-Holland, Edinburgh, 1984.

Mary S. Neff, Brigitte Bläser, Jean-Marc Langé, Hubert Lehmann, and Isabel Zapata Dominguez. Get it where you can: Acquiring and maintaining bilingual lexicons for machine translation. In *Working Notes of the AAAI Spring Syposium on Building Lexicons for Machine Translation*, pages 98–98, Stanford CA, 1993. AAAI Press.

Craig G. Nevill-Manning, Ian H. Witten, and Gordon W. Paynter. Browsing in digital libraries: a phrase-based approach. In *Proceedings of ACM Digital Libraries*, pages 230–236, Philadelphia PA, 1997. Association for Computing Machinery.

Frederick J. Newmeyer. *The Cambridge Survey*. Cambridge University Press, Cambridge, England, 1988.

Hermann Ney and Ute Essen. Estimating 'small' probabilities by leaving-one-out. In *Eurospeech '93*, volume 3, pages 2239–2242. ESCA, 1993.

Hermann Ney, Ute Essen, and Reinhard Kneser. On structuring probabilistic dependencies in stochastic language modeling. *Computer Speech and Language*, 8: 1–28, 1994.

Hermann. Ney, S. Martin, and F. Wessel. Statistical language modeling using leaving-one-out. In Steve Young and Gerrit Bloothooft, editors, *Corpus-Based Methods in Language and Speech Processing*. Kluwer Academic, Dordrecht, 1997.

Hwee Tou Ng and Hian Beng Lee. Integrating multiple knowledge sources to disambiguate word sense: An exemplar-based approach. In *Proceedings of ACL 34*, pages 40–47, 1996.

Jian-Yun Nie, Pierre Isabelle, Pierre Plamondon, and George Foster. Using a probablistic translation model for cross-language information retrieval. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 18–27, New Brunswick NJ, 1998. Association for Computational Linguistics.

S. Nießen, S. Vogel, H. Ney, and C. Tillmann. A DP based search algorithm for statistical machine translation. In *Proceedings of COLING-ACL*, pages 960–967, Montreal, Canada, 1998.

Geoff Nunberg and Annie Zaenen. Systematic polysemy in lexicology and lexicography. In *Proceedings of Euralex II*, Tampere, Finland, 1992.

Geoffrey Nunberg. *The Linguistics of Punctuation*. CSLI Publications, Stanford, CA, 1990.

Douglas W. Oard and Nicholas DeClaris. Cognitive models for text filtering. Unpublished Draft, 1996.

Nicholas Ostler and B. T. S. Atkins. Predictable meaning shift: Some linguistic properties of lexical implication rules. In James Pustejovsky and Sabine Bergler, editors, *Lexical Semantics and Knowledge Representation: Proceedings fof the 1st SIGLEX Workshop*, pages 76–87. Springer Verlag, Berlin, 1992.

Woojin Paik, Elizabeth D. Liddy, Edmund Yu, and Mary McKenna. Categorizing and standardizing proper nouns for efficient information retrieval. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 61–73. MIT Press, Cambridge MA, 1995.

David D. Palmer and Marti A. Hearst. Adaptive sentence boundary disambiguation. In *Proceedings of the Conference on Applied Natural Language Processing (ANLP '94)*, pages 2239–2242. ACL, 1994.

David D. Palmer and Marti A. Hearst. Adaptive multilingual sentence boundary disambiguation. *Computational Linguistics*, 23:241–267, 1997.

Douglas B. Paul. Speech recognition using hidden markov models. *The Lincoln Laboratory Journal*, 3(1):41–62, 1990.

N. Pearlmutter and M. MacDonald. Plausibility and syntactic ambiguity resolution. In *Proceedings of the 14th Annual COnference of the Cognitive Society*, 1992.

Jan O. Pedersen and Craig Silverstein. Almost-constant-time clustering of arbitrary corpus subsets. In *Proceedings of SIGIR '97*, 1997.

Ted Pedersen. Fishing for exactness. In *Proceedings of the South-Central SAS Users Group Conference*, Austin TX, 1996.

