Bibliography

Abbreviations and symbols:

AAAI	Proceedings of the National Conference on Artificial Intelligence
ACL	Proceedings of the Annual Conference of the Association for Computational Linguistics
CLS	Papers from the Annual Regional Meeting of the Chicago Linguistics Society
COGSCI	Proceedings of the Annual Conference of the Cognitive Science Society
COLINGX	Proceedings of the International Conference on Computational Linguistics
EMNLP	Proceedings of the Conference on Empirical Methods in Natural Language Processing
EUROSPEECH	Proceedings of the European Conference on Speech Communication and Technology
INTERSPEECH	Proceedings of the Annual INTERSPEECH Conference
ICPhS	Proceedings of the International Congress of Phonetic Sciences
ICSLP	Proceedings of the International Conference on Spoken Language Processing
IEEE ICASSP	Proceedings of the IEEE International Conference on Acoustics, Speech, & Signal Processing
NAACL-HLT	Proceedings of the North American Chapter of the ACL/Human Language Technology Conference
IJCAI	Proceedings of the International Joint Conference on Artificial Intelligence

Abney, S. (1996). Partial parsing via finite-state cascades. *Natural Language Engineering*, *2*.

Abney, S. P. (1991). Parsing by chunks. In Berwick, R. C., Abney, S. P., and Tenny, C. (Eds.), *Principle-Based Parsing: Computation and Psycholinguistics*, pp. 257–278. Kluwer, Dordrecht.

Abney, S. P. (1997). Stochastic attribute-value grammars. *Computational Linguistics*, *23*(4), 597–618.

Abney, S. P. (2002). Bootstrapping. In *ACL-02*.

Abney, S. P. (2004). Understanding the Yarowsky algorithm. *Computational Linguistics*, *30*(3), 365–395.

Abney, S. P., McAllester, D. A., and Pereira, F. C. N. (1999a). Relating probabilistic grammars and automata. In *ACL-99*.

Abney, S. P., Schapire, R. E., and Singer, Y. (1999b). Boosting applied to tagging and PP attachment. In *EMNLP/VLC-99*, College Park, MD, pp. 38–45.

Ades, A. E. and Steedman, M. (1982). On the order of words. *Linguistics and Philosophy*, *4*, 517–558.

Adjukiewicz, K. (1935). Die syntaktische Konnexität. *Studia Philosophica*, *1*, 1–27. English translation "Syntactic Connexion" by H. Weber in McCall, S. (Ed.) *Polish Logic*, pp. 207–231, Oxford University Press, Oxford, 1967.

Adriaans, P. and van Zaanen, M. (2004). Computational grammar induction for linguists. *Grammars; special issue with the theme "Grammar Induction"*, *7*, 57–68.

Agichtein, E. and Gravano, L. (2000). Snowball: Extracting relations from large plain-text collections. In *Proceedings of the 5th ACM International Conference on Digital Libraries*.

Agirre, E. and de Lacalle, O. L. (2003). Clustering WordNet word senses. In *RANLP 2003*.

Agirre, E. and Edmonds, P. (Eds.). (2006). *Word Sense Disambiguation: Algorithms and Applications*. Kluwer.

Ahn, D., Adafre, S. F., and de Rijke, M. (2005). Extracting temporal information from open domain text: A comparative exploration. In *Proceedings of the 5th Dutch-Belgian Information Retrieval Workshop (DIR'05)*.

Aho, A. V., Sethi, R., and Ullman, J. D. (1986). *Compilers: Principles, Techniques, and Tools*. Addison-Wesley, Reading, MA.

Aho, A. V. and Ullman, J. D. (1972). *The Theory of Parsing, Translation, and Compiling*, Vol. 1. Prentice-Hall.

Aït-Kaci, H. (1984). *A Lattice-Theoretic Approach to Computation Based on a Calculus of Partially Ordered Types*. Ph.D. thesis, University of Pennsylvania.

Aist, G., Dowding, J., Hockey, B. A., and Hieronymus, J. L. (2002). An intelligent procedure assistant for astronaut training and support. In *ACL-02*, Philadelphia, PA.

Albright, A. (2007). How many grammars am i holding up? Discovering phonological differences between word classes. In *WCCFL 26*, pp. 34–42.

Albright, A. and Hayes, B. (2003). Rules vs. analogy in english past tenses: A computational/experimental study. *Cognition*, *90*, 119–161.

Alderete, J., Brasoveanu, A., Merchant, N., Prince, A., and Tesar, B. (2005). Contrast analysis aids in the learning of phonological underlying forms. In *WCCFL 24*, pp. 34–42.

Algoet, P. H. and Cover, T. M. (1988). A sandwich proof of the Shannon-McMillan-Breiman theorem. *The Annals of Probability*, *16*(2), 899–909.

Allen, J. (1984). Towards a general theory of action and time. *Artificial Intelligence*, *23*(2), 123–154.

Allen, J. (1995). *Natural Language Understanding*. Benjamin Cummings, Menlo Park, CA. 2nd ed.

Allen, J. and Core, M. (1997). Draft of DAMSL: Dialog act markup in several layers. Unpublished manuscript.

Allen, J., Ferguson, G., and Stent, A. (2001). An architecture for more realistic conversational systems. In *IUI '01: Proceedings of the 6th international conference on Intelligent user interfaces*, Santa Fe, NM, pp. 1–8. ACM Press.

Allen, J. and Perrault, C. R. (1980). Analyzing intention in utterances. *Artificial Intelligence*, *15*, 143–178.

Allen, J., Hunnicut, M. S., and Klatt, D. H. (1987). *From Text to Speech: The MITalk system*. Cambridge University Press.

Allwood, J. (1995). An activity-based approach to pragmatics. *Gothenburg Papers in Theoretical Linguistics*, *76*.

Allwood, J., Nivre, J., and Ahlsén, E. (1992). On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, *9*, 1–26.

Alshawi, H. (1987). *Memory and Context for Language Interpretation*. Cambridge University Press.

923

Alshawi, H. (Ed.). (1992). *The Core Language Engine*. MIT Press.

Althaus, E., Karamanis, N., and Koller, A. (2004). Computing locally coherent discourses. In *ACL-04*.