Ted Pedersen and Rebecca Bruce. Distinguishing word senses in untagged text. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, Somerset NJ, 1997. Association for Computational Linguistics.

Fernando Pereira and Yves Schabes. Inside-outside reestimation from partially bracketed corpora. In *ACL 30*, pages 128–135, 1992.

Fernando Pereira, Naftali Tishby, and Lillian Lee. Distributional clustering of English words. In *Proceedings of ACL 31*, pages 183–190, Morristown NJ, 1993. Association of Computational Linguistics.

Steven Pinker. *The Language Instinct*. William Morrow, New York, 1994.

Stuart L. Pook and Jason Catlett. Making sense out of searching. In *Information Online 88*, pages 148–157, Sydney, 1988. The Information Science Section of the Library Association of Australia.

M. F. Porter. An algorithm for suffix stripping. *Program*, 14:130–137, 1980.

Victor Poznański and Antonio Sanfilippo. Detecting dependencies between semantic verb subclasses and subcategorization frames in text corpora. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 175–190. MIT Press, Cambridge MA, 1995.

W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 1988.

P. Procter, editor. *Longman dictionary of contemporary English*. Longman Group, Harlow, England, 1978.

E. Prokosch. Review of selected studies of the principle of relative frequency in language. *Language*, 9:89–92, 1933.

James Pustejovsky. The generative lexicon. *Computational Linguistics*, 17(4):409–441, 1991.

James Pustejovsky, Sabine Bergler, and Peter Anick. Lexical semantic techniques for corpus analysis. *Computational Linguistics*, 19(2):331–358, 1993.

Yonggang Qiu and H.P. Frei. Concept based query expansion. In *Proceedings of SIGIR '93*, pages 160–169, 1993.

J. R. Quinlan. Induction of decision trees. *Machine Learning*, 1:81–106, 1986. Reprinted in (Shavlik and Dietterich 1990).

Lawrence Rabiner and Biing-Hwang Juang. *Fundamentals of Speech Recognition*. PTR Prentice-Hall, Englewood Cliffs, NJ, 1993.

Lawrence R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of IEEE*, 77(2):257–286, 1989.

Fred L. Ramsey and Daniel W. Schafer. *The statistical sleuth : a course in methods of data analysis*. Duxbury Press, Belmont CA, 1997.

Lance A. Ramshaw and Mitchell P. Marcus. Exploring the statistical derivation of transformational rule sequences for part-of-speech tagging. In *The Balancing Act. Proceedings of the Workshop.*, pages 86–95, Morristown NJ, 1994. Association of Computational Linguistics.

Edie Rasmussen. Clustering algorithms. In William B. Frakes and Ricardo Baeza-Yates, editors, *Information Retrieval*, pages 419–442. Prentice Hall, Englewood Cliffs NJ, 1992.

Adwait Ratnaparkhi. A maximum entropy model for part-of-speech tagging. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 133–142, Somerset NJ, 1996. Association for Computational Linguistics.

Adwait Ratnaparkhi. A simple introduction to maximum entropy models for natural language processing. Technical Report IRCS Report 97–08, Institute for Research in Cognitive Science, Philadelphia PA, 1997.

Adwait Ratnaparkhi. Unsupervised statistical models for prepositional phrase attachment. In *Proceedings of the Seventeenth International Conference on Computational Linguistics*, Montreal, 1998.

Adwait Ratnaparkhi, Jeff Reynar, and Salim Roukos. A maximum entropy model for prepositional phrase attachment. In *Proceedings of the ARPA Workshop on Human Language Technology*, pages 250–255, Plainsboro, NJ, 1994.

Timothy R. C. Read and Noel A. C. Cressie. *Goodness-of-fit statistics for discrete multivariate data*. Springer Verlag, New York, 1988.

Philip Resnik. Probabilistic tree-adjoining grammar as a framework for statistical natural language processing. In *COLING 14*, pages x–y, 1992.

Philip Resnik. Selectional constraints: an information-theoretic model and its computational realization. *Cognition*, 61:127–159, 1996.

Philip Resnik and Marti Hearst. Structural ambiguity and conceptual relations. In *Proceedings of the Workshop on Very Large Corpora*, pages 58–64, 1993.