Amsler, R. A. (1980). *The Structure of the Merriam-Webster Pocket Dictionary*. Ph.D. thesis, University of Texas, Austin, Texas.

Amsler, R. A. (1981). A taxonomy of English nouns and verbs. In *ACL-81*, Stanford, CA, pp. 133–138. ACL.

Anderson, M. J., Pierrehumbert, J. B., and Liberman, M. Y. (1984). Improving intonational phrasing with syntactic information. In *IEEE ICASSP-84*, pp. 2.8.1–2.8.4.

Anderson, S. R. (1985). *Phonology in the Twentieth Century*. Cambridge University Press.

Antworth, E. L. (1990). *PC-KIMMO: A Two-level Processor for Morphological Analysis*. Summer Institute of Linguistics, Dallas, TX.

Aone, C. and Bennett, S. W. (1995). Evaluating automated and manual acquisition of anaphora resolution strategies. In *ACL-95*, Cambridge, MA, pp. 122–129.

Appelt, D. E., Hobbs, J. R., Bear, J., Israel, D., Kameyama, M., Kehler, A., Martin, D., Myers, K., and Tyson, M. (1995). SRI International FASTUS system MUC-6 test results and analysis. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, San Francisco, pp. 237–248. Morgan Kaufmann.

Appelt, D. E. and Israel, D. (1997). ANLP-97 tutorial: Building information extraction systems. Available as www.ai.sri.com/~appelt/ie-tutorial/.

Archangeli, D. (1984). *Underspecification in Yawelmani Phonology and Morphology*. Ph.D. thesis, MIT.

Archangeli, D. (1997). Optimality theory: An introduction to linguistics in the 1990s. In Archangeli, D. and Langendoen, D. T. (Eds.), *Optimality Theory: An Overview*. Blackwell, Oxford.

Arens, Y., Granacki, J., and Parker, A. (1987). Phrasal analysis of long noun sequences. In *ACL-87*, Stanford, CA, pp. 59–64. ACL.

Argamon, S., Dagan, I., and Krymolowski, Y. (1998). A memory-based approach to learning shallow natural language patterns. In *COLING/ACL-98*, Montreal, pp. 67–73.

Ariel, M. (2001). Accessibility theory: An overview. In Sanders, T., Schilperoord, J., and Spooren, W. (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects*, pp. 29–87. Benjamins.

Ariel, M. (1990). *Accessing Noun Phrase Antecedents*. Routledge.

Artiles, J., Gonzalo, J., and Sekine, S. (2007). The SemEval-2007 WePS evaluation: Establishing a benchmark for the web people search task. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, Czech Republic, pp. 64–69.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29.

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. SLAP 50, Dordrecht, Kluwer.

Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Cambridge University Press.

Atal, B. S. (1974). Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification. *The Journal of the Acoustical Society of America*, *55*(6), 1304–1312.

Atal, B. S. and Hanauer, S. (1971). Speech analysis and synthesis by prediction of the speech wave. *Journal of the Acoustical Society of America*, *50*, 637–655.

Atkins, S. (1993). Tools for computer-aided corpus lexicography: The Hector project. *Acta Linguistica Hungarica*, *41*, 5–72.

Atkinson, M. and Drew, P. (1979). *Order in Court*. Macmillan, London.

Attar, R. and Fraenkel, A. S. (1977). Local feedback in full-text retrieval systems. *Journal of the ACM*, *24*(3), 398–417.

Aubert, X. and Ney, H. (1995). Large vocabulary continuous speech recognition using word graphs. In *IEEE ICASSP*, Vol. 1, pp. 49–52.

Austin, J. L. (1962). *How to Do Things with Words*. Harvard University Press.

Austin, S., Schwartz, R., and Placeway, P. (1991). The forward-backward search algorithm. In *IEEE ICASSP-91*, Vol. 1, pp. 697–700.

Aylett, M. P. (1999). Stochastic suprasegmentals - relationships between redundancy, prosodic structure and syllable duration. In *ICPhS-99*, San Francisco, CA.

Aylett, M. P. and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31–56.

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database (Release 2) [CD-ROM]*. Linguistic Data Consortium, University of Pennsylvania [Distributor], Philadelphia, PA.

Baayen, R. H., Lieber, R., and Schreuder, R. (1997). The morphological complexity of simplex nouns. *Linguistics*, *35*(5), 861–877.

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

Babyonyshev, M. and Gibson, E. (1999). The complexity of nested structures in Japanese. *Language*, *75*(3), 423–450.

Bacchiani, M. and Roark, B. (2003). Unsupervised language model adaptation. In *IEEE ICASSP-03*, pp. 224–227.

Bacchiani, M., Roark, B., and Saraclar, M. (2004). Language model adaptation with MAP estimation and the perceptron algorithm. In *HLT-NAACL-04*, pp. 21–24.

Bach, E. (1976). An extension of classical transformational grammar. In *Problems of Linguistic Metatheory (Proceedings of the 1976 Conference)*. Michigan State University.

Bach, E. (1988). Categorial grammars as theories of language. In Oehrle, R. T., Bach, E., and Wheeler, D. W. (Eds.), *Categorial Grammars and Natural Language Structures*, pp. 17–34. D. Reidel, Dordrecht.

Bachenko, J. and Fitzpatrick, E. (1990). A computational grammar of discourse-neutral prosodic phrasing in English. *Computational Linguistics*, *16*(3), 155–170.

Backus, J. W. (1959). The syntax and semantics of the proposed international algebraic language of the Zurich ACM-GAMM Conference. In *Information Processing: Proceedings of the International Conference on Information Processing, Paris*, pp. 125–132. UNESCO.

Backus, J. W. (1996). Transcript of question and answer session. In Wexelblat, R. L. (Ed.), *History of Programming Languages*, p. 162. Academic Press.

Bacon, F. (1620). *Novum Organum*. Annotated edition edited by Thomas Fowler published by Clarendon Press, Oxford, 1889.

Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. ACM Press, New York.

Bagley, W. C. (1900–1901). The apperception of the spoken sentence: A study in the psychology of language. *The American Journal of Psychology*, *12*, 80–130.

Bahl, L. R., Brown, P. F., de Souza, P. V., and Mercer, R. L. (1986). Maximum mutual information estimation of hidden Markov model parameters for speech recognition. In *IEEE ICASSP-86*, Tokyo, pp. 49–52.

Bahl, L. R., de Souza, P. V., Gopalakrishnan, P. S., Nahamoo, D., and Picheny, M. A. (1992). A fast match for continuous speech recognition using allophonic models. In *IEEE ICASSP-92*, San Francisco, CA, pp. I.17–20.

Bahl, L. R. and Mercer, R. L. (1976). Part of speech assignment by a statistical decision algorithm. In *Proceedings IEEE International Symposium on Information Theory*, pp. 88–89.

Bahl, L. R., Jelinek, F., and Mercer, R. L. (1983). A maximum likelihood approach to continuous speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *5*(2), 179–190.

Bailey, D., Feldman, J., Narayanan, S., and Lakoff, G. (1997). Modeling embodied lexical development. In *COGSCI-97*, Stanford, CA, pp. 19–24. Lawrence Erlbaum.

Bailey, T. and Hahn, U. (2001). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, *44*, 568–591.

Baker, C. F., Fillmore, C. J., and Cronin, B. (2003). The structure of the framenet database. *International Journal of Lexicography*, *16:3*, 281–296.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The Berkeley FrameNet project. In *COLING/ACL-98*, Montreal, Canada, pp. 86–90.

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

Baker, J. K. (1975). The DRAGON system – An overview. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *ASSP-23*(1), 24–29.

Baker, J. K. (1975/1990). Stochastic modeling for automatic speech understanding. In Waibel, A. and Lee, K.-F. (Eds.), *Readings in Speech Recognition*, pp. 297–307. Morgan Kaufmann, Los Altos. Originally appeared in *Speech Recognition*, Academic Press, 1975.

Baldridge, J., Asher, N., and Hunter, J. (2007). Annotation for and robust parsing of discourse structure on unrestricted texts. *Zeitschrift für Sprachwissenschaft*. In press.

Banerjee, S. and Lavie, A. (2005). METEOR: An automatic metric for Mt evaluation with improved correlation with human judgments. In *Proceedings of ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization*.

Banerjee, S. and Pedersen, T. (2003). Extended gloss overlaps as a measure of semantic relatedness. In *IJCAI 2003*, pp. 805–810.

Bangalore, S. and Joshi, A. K. (1999). Supertagging: An approach to almost parsing. *Computational Linguistics*, *25*(2), 237–265.

Banko, M. and Moore, R. C. (2004). A study of unsupervised part-of-speech tagging. In *COLING-04*.

Bar-Hillel, Y. (1953). A quasi-arithmetical notation for syntactic description. *Language*, *29*, 47–58. Reprinted in Y. Bar-Hillel. (1964). *Language and Information: Selected Essays on their Theory and Application*, Addison-Wesley 1964, 61–74.

Bar-Hillel, Y. (1960). The present status of automatic translation of languages. In Alt, F. (Ed.), *Advances in Computers 1*, pp. 91–163. Academic Press.

Bar-Hillel, Y., Perles, M., and Shamir, E. (1961). On formal properties of simple phrase structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, *14*, 143–172. Reprinted in Y. Bar-Hillel. (1964). *Language and Information: Selected Essays on their Theory and Application*, Addison-Wesley 1964, 116–150.

Baroni, M., Matiasek, J., and Trost, H. (2002). Unsupervised discovery of morphologically related words based on orthographic and semantic similarity. In *Proceedings of ACL SIGPHON*, Philadelphia, PA.

Barton, Jr., G. E., Berwick, R. C., and Ristad, E. S. (1987). *Computational Complexity and Natural Language*. MIT Press.

Barzilay, R. and Elhadad, M. (1997). Using lexical chains for text summarization. In *Proceedings of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 10–17.

Barzilay, R., Elhadad, N., and McKeown, K. R. (2002). Inferring strategies for sentence ordering in multidocument news summarization. *Journal of Artificial Intelligence Research*, *17*, 35–55.

Barzilay, R. and Lapata, M. (2005). Modeling local coherence: an entity-based approach. In *ACL-05*, pp. 141–148.

Barzilay, R. and Lapata, M. (2007). Modeling local coherence: an entity-based approach. *Computational Linguistics*. To appear.

Barzilay, R. and Lee, L. (2003). Learning to paraphrase: an unsupervised approach using multiple-sequence alignment. In *HLT-NAACL-03*, pp. 16–23.

Barzilay, R. and Lee, L. (2004). Catching the drift: Probabilistic content models, with applications to generation and summarization. In *HLT-NAACL-04*, pp. 113–120.

Barzilay, R. and McKeown, K. R. (2005). Sentence fusion for multidocument news summarization. *Computational Linguistics*, *31*(3), 297–328.

Bates, R. (1997). The corrections officer: Can John Kidd save Ulysses. *Lingua Franca*, *7*(8). October.

Bauer, L. (1983). *English word-formation*. Cambridge University Press.

Baum, L. F. (1900). *The Wizard of Oz*. Available at Project Gutenberg.

Baum, L. E. (1972). An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes. In Shisha, O. (Ed.), *Inequalities III: Proceedings of the Third Symposium on Inequalities*, University of California, Los Angeles, pp. 1–8. Academic Press.

Baum, L. E. and Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, *73*(3), 360–363.

Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite-state Markov chains. *Annals of Mathematical Statistics*, *37*(6), 1554–1563.

Baumgartner, Jr., W. A., Cohen, K. B., Fox, L., Acquaah-Mensah, G. K., and Hunter, L. (2007). Manual curation is not sufficient for annotation of genomic databases. *Bioinformatics*, *23*, i41–i48.

Baumgartner, Jr., W. A., Lu, Z., Johnson, H. L., Caporaso, J. G., Paquette, J., Lindemann, A., White, E. K., Medvedeva, O., Cohen, K. B., and Hunter, L. (2006). An integrated approach to concept recognition in biomedical text. In *Proceedings of BioCreative 2006*.

Bayes, T. (1763). *An Essay Toward Solving a Problem in the Doctrine of Chances*, Vol. 53. Reprinted in *Facsimiles of two papers by Bayes*, Hafner Publishing Company, New York, 1963.