Philip Resnik and David Yarowsky. A perspective on word sense disambiguation methods and their evaluation. In *Proceedings of the SIGLEX workshop Tagging Text with Lexical Semantics*, pages 79–86, Washington DC, 1998.

Philip Stuart Resnik. *Selection and Information: A Class-Based Approach to Lexical Relationships*. PhD thesis, University of Pennsylvania, 1993.

Jeffrey C. Reynar and Adwait Ratnaparkhi. A maximum entropy approach to identifying sentence boundaries. In *Proceedings of the Fifth Conference on Applied Natural Language Processing*, 1997.

Michael D. Riley. Some applications of tree-based modeling to speech and language indexing. In *Proceedings of the DARPA Speech and Natural Language Workshop*, pages 339–352. Morgan Kaufmann, 1989.

Ellen Riloff and Jessica Shepherd. A corpus-based approach for building semantic lexicons. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, pages 117–124, Somerset NJ, 1997. Association for Computational Linguistics.

Eric Sven Ristad. A natural law of succession. Technical Report CS-TR-495-95, Princeton University, 1995.

Eric Sven Ristad. Maximum entropy modeling toolkit. Available from the LANL e-print archive., 1996.

Eric Sven Ristad and Robert G. Thomas. Hierarchical non-emitting markov models. In *Proceedings of ACL 35 and EACL 8*, pages 381–385, Morristown NJ, 1997. Association of Computational Linguistics.

Brian Roark and Eugene Charniak. Noun-phrase co-occurrence statistics for semi-automatic semantic lexicon construction. In *Proceedings of COLING-ACL*, pages 1110–1116, Montreal, Canada, 1998.

S.E. Robertson and K. Sparck Jones. Relevance weighting of search terms. *Journal of the American Society for Information Science*, 27:129–146, 1976.

J. J. Rocchio. Relevance feedback in information retrieval. In Gerard Salton, editor, *The Smart Retrieval System - Experiments in Automatic Document Processing*, pages 313–323. Prentice-Hall, Englewood Cliffs NJ, 1971.

Emmanuel Roche and Yves Schabes. Deterministic part-of-speech tagging with finite-state transducers. *Computational Linguistics*, 21(2):227–253, 1995.

Emmanuel Roche and Yves Schabes. *Finite-State Language Processing*. MIT Press, Boston MA, 1997.

P. M. Roget. *Roget's International Thesaurus*. Thomas Y. Crowell Company, New York, 1946.

Frank Rosenblatt, editor. *Principles of neurodynamics; perceptrons and the theory of brain mechanisms*. Spartan Books, Washington DC, 1962.

Ronald Rosenfeld. *Adaptive Statistical Language Modeling: A Maximum Entropy Approach*. PhD thesis, CMU, 1994. Technical report CMU-CS-94-138.

Ronald Rosenfeld and Xuedong Huang. Improvements in stochastic language modeling. In *Proceedings of the DARPA Speech and Natural Language Workshop*, pages 107–111. Morgan Kaufmann, 1992.

Roni Rosenfeld. A maximum entropy approach to adaptive statistical language modelling. *Computer Speech and Language*, 10:187–228, 1996.

Stanley J. Rosenkrantz and Philip M. Lewis, II. Deterministic left corner parser. In *IEEE Conference Record of the 11th Annual Syposium on Switching and Automata*, pages 139–152, 1970.

Ian C. Ross and John W. Tukey. Introduction to these volumes. In John Wilder Tukey, editor, *Index to Statistics and Probability*, pages iv–x. R & D Press, Los Altos CA, 1975.

D. E. Rumelhart and J. L. McClelland. On learning the past tenses of English verbs. In James L. McClelland, David E. Rumelhart, and the PDP Research Group, editors, *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*, pages 216–271. The MIT Press, Cambridge MA, 1986.

David E. Rumelhart, James L. McClelland, and the PDP research group, editors. *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Volume 1: Foundations*. The MIT Press, Cambridge MA, 1986.

David E. Rumelhart and David Zipser. Feature discovery by competitive learning. *Cognitive Science*, 9:75–112, 1985.

Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Englewood Cliffs, NJ, 1995.

G. Salton, J. Allan, C. Buckley, and A. Singhal. Automatic analysis, theme generation and summarization of machine-readable texts. *Science*, 264:1421–1426, 1994.

Gerard Salton. Experiments in automatic thesaurus construction for information retrieval. In *Proceedings IFIP Congress*, pages 43–49, 1971.

Gerard Salton. *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Addison Wesley, Reading, MA, 1989.

Gerard Salton and James Allen. Selective text utilization and text traversal. In *Proceedings of ACM Hypertext 93*, New York, 1993. Association for Computing Machinery.

Gerard Salton and Chris Buckley. Global text matching for information retrieval. *Science*, 253:1012–1015, 1991.

Gerard Salton, Edward A. Fox, and Harry Wu. Extended boolean information retrieval. *Communications of the ACM*, 26(11):1022–1036, 1983.

Gerard Salton and Michael J. McGill. *Introduction to modern information retrieval*. McGraw-Hill, New York, 1983.

Geoffrey Sampson. How fully does a machine-usable dictionary cover english text? *Literary and Linguistic Computing*, 4:29–35, 1989.

Geoffrey Sampson. *English for the Computer*. Oxford University Press, New York, 1995.

Geoffrey Sampson. *Educating Eve*. Cassell, London; Washington, 1997.

Ken Samuel, Sandra Carberry, and K. Vijay-Shanker. Dialogue act tagging with transformation-based learning. In *Proceedings of COLING-ACL*, pages 1150–1156, Montreal, Canada, 1998.

Christer Samuelsson. Morphological tagging based entirely on bayesian inference. In *9th Nordic Conference on Computational Linguistics*, Stockholm University, Stockholm, Sweden, 1993.

Christer Samuelsson. Handling sparse data by successive abstraction. In *COLING 16*, pages x–y, 1996.

Christer Samuelsson and Atro Voutilainen. Comparing a linguistic and a stochastic tagger. In *Proceedings of ACL 35 and EACL 8*, pages 246–253, Morristown NJ, 1997. Association of Computational Linguistics.

Mark Sanderson and C. J. van Rijsbergen. The impact on retrieval effectiveness of the skewed frequency distribution of a word's senses. *ACM TOIS*, 1998. To appear.

D. Sankoff. Branching processes with terminal types: applications to context-free grammars. *Journal of Applied Probability*, 8:233–240, 1971.

Beatrice Santorini. Part-of-speech tagging guidelines for the Penn treebank project. 3rd Revision, 2nd printing, Feb. 1995. University of Pennsylvania, 1990.

Edward Sapir. *Language: an introduction to the study of speech*. Harcourt Brace and co., New York, 1921.

Satoshi Sato. CTM: An example-based translation aid system. In *Proceedings of COLING*, pages 1259–1263. Association of Computational Linguistics, 1992.

Eric Saund. Unsupervised learning of mixtures of multiple causes in binary data. In J. Cowan, G. Tesauro, and J. Alspector, editors, *Advances in Neural Information Processing Systems 6*. Morgan Kaufmann Publishers, San Mateo CA, 1994.

Yves Schabes. Stochastic lexicalized tree-adjoining grammars. In *COLING 14*, pages x–y, 1992.

Yves Schabes, Anne Abeillé, and Aravind Joshi. Parsing strategies with lexicalized grammars: Tree adjoining grammars. In *COLING 12*, pages 578–583, 1988.

Yves Schabes, Michal Roth, and Randy Osborne. Parsing the Wall Street Journal with the Inside-Outside algorithm. In *EACL 6*, pages 341–347, 1993.

Robert E. Schapire, Yoram Singer, and Amit Singhal. Boosting and rocchio applied to text filtering. In *Proceedings of SIGIR '98*, pages 215–223, Melbourne, 1998.

Helmut Schmid. Probabilistic part-of-speech tagging using decision trees. In *International Conference on New Methods in Language Processing*, pages 44–49, Manchester, England, 1994.