Bazell, C. E. (1952/1966). The correspondence fallacy in structural linguistics. In Hamp, E. P., Householder, F. W., and Austerlitz, R. (Eds.), *Studies by Members of the English Department, Istanbul University (3), reprinted in Readings in Linguistics II (1966)*, pp. 271–298. University of Chicago Press, Chicago.

Bean, D. and Riloff, E. (1999). Corpus-based identification of non-anaphoric noun phrases. In *ACL-99*, pp. 373–380.

Bean, D. and Riloff, E. (2004). Unsupervised learning of contextual role knowledge for coreference resolution. In *HLT-NAACL-04*.

Becker (1975). The phrasal lexicon. In Schank, R. and Nash-Webber, B. L. (Eds.), *Theoretical Issues in Natural Language Processing*. ACL, Cambridge, MA.

Beckman, M. E. and Ayers, G. M. (1997). Guidelines for ToBI labelling..

Beckman, M. E. and Hirschberg, J. (1994). The ToBI annotation conventions. Manuscript, Ohio State University.

Beckman, M. E. and Pierrehumbert, J. B. (1986). Intonational structure in English and Japanese. *Phonology Yearbook*, *3*, 255–310.

Beeferman, D., Berger, A., and Lafferty, J. D. (1999). Statistical Models for Text Segmentation. *Machine Learning*, *34*(1), 177–210.

Beesley, K. R. (1996). Arabic finite-state morphological analysis and generation. In *COLING-96*, Copenhagen, pp. 89–94.

Beesley, K. R. and Karttunen, L. (2000). Finite-state non-concatenative morphotactics. In *Proceedings of ACL SIGPHON*, Luxembourg, pp. 50–59.

Beesley, K. R. and Karttunen, L. (2003). *Finite-State Morphology*. CSLI Publications, Stanford University.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M. L., and Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, *113*(2), 1001–1024.

Bellegarda, J. R. (2000). Exploiting latent semantic information in statistical language modeling. *Proceedings of the IEEE*, *89*(8), 1279–1296.

Bellegarda, J. R. (1999). Speech recognition experiments using multi-span statistical language models. In *IEEE ICASSP-99*, pp. 717–720.

Bellegarda, J. R. (2005). Unsupervised, language-independent grapheme-to-phoneme conversion by latent analogy. *Speech Communication*, *46*(2), 140–152.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press, Princeton, NJ.

Bennett, C. (2005). Large scale evaluation of corpus-based synthesizers: Results and lessons from the blizzard challenge 2005. In *EUROSPEECH-05*.

Benoît, C., Grice, M., and Hazan, V. (1996). The SUS test: A method for the assessment of text-to-speech synthesis intelligibility using Semantically Unpredictable Sentences. *Speech Communication*, *18*(4), 381–392.

Berger, A., Della Pietra, S. A., and Della Pietra, V. J. (1996). A maximum entropy approach to natural language processing. *Computational Linguistics*, *22*(1), 39–71.

Berger, A. and Miller, R. (1998). Just-in-time language modeling. In *IEEE ICASSP-98*, Vol. II, pp. 705–708.

Bergsma, S. and Lin, D. (2006). Bootstrapping path-based pronoun resolution. In *COLING/ACL 2006*, Sydney, Australia, pp. 33–40.

Bestgen, Y. (2006). Improving Text Segmentation Using Latent Semantic Analysis: A Reanalysis of Choi, Wiemer-Hastings, and Moore (2001). *Computational Linguistics*, *32*(1), 5–12.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In Hayes, J. R. (Ed.), *Cognition and the Development of Language*, pp. 279–352. Wiley.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Pearson ESL, Harlow.

Bickel, B. (2003). Referential density in discourse and syntactic typology. *Language*, *79*(2), 708–736.

Bies, A., Ferguson, M., Katz, K., and MacIntyre, R. (1995). Bracketing guidelines for Treebank II style Penn Treebank Project..

Bikel, D. M. (2004). Intricacies of Collins' parsing model. *Computational Linguistics*, *30*(4), 479–511.

Bikel, D. M., Miller, S., Schwartz, R., and Weischedel, R. (1997). Nymble: a high-performance learning name-finder. In *Proceedings of ANLP-97*, pp. 194–201.

Bikel, D. M., Schwartz, R., and Weischedel, R. (1999). An algorithm that learns what's in a name. *Machine Learning*, *34*, 211–231.

Bilmes, J. (1997). A gentle tutorial on the EM algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. Tech. rep. ICSI-TR-97-021, ICSI, Berkeley.

Bilmes, J. (2003). Buried Markov Models: A graphical-modeling approach to automatic speech recognition. *Computer Speech and Language*, *17*(2-3).

Bilmes, J. and Bartels, C. (2005). Graphical model architectures for speech recognition. *IEEE Signal Processing Magazine*, *22*(5), 89–100.

Bird, S. and Ellison, T. M. (1994). One-level phonology: Autosegmental representations and rules as finite automata. *Computational Linguistics*, *20*(1).

Bird, S. and Loper, E. (2004). NLTK: The Natural Language Toolkit. In *Proceedings of the ACL 2004 demonstration session*, Barcelona, Spain, pp. 214–217.

Birnbaum, L. and Selfridge, M. (1981). Conceptual analysis of natural language. In Schank, R. C. and Riesbeck, C. K. (Eds.), *Inside Computer Understanding: Five Programs plus Miniatures*, pp. 318–353. Lawrence Erlbaum.

Bisani, M. and Ney, H. (2002). Investigations on joint-multigram models for grapheme-to-phoneme conversion. In *ICSLP-02*, Vol. 1, pp. 105–108.

Black, A. W. and Taylor, P. (1994). CHATR: a generic speech synthesis system. In *COLING-94*, Kyoto, Vol. II, pp. 983–986.

Black, A. W. and Hunt, A. J. (1996). Generating F0 contours from ToBI labels using linear regression. In *ICSLP-96*, Vol. 3, pp. 1385–1388.

Black, A. W., Lenzo, K., and Pagel, V. (1998). Issues in building general letter to sound rules. In *3rd ESCA Workshop on Speech Synthesis, Jenolan Caves, Australia*.

Black, A. W., Taylor, P., and Caley, R. (1996-1999). The Festival Speech Synthesis System system. Manual and source code available at www.cstr.ed.ac.uk/projects/festival.html.

Black, A. W. and Tokuda, K. (2005). The Blizzard Challenge–2005: Evaluating corpus-based speech synthesis on common datasets. In *EUROSPEECH-05*.

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

Black, E. (1988). An experiment in computational discrimination of English word senses. *IBM Journal of Research and Development*, *32*(2), 185–194.

Black, E., Jelinek, F., Lafferty, J. D., Magerman, D. M., Mercer, R. L., and Roukos, S. (1992). Towards history-based grammars: Using richer models for probabilistic parsing. In *Proceedings DARPA Speech and Natural Language Workshop*, Harriman, NY, pp. 134–139. Morgan Kaufmann.

Blair, C. R. (1960). A program for correcting spelling errors. *Information and Control*, *3*, 60–67.

Blair-Goldensohn, S., McKeown, K. R., and Schlaikjer, A. H. (2004). Answering definitional questions: A hybrid approach. In Maybury, M. T. (Ed.), *New Directions in Question Answering*, pp. 47–58. AAAI Press.

Blaschke, C., Leon, E. A., Krallinger, M., and Valencia, A. (2005). Evaluation of BioCreative assessment of task 2. *BMC Bioinformatics*, *6*(2).

Bledsoe, W. W. and Browning, I. (1959). Pattern recognition and reading by machine. In *1959 Proceedings of the Eastern Joint Computer Conference*, pp. 225–232. Academic.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(5), 993–1022.

Blevins, J. (1995). The handbook of phonological theory. In Goldsmith, J. (Ed.), *The syllable in phonological theory*. Blackwell, Oxford.

Blitzer, J., McDonald, R., and Pereira, F. C. N. (2006). Domain adaptation with structural correspondence learning. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Sydney, Australia.

Bloomfield, L. (1914). *An Introduction to the Study of Language*. Henry Holt and Company, New York.

Bloomfield, L. (1933). *Language*. University of Chicago Press, Chicago.

Blunsom, P. and Baldwin, T. (2006). Multilingual deep lexical acquisition for hpsgs via supertagging. In *EMNLP 2006*.

Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., Thompson, H., and Winograd, T. (1977). GUS, A frame driven dialog system. *Artificial Intelligence*, *8*, 155–173.

Bobrow, D. G. and Winograd, T. (1977). An overview of KRL, a knowledge representation language. *Cognitive Science*, *1*(1), 3–46.

Bobrow, R. J. and Webber, B. L. (1980). Knowledge representation for syntactic/semantic processing. In *AAAI-80*, Stanford, CA, pp. 316–323. Morgan Kaufmann.

Bock, K. and Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.

Bod, R. (1993). Using an annotated corpus as a stochastic grammar. In *EACL-93*, pp. 37–44.

Boersma, P. and Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, *32*, 45–86.

Boersma, P. and Weenink, D. (2005). Praat: doing phonetics by computer (version 4.3.14). [Computer program]. Retrieved May 26, 2005, from http://www.praat.org/.

Boguraev, B. and Briscoe, T. (Eds.). (1989). *Computational Lexicography for Natural Language Processing*. Longman, London.

Bohus, D. and Rudnicky, A. I. (2005). Sorry, I Didn't Catch That! - An Investigation of Non-understanding Errors and Recovery Strategies. In *SIGdial-2005*, Lisbon, Portugal.

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language*, *48*(3), 633–644.

Bolinger, D. (1981). Two kinds of vowels, two kinds of rhythm. Indiana University Linguistics Club.

Booth, T. L. (1969). Probabilistic representation of formal languages. In *IEEE Conference Record of the 1969 Tenth Annual Symposium on Switching and Automata Theory*, pp. 74–81.

Booth, T. L. and Thompson, R. A. (1973). Applying probability measures to abstract languages. *IEEE Transactions on Computers*, *C-22*(5), 442–450.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *COLT '92*, pp. 144–152.

Bourlard, H. and Morgan, N. (1994). *Connectionist Speech Recognition: A Hybrid Approach*. Kluwer Press.

Bouwman, G., Sturm, J., and Boves, L. (1999). Incorporating confidence measures in the Dutch train timetable information system developed in the Arise project. In *IEEE ICASSP-99*, pp. 493–496.

Brachman, R. J. (1979). On the epistemogical status of semantic networks. In Findler, N. V. (Ed.), *Associative Networks: Representation and Use of Knowledge by Computers*, pp. 3–50. Academic Press.

Brachman, R. J. and Levesque, H. J. (Eds.). (1985). *Readings in Knowledge Representation*. Morgan Kaufmann.

Brachman, R. J. and Schmolze, J. G. (1985). An overview of the KL-ONE knowledge representation system. *Cognitive Science*, *9*(2), 171–216.

Branco, A., McEnery, T., and Mitkov, R. (2002). *Anaphora Processing*. John Benjamins.

Brants, T. (2000). TnT: a statistical part-of-speech tagger. In *Proceedings of the sixth conference on Applied natural language processing*, Seattle, WA, pp. 224–231. Morgan Kaufmann.

Brants, T., Chen, F., and Tsochantaridis, I. (2002). Topic-based document segmentation with probabilistic latent semantic analysis. In *CIKM '02: Proceedings of the eleventh international conference on Information and knowledge management*, pp. 211–218.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth & Brooks, Pacific Grove, CA.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes*, *10*, 137–167.

Brennan, S. E., Friedman, M. W., and Pollard, C. (1987). A centering approach to pronouns. In *ACL-87*, Stanford, CA, pp. 155–162.

Brent, M. R. (1999). An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning*, *34*(1–3), 71–105.

Bresnan, J. (Ed.). (1982). *The Mental Representation of Grammatical Relations*. MIT Press.

Bresnan, J. and Kaplan, R. M. (1982). Introduction: Grammars as mental representations of language. In Bresnan, J. (Ed.), *The Mental Representation of Grammatical Relations*. MIT Press.