Carson T. Schütze. *he empirical base of linguistics: grammaticality judgments and linguistic methodology*. University of Chicago Press, Chicago, IL, 1996.

Hinrich Schütze. Context space. In Robert Goldman, Peter Norvig, Eugene Charniak, and Bill Gale, editors, *Working Notes of the AAAI Fall Symposium on Probabilistic Approaches to Natural Language*, pages 113–120, Menlo Park CA, 1992a. AAAI Press.

Hinrich Schütze. Dimensions of meaning. In *Proceedings of Supercomputing '92*, pages 787–796, Los Alamitos CA, 1992b. IEEE Computer Society Press.

Hinrich Schütze. Part-of-speech induction from scratch. In *Proceedings of ACL 31*, pages 251–258, Morristown NJ, 1993. Association of Computational Linguistics.

Hinrich Schütze. Distributional part-of-speech tagging. In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics*, pages 141–148, San Francisco CA, 1995. Morgan Kaufmann.

Hinrich Schütze. *Ambiguity Resolution in Language Learning*. CSLI Publications, Stanford CA, 1997.

Hinrich Schütze. Automatic word sense discrimination. *Computational Linguistics*, 24(1):97–124, 1998.

Hinrich Schütze, David A. Hull, and Jan O. Pedersen. A comparison of classifiers and document representations for the routing problem. In *Proceedings of SIGIR 95*, pages 229–237, 1995.

Hinrich Schütze and Jan O. Pedersen. Information retrieval based on word senses. In *Fourth Annual Symposium on Document Analysis and Information Retrieval*, pages 161–175, Las Vegas NV, 1995.

Hinrich Schütze and Jan O. Pedersen. A cooccurrence-based thesaurus and two applications to information retrieval. *Information Processing & Management*, 33 (3):307–318, 1997.

Hinrich Schütze and Yoram Singer. Part-of-speech tagging using a variable memory Markov model. In *Proceedings of the ACL 32*, pages 181–187, Morristown NJ, 1994. Association of Computational Linguistics.

Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.

Claude E. Shannon. Prediction and entropy of printed English. *Bell System Technical Journal*, 30:50–64, 1951.

Jude W. Shavlik and Thomas G. Dietterich, editors. *Readings in Machine Learning*. Morgan Kaufman, San Mateo CA, 1990.

Hadar Shemtov. Text alignment in a tool for translating revised documents. In *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics*, pages 449–453, 1993.

Paraic Sheridan and Alan F. Smeaton. The application of morpho-syntactic language processing to effective phrase matching. *Information Processing & Management*, 28(3):349–370, 1992.

Paraic Sheridan, Martin Wechsler, and Peter Schäuble. Cross language speech retrieval: Establishing a baseline performance. In *SIGIR 97*, pages 99–108, 1997.

Sayori Shimohata, Toshiyuko Sugio, and Junji Nagata. Retrieving collocations by co-occurrences and word order constraints. In *Proceedings of ACL 35 and EACL 8*, pages 476–481, San Francisco CA, 1997. Morgan Kaufmann Publishers.

Sidney Siegel and N. John Castellan, Jr. *Nonparametric statistics for the behavioral sciences*. McGraw Hill, New York, 1988.

K. Sima'an, Rens Bod, S. Krauwer, and Remko Scha. Efficient disambiguation by means of stochastic tree substitution grammars. In *Proceedings International Conference on New Methods in Language Processing*, pages x–y, 1994.

Michel Simard, G. F. Foster, and P. Isabelle. Using cognates to align sentences in bilingual corpora. In *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-92)*, pages 67–81, 1992.

Michel Simard and Pierre Plamondon. Bilingual sentence alignment: Balancing robustness and accuracy. In *Proceedings of the First Conference of the Association for Machine Translation in the Americas (AMTA-96)*, pages 135–144, 1996.

John Sinclair, editor. *Collins COBUILD English dictionary*. Harper Collins, London, 1995. New edition, completely revised.

Amit Singhal, Gerard Salton, and Chris Buckley. Length normalization in degraded text collections. In *Fifth Annual Symposium on Document Analysis and Information Retrieval*, pages 149–162, Las Vegas NV, 1996.