Brew, C. (1992). Letting the cat out of the bag: generation for shake-and-bake mt. In *COLING-92*, pp. 610–616.

Brill, E. (1995). Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational Linguistics*, *21*(4), 543–566.

Brill, E. (1997). Unsupervised learning of disambiguation rules for part of speech tagging. Unpublished manuscript.

Brill, E., Dumais, S. T., and Banko, M. (2002). An analysis of the AskMSR question-answering system. In *EMNLP 2002*, pp. 257–264.

Brill, E. and Moore, R. C. (2000). An improved error model for noisy channel spelling correction. In *ACL-00*, Hong Kong, pp. 286–293.

Brill, E. and Resnik, P. (1994). A rule-based approach to prepositional phrase attachment disambiguation. In *COLING-94*, Kyoto, pp. 1198–1204.

Brill, E. and Wu, J. (1998). Classifier combination for improved lexical disambiguation. In *COLING/ACL-98*, Montreal, Canada, pp. 191–195.

Brin, S. (1998). Extracting patterns and relations from the World Wide Web. In *Proceedings World Wide Web and Databases International Workshop, Number 1590 in LNCS*, pp. 172–183. Springer.

Briscoe, T. and Carroll, J. (1993). Generalized Probabilistic LR parsing of natural language (corpora) with unification-based grammars. *Computational Linguistics*, *19*(1), 25–59.

Bromberger, S. and Halle, M. (1989). Why phonology is different. *Linguistic Inquiry*, *20*, 51–70.

Brookes, D. M. and Loke, H. P. (1999). Modelling energy flow in the vocal tract with applications to glottal closure and opening detection. In *IEEE ICASSP-99*, pp. 213–216.

Broschart, J. (1997). Why Tongan does it differently. *Linguistic Typology*, *1*, 123–165.

Browman, C. P. and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252.

Browman, C. P. and Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

Browman, C. P. and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*, 201–250.

Browman, C. P. and Goldstein, L. (1995). Dynamics and articulatory phonology. In Port, R. and v. Gelder, T. (Eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*, pp. 175–193. MIT Press.

Brown, J. S. and Burton, R. R. (1975). Multiple representations of knowledge for tutorial reasoning. In Bobrow, D. G. and Collins, A. (Eds.), *Representation and Understanding*, pp. 311–350. Academic Press.

Brown, P. F., Cocke, J., Della Pietra, S. A., Della Pietra, V. J., Jelinek, F., Lai, J. C., and Mercer, R. L. (1995). Method and system for natural language translation. U.S. Patent 5,477,451.

Brown, P. F., Cocke, J., Della Pietra, S. A., Della Pietra, V. J., Jelinek, F., Lafferty, J. D., Mercer, R. L., and Roossin, P. S. (1990). A statistical approach to machine translation. *Computational Linguistics*, *16*(2), 79–85.

Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., Lai, J. C., and Mercer, R. L. (1992). An estimate of an upper bound for the entropy of English. *Computational Linguistics*, *18*(1), 31–40.

Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., and Mercer, R. L. (1993). The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, *19*(2), 263–311.

Brown, P. F., Della Pietra, V. J., de Souza, P. V., Lai, J. C., and Mercer, R. L. (1992). Class-based *n*-gram models of natural language. *Computational Linguistics*, *18*(4), 467–479.

Brown, P. F., Lai, J. C., and Mercer, R. L. (1991). Aligning sentences in parallel corpora. In *ACL-91*, Berkeley, CA, pp. 169–176.

Bruce, R. and Wiebe, J. (1994). Word-sense disambiguation using decomposable models. In *ACL-94*, Las Cruces, NM, pp. 139–145.

Budanitsky, A. and Hirst, G. (2001). Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures. In *Proceedings of the NAACL 2001 Workshop on WordNet and Other Lexical Resources*, Pittsburgh, PA.

Budanitsky, A. and Hirst, G. (2006). Evaluating WordNet-based measures of lexical semantic relatedness. *Computational Linguistics*, *32*(1), 13–47.

Bulut1, M., Narayanan, S. S., and Syrdal, A. K. (2002). Expressive speech synthesis using a concatenative synthesizer. In *ICSLP-02*.

Bulyko, I., Kirchhoff, K., Ostendorf, M., and Goldberg, J. (2005). Error-sensitive response generation in a spoken langugage dialogue system. *Speech Communication*, *45*(3), 271–288.

Bulyko, I. and Ostendorf, M. (2001). Unit selection for speech synthesis using splicing costs with weighted finite state transducers. In *EUROSPEECH-01*, Vol. 2, pp. 987–990.

Bulyko, I., Ostendorf, M., and Stolcke, A. (2003). Getting more mileage from web text sources for conversational speech language modeling using class-dependent mixtures. In *HLT-NAACL-03*, Edmonton, Canada, Vol. 2, pp. 7–9.

Bunescu, R. C. and Mooney, R. J. (2005). A shortest path dependency kernel for relation extraction. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 724–731.

Bunt, H. (1994). Context and dialogue control. *Think*, *3*, 19–31.

Bunt, H. (2000). Dynamic interpretation and dialogue theory, volume 2. In Taylor, M. M., Neel, F., and Bouwhuis, D. G. (Eds.), *The structure of multimodal dialogue*, pp. 139–166. John Benjamins, Amsterdam.

Bybee, J. L. (2000). The phonology of the lexicon: evidence from lexical diffusion. In Barlow, M. and Kemmer, S. (Eds.), *Usage-based Models of Language*, pp. 65–85. CSLI, Stanford.

Byrd, R. J. and Chodorow, M. S. (1985). Using an On-Line dictionary to find rhyming words and pronunciations for unknown words. In *ACL-85*, pp. 277–283.

Cahn, J. E. (1990). The generation of affect in synthesized speech. In *Journal of the American Voice I/O Society*, Vol. 8, pp. 1–19.

Calder, J. (1987). Typed unification for natural language processing. In Kahn, G., MacQueen, D., and Plotkin, G. (Eds.), *Categories, Polymorphism, and Unification*. Centre for Cognitive Science, University of Edinburgh, Edinburgh, Scotland.

Callison-Burch, C., Osborne, M., and Koehn, P. (2006). Re-evaluating the role of BLEU in machine translation research. In *EACL-06*.