Michael Sipser. *Introduction to the theory of computation*. PWS Publishing Company, Boston MA, 1996.

Jeffrey Mark Siskind. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61:39–91, 1996.

Wojciech Skut and Thorsten Brants. A maximum-entropy partial parser for unrestricted text. In *Sixth Workshop on Very Large Corpora*, Montreal, Canada, 1998a.

Wojciech Skut and Thorsten Brants. A maximum-entropy partial parser for unrestricted text. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 143–151, New Brunswick NJ, 1998b. Association for Computational Linguistics.

Frank Smadja. Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1):143–177, 1993.

Frank Smadja, Kathleen R. McKeown, and Vasileios Hatzivassiloglou. Translating collocations for bilingual lexicons: A statistical approach. *Computational Linguistics*, 22(1):1–38, 1996.

Frank A. Smadja and Kathleen R. McKeown. Automatically extracting and representing collocations for language generation. In *Proceedings of the ACL 28*, pages 252–259, Morristown NJ, 1990. Association of Computational Linguistics.

Alan F. Smeaton. Progress in the application of natural language processing to information retrieval tasks. *The Computer Journal*, 35(3):268–278, 1992.

George Waddel Snedecor and William G. Cochran. *Statistical methods*. Iowa State University Press, Ames, 1989. 8th edition.

Karen Sparck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1):11–21, 1972.

Karen Sparck Jones and Peter Willett, editors. *Readings in Information Retrieval*. Morgan Kaufmann, San Francisco, 1998.

Richard W. Sproat, Chilin Shih, William Gale, and Nancy Chang. A stochastic finite-state word-segmentation algorithm for chinese. *Computational Linguistics*, 22(3): 377–404, 1996.

Richard William Sproat. *Morphology and computation*. MIT Press, Cambridge, MA, 1992.

B. Srinivas, Beth Ann Hockey, Martin Zaidel, and Dania Egedi. Xtag system - a wide coverage grammar for english. In *Proceedings of COLING*, Kyoto, Japan, 1994.

Simon St. Laurent. *XML: A Primer*. MIS Press/IDG Books, Foster City, CA, 1998.

Amy M. Steier and Richard K. Belew. Exporting phrases: A statistical analysis of topical language. In R. Casey and B. Croft, editors, *Second Annual Symposium on Document Analysis and Information Retrieval*, pages 179–190, Las Vegas NV, 1993.

A. Stolcke, E. Shriberg, R. Bates, N. Coccaro, D. Jurafsky, R. Martin, M. Meteer, K. Ries, P. Taylor, and C. Van Ess-Dykema. Dialog act modeling for conversational speech. In *Applying Machine Learning to Discourse Processing*, pages 98–105, Menlo Park CA, 1998. AAAI Press.

Andreas Stolcke. An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational Linguistics*, 21:165–202, 1995.

Andreas Stolcke and Stephen M. Omohundro. Best-first model merging for hidden markov model induction. Technical Report TR-94-003, International Computer Science Institute, Berkeley, California, USA, 1994.

Tomek Strzalkowski. Natural language information retrieval. *Information Processing & Management*, 31(3):397–417, 1995.

Michael Stubbs. *Text and corpus analysis: computer-assisted studies of language and culture*. Blackwell, Oxford, 1996.

Patrick Suppes. Probabilistic grammars for natural languages. *Synthese*, 22:95–116, 1970.

Patrick Suppes, Michael Böttner, and Lin Liang. Machine learning comprehension grammars for ten languages. *Computational Linguistics*, 22(3):329–350, 1996.

Whitney Tabor. *Syntactic Innovation: A Connectionist Model*. PhD thesis, Stanford, 1994.

Jean Tague-Sutcliffe. The pragmatics of information retrieval experimetnation, revisited. *Information Processing & Management*, 28:467–490, 1992. Reprinted in (Sparck Jones and Willett 1998).

Leonard Talmy. Lexicalization patterns: Semantic structure in lexical form. In Timothy Shopen, editor, *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon*, pages 57–149. Cambridge University Press, Cambridge MA, 1985.

Tanenhaus. xx. *yy*, 23:241–268, 1993.