Callison-Burch, C., Talbot, D., and Osborne, M. (2005). Statistical marchine translation with word- and sentences-aligned parallel corpora. In *ACL-05*, pp. 176–183.

Cao, X. (1792). *The Story of the Stone.* (Also known as *The Dream of the Red Chamber*). Penguin Classics, London. First published in Chinese in 1792, translated into English by David Hawkes and published by Penguin in 1973.

Caraballo, S. A. (1999). Automatic construction of a hypernym-labeled noun hierarchy from text. In *ACL-99*, College Park, MD.

Caraballo, S. A. (2001). *Automatic Acquisition of a hypernym-labeled noun hierarchy from text*. Ph.D. thesis, Brown University.

Caramazza, A., Grober, E., Garvey, C., and Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behaviour*, *16*, 601–609.

Carberry, S. (1990). *Plan Recognition in Natural Language Dialog*. MIT Press.

Carbonell, J. (1982). Metaphor: An inescapable phenomenon in natural language comprehension. In Lehnert, W. G. and Ringle, M. (Eds.), *Strategies for Natural Language Processing*, pp. 415–434. Lawrence Erlbaum.

Carbonell, J., Cullingford, R. E., and Gershman, A. V. (1981). Steps toward knowledge-based machine translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *3*(4), 376–392.

Carbonell, J. and Goldstein, J. (1998). The use of mmr, diversity-based reranking for reordering documents and producing summaries. In *SIGIR 1998*, pp. 335–336.

Carbonell, J., Mitamura, T., and Nyberg, E. H. (1992). The KANT perspective: A critique of pure transfer (and pure interlingua, pure statistics, ...). In *International Conference on Theoretical and Methodological Issues in Machine Translation*.

Cardie, C. and Wagstaff, K. (1999). Noun phrase coreference as clustering. In *EMNLP/VLC-99*, College Park, MD.

Cardie, C. (1993). A case-based approach to knowledge acquisition for domain specific sentence analysis. In *AAAI-93*, pp. 798–803. AAAI Press.

Cardie, C. (1994). *Domain-Specific Knowledge Acquisition for Conceptual Sentence Analysis*. Ph.D. thesis, University of Massachusetts, Amherst, MA. Available as CMPSCI Technical Report 94-74.

Cardie, C., Daelemans, W., Ndellec, C., and Sang, E. T. K. (Eds.). (2000). *Proceedings of the Fourth Conference on Computational Language Learning*, Lisbon, Portugal.

Carletta, J., Dahlbäck, N., Reithinger, N., and Walker, M. A. (1997a). Standards for dialogue coding in natural language processing. Tech. rep. Report no. 167, Dagstuhl Seminars. Report from Dagstuhl seminar number 9706.

Carletta, J., Isard, A., Isard, S., Kowtko, J. C., Doherty-Sneddon, G., and Anderson, A. H. (1997b). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, *23*(1), 13–32.

Carlson, L. and Marcu, D. (2001). Discourse tagging manual. Tech. rep. ISI-TR-545, ISI.

Carlson, L., Marcu, D., and Okurowski, M. E. (2001). Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of SIGDIAL*.

Carlson, L., Marcu, D., and Okurowski, M. E. (2002). Building a discourse-tagged corpus in the framework of rhetorical structure theory. In van Kuppevelt, J. and Smith, R. (Eds.), *Current Directions in Discourse and Dialogue*. Kluwer.

Carpenter, B. (1991). The generative power of categorial grammars and head-driven phrase structure grammars with lexical rules. *Computational Linguistics*, *17*(3), 301–313.

Carpenter, B. (1992). *The Logic of Typed Feature Structures*. Cambridge University Press.

Carpenter, B. and Penn, G. (1994). The Attribute Logic Engine Users's Guide Version 2.0.1. Tech. rep., Carnegie Mellon University.

Carpuat, M. and Wu, D. (2007). Improving statistical machine translation using word sense disambiguation. In *EMNLP/CoNLL 2007*, Prague, Czech Republic, pp. 61–72.

Carroll, G. and Charniak, E. (1992). Two experiments on learning probabilistic dependency grammars from corpora. Tech. rep. CS-92-16, Brown University.

Carroll, J., Briscoe, T., and Sanfilippo, A. (1998). Parser evaluation: a survey and a new proposal. In *LREC-98*, Granada, Spain, pp. 447–454.

Cedergren, H. J. and Sankoff, D. (1974). Variable rules: performance as a statistical reflection of competence. *Language*, *50*(2), 333–355.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. N. (Ed.), *Subject and Topic*, pp. 25–55. Academic Press.

Chan, Y. S., Ng, H. T., and Chiang, D. (2007). Word sense disambiguation improves statistical machine translation. In *ACL-07*, Prague, Czech Republic, pp. 33–40.

Chandioux, J. (1976). MÉTÉO: un système opérationnel pour la traduction automatique des bulletins météorologiques destinés au grand public. *Meta*, *21*, 127–133.

Chandler, S. (1991). Metaphor comprehension: A connectionist approach to implications for the mental lexicon. *Metaphor and Symbolic Activity*, *6*(4), 227–258.

Chang, N., Gildea, D., and Narayanan, S. (1998). A dynamic model of aspectual composition. In *COGSCI-98*, Madison, WI, pp. 226–231. Lawrence Erlbaum.

Charniak, E. (1993). *Statistical Language Learning*. MIT Press.

Charniak, E. and Johnson, M. (2005). Coarse-to-fine *n*-best parsing and MaxEnt discriminative reranking. In *ACL-05*, Ann Arbor.

Charniak, E. and Shimony, S. E. (1990). Probabilistic semantics for cost based abduction. In Dietterich, T. S. W. (Ed.), *AAAI-90*, pp. 106–111. MIT Press.

Charniak, E. (1997). Statistical parsing with a context-free grammar and word statistics. In *AAAI-97*, Menlo Park, pp. 598–603. AAAI Press.

Charniak, E. (2000). A maximum-entropy-inspired parser. In *NAACL 2000*, Seattle, Washington, pp. 132–139.

Charniak, E. (2001). Immediate-head parsing for language models. In *ACL-01*, Toulouse, France.