Tesnière. *xx*. yy, zz, 1959.

Masaru Tomita, editor. *Generalized LR parsing*. Kluwer Academic, Boston, 1991.

Geoffrey Towell and Ellen M. Voorhees. Disambiguating highly ambiguous words. *Computational Linguistics*, 24(1):125–146, 1998.

Robert Lawrence Trask. *A dictionary of grammatical terms in linguistics*. Routledge, London; New York, 1993.

Howard Turtle and W. Bruce Croft. Evaluation of an inference network-based retrieval model. *ACM Transactions on Information Systems*, 9(3):187–222, 1991.

Hans van Halteren, Jakub Zavrel, and Walter Daelemans. Improving data driven wordclass tagging by system combination. In *Proceedings of COLING-ACL*, pages 491–497, Montreal, Canada, 1998.

Henk van Riemsdijk and Edwin Williams. *Introduction to the Theory of Grammar*. MIT Press, Cambridge, MA, 1986.

C. J. van Rijsbergen. *Information Retrieval*. Butterworths, London, 1979. Second Edition.

Paola Velardi and Maria Teresa Pazienza. Computer aided interpretation of lexical cooccurrences. In *Proceedings of ACL 27*, pages 185–192, Vancouver, Canada, 1989.

Evelyne Viegas, Boyan Onyshkevych, Victor Raskin, and Sergei Nirenburg. From submit to submitted via submission: On lexical rules in large-scale lexicon acquisition. In *Proceedings of ACL 34*, pages 32–39, 1996.

Stephan Vogel, Hermann Ney, and Christoph Tillmann. Hmm-based word alignment in statistical translation. In *Proceedings of COLING 96*, pages 836–841, Copenhagen, Denmark, 1996.

A. Voutilainen. A syntax-based part of speech analyser. In *Proceedings of EACL 7*, pages 157–164, Dublin, Ireland, 1995. Association for Computational Linguistics.

Donald E. Walker. Knowledge resource tools for accessing large text files. In Sergei Nirenburg, editor, *Machine Translation: Theoretical and methodological issues*, pages 247–261. Cambridge University Press, Cambridge, 1987.

Donald E. Walker and Robert A. Amsler. The use of machine-readable dictionaries in sublanguage analysis. In Ralph Grishman and Richard Kittredge, editors, *Analyzing language in restricted domains: sublanguage description and processing*, pages 69–84. L. Erlbaum Associates, Hillsdale NJ, 1986.

Marilyn A. Walker, Jeanne C. Fromer, and Shrikanth Narayanan. Learning optimal dialogue strategies: A case study of a spoken dialogue agent for email. In *Proceedings of COLING-ACL*, pages 1345–1351, Montreal, Canada, 1998.

Marilyn A. Walker and Johanna D. Moore. Empirical studies in discourse. *Computational Linguistics*, 23(1):1–12, 1997.

Ye-Yi Wang and Alex Waibel. Decoding algorithm in statistical machine translation. In *Proceedings of ACL 35 and EACL 8*, pages 366–372, San Francisco CA, 1997. Morgan Kaufmann Publishers.

Ye-Yi Wang and Alex Waibel. Modeling with structures in statistical machine translation. In *Proceedings of COLING-ACL*, pages 1357–1363, Montreal, Canada, 1998.

Scott A. Waterman. Distinguished usage. In Branimir Boguraev and James Pustejovsky, editors, *Corpus Processing for Lexical Acquisition*, pages 143–172. MIT Press, Cambridge MA, 1995.

Mort Webster and Mitch Marcus. Automatic acquisition of the lexical semantics of verbs from sentence frames. In *Proceedings of ACL 27*, pages 177–184, Vancouver, Canada, 1989.

Sharon L. Weinberg and Kenneth P. Goldberg. *Statistics for the behavioral sciences*. Cambridge University Press, Cambrige; New York, 1990.

Ralph Weischedel, Marie Meteer, Richard Schwartz, Lance Ramshaw, and Jeff Palmucci. Coping with ambiguity and unknown words through probabilistic models. *Computational Linguistics*, 19(2):359–382, 1993.

Erich Wiener, Jan Pedersen, and Andreads Weigend. A neural network approach to topic spotting. In *Proc. SDAIR 95*, pages 317–332, Las Vegas, NV, 1995.

P. Willett and V. Winterman. A comparison of some measures for the determination of inter-molecular structural similarity. *Quant. Struct. Activ. Relat.*, 5:18–25, 1986.

Peter Willett. Recent trends in hierarchic document clustering: A critical review. *Information Processing & Management*, 24(5):577–597, 1988.

Ian H. Witten and Timothy C. Bell. The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression. *IEEE Transactions on Information Theory*, 37:1085–1094, 1991.

S. K. M. Wong and Y. Y. Yao. An information-theoretic measure of term specificity. *Journal of the American Society for Information Science*, 43(1):54–61, 1992.

Henry Bosley Woolf, editor. *Webster's new collegiate dictionary*. G. & C. Merriam Co., Springfield MA, 1973.

Dekai Wu. Aligning a parallel English-Chinese corpus statistically with lexical criteria. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pages 80–87, 1994.

Dekai Wu. Grammarless extraction of phrasal examples from parallel texts. In *Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*, 1995.

Dekai Wu. A polynomial-time algorithm for statistical machine translation. In *Proceedings of ACL 34*, pages 152–158, 1996.

Dekai Wu and Hongsing Wong. Machine translation with a stochastic grammatical channel. In *Proceedings of COLING-ACL*, pages 1408–1415, Montreal, Canada, 1998.

Mikio Yamamoto and Kenneth W. Church. Using suffix arrays to compute term frequency and document frequency for all substrings in a corpus. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 28–37, New Brunswick NJ, 1998. Association for Computational Linguistics.

Yiming Yang. Expert network: Effective and efficient learning from human decisions in text categorization and retrieval. In *SIGIR 94*, pages 13–22, 1994.

Yiming Yang. Noise reduction in a statistical approach to text categorization. In *SIGIR 95*, pages 256–263, 1995.

Yiming Yang. An evaluation of statistical approaches to text categorization. *Information Retrieval*, 1998.

David Yarowsky. Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. In *Proceedings of Coling-92*, pages 454–460. Association of Computational Linguistics, 1992.

David Yarowsky. Decision lists for lexical ambiguity resolution: Application to accent restoration in Spanish and French. In *Proceedings of ACL 32*, Morristown NJ, 1994. Association of Computational Linguistics.

David Yarowsky. Unsupervised word sense disambiguation rivaling supervised methods. In *Proceedings of ACL 33*, 1995.

Gilbert Youmans. A new tool for discourse analysis: The vocabulary-management profile. *Language*, 67(4):763–789, 1991.

J. Zavrel, W. Daelemans, and J. Veenstra. Resolving pp attachment ambiguities with memory-based learning. In *Proceedings of the Conference on Natural Language Learning (CoNLL97)*, pages x–y, Madrid, 1997a.

Jakub Zavrel and Walter Daelemans. Memory-based learning: Using similarity for smoothing. In *Proceedings of ACL 35 and EACL 8*, pages 436–443, San Francisco CA, 1997. Morgan Kaufmann Publishers.

Jakub Zavrel, Walter Daelemans, and Jorn Veenstra. Resolving pp attachment ambiguities with memory-based learning. In *Proceedings of the Workshop on Computational Natural Language Learning*, pages 136–144, Somerset NJ, 1997b. Association for Computational Linguistics.

Uri Zernik. Introduction. In *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 1–26. Lawrence Erlbaum, Hillsdale NJ, 1991a.

Uri Zernik. Train1 vs. train2: Tagging word sense in corpus. In Uri Zernik, editor, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pages 91–112. Lawrence Erlbaum, Hillsdale NJ, 1991b.

George Kingsley Zipf. Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology*, 40:1–95, 1929.

George Kingsley Zipf. *The Psycho-Biology of Language*. Houghton Mifflin, Boston, MA, 1935.

George Kingsley Zipf. *Human Behavior and the Principle of Least Effort*. Addison-Wesley, Cambridge, MA, 1949.