Charniak, E. and Goldman, R. (1988). A logic for semantic interpretation. In *ACL-88*, Buffalo, NY.

Charniak, E., Hendrickson, C., Jacobson, N., and Perkowitz, M. (1993). Equations for part-of-speech tagging. In *AAAI-93*, Washington, D.C., pp. 784–789. AAAI Press.

Charniak, E. and McDermott, D. (1985). *Introduction to Artificial Intelligence*. Addison Wesley.

Chelba, C. and Jelinek, F. (2000). Structured language modeling. *Computer Speech and Language*, *14*, 283–332.

Chen, J. N. and Chang, J. S. (1998). Topical clustering of MRD senses based on information retrieval techniques. *Computational Linguistics*, *24*(1), 61–96.

Chen, J. and Rambow, O. (2003). Use of deep linguistic features for the recognition and labeling of semantic arguments. In *EMNLP 2003*, pp. 41–48.

Chen, S. F. (2003). Conditional and joint models for grapheme-to-phoneme conversion. In *EUROSPEECH-03*.

Chen, S. F. and Goodman, J. (1996). An empirical study of smoothing techniques for language modeling. In *ACL-96*, Santa Cruz, CA, pp. 310–318.

Chen, S. F. and Goodman, J. (1998). An empirical study of smoothing techniques for language modeling. Tech. rep. TR-10-98, Computer Science Group, Harvard University.

Chen, S. F. and Goodman, J. (1999). An empirical study of smoothing techniques for language modeling. *Computer Speech and Language*, *13*(359–394).

Chen, S. F. and Rosenfeld, R. (2000). A survey of smoothing techniques for ME models. *IEEE Transactions on Speech and Audio Processing*, *8*(1), 37–50.

Chen, S. F., Seymore, K., and Rosenfeld, R. (1998). Topic adaptation for language modeling using unnormalized exponential models. In *IEEE ICASSP-98*, pp. 681–684. IEEE.

Chi, Z. (1999). Statistical Properties of Probabilistic Context-Free Grammars. *Computational Linguistics*, *25*(1), 131–160.

Chiang, D. (2005). A hierarchical phrase-based model for statistical machine translation. In *ACL-05*, Ann Arbor, MI, pp. 263–270.

Chierchia, G. and McConnell-Ginet, S. (1991). *Meaning and Grammar*. MIT Press.

Chinchor, N., Hirschman, L., and Lewis, D. L. (1993). Evaluating Message Understanding systems: An analysis of the third Message Understanding Conference. *Computational Linguistics*, *19*(3), 409–449.

Chklovski, T. and Mihalcea, R. (2003). Exploiting Agreement and Disagreement of Human Annotators for Word Sense Disambiguation. In *RANLP 2003*.

Chklovski, T. and Mihalcea, R. (2002). Building a sense tagged corpus with open mind word expert. In *ACL-02 Workshop on Word Sense Disambiguation: Recent Successes and Future Directions*, pp. 116–122.

Chklovski, T. and Pantel, P. (2004). Verb ocean: Mining the Web for fine-grained semantic verb relations. In *EMNLP 2004*, pp. 25–26.

Choi, F. Y. Y. (2000). Advances in domain independent linear text segmentation. In *NAACL 2000*, pp. 26–33.

Choi, F. Y. Y., Wiemer-Hastings, P., and Moore, J. (2001). Latent semantic analysis for text segmentation. In *EMNLP 2001*, pp. 109–117.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*(3), 113–124.

Chomsky, N. (1956/1975). *The Logical Structure of Linguistic Theory*. Plenum.

Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.

Chomsky, N. (1959a). On certain formal properties of grammars. *Information and Control*, *2*, 137–167.

Chomsky, N. (1959b). A review of B. F. Skinner's "Verbal Behavior". *Language*, *35*, 26–58.

Chomsky, N. (1963). Formal properties of grammars. In Luce, R. D., Bush, R., and Galanter, E. (Eds.), *Handbook of Mathematical Psychology*, Vol. 2, pp. 323–418. Wiley.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N. (1969). Quine's empirical assumptions. In Davidson, D. and Hintikka, J. (Eds.), *Words and objections. Essays on the work of W. V. Quine*, pp. 53–68. D. Reidel, Dordrecht.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. Harper and Row.

Chomsky, N. and Miller, G. A. (1958). Finite-state languages. *Information and Control*, *1*, 91–112.

Chomsky, N. and Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In Luce, R. D., Bush, R., and Galanter, E. (Eds.), *Handbook of Mathematical Psychology*, Vol. 2, pp. 269–322. Wiley.

Chou, W., Lee, C.-H., and Juang, B. H. (1993). Minimum error rate training based on *n*-best string models. In *IEEE ICASSP-93*, pp. 2.652–655.

Christiansen, M. H., Allen, J., and Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, *13*(2), 221–268.

Chu-Carroll, J. (1998). A statistical model for discourse act recognition in dialogue interactions. In Chu-Carroll, J. and Green, N. (Eds.), *Applying Machine Learning to Discourse Processing. Papers from the 1998 AAAI Spring Symposium*. Tech. rep. SS-98-01, Menlo Park, CA, pp. 12–17. AAAI Press.

Chu-Carroll, J. and Brown, M. K. (1997). Tracking initiative in collaborative dialogue interactions. In *ACL/EACL-97*, Madrid, Spain, pp. 262–270.

Chu-Carroll, J. and Carberry, S. (1998). Collaborative response generation in planning dialogues. *Computational Linguistics*, *24*(3), 355–400.

Chu-Carroll, J. and Carpenter, B. (1999). Vector-based natural language call routing. *Computational Linguistics*, *25*(3), 361–388.

Chung, G. (2004). Developing a flexible spoken dialog system using simulation. In *ACL-04*, Barcelona, Spain.

Church, A. (1940). A formulation of a simple theory of types. *Journal of Symbolic Logic*, *5*, 56–68.

Church, K. W. (1983). *Phrase-Structure Parsing: A Method for Taking Advantage of Allophonic Constraints*. Ph.D. thesis, MIT.

Church, K. W. (1988). A stochastic parts program and noun phrase parser for unrestricted text. In *Second Conference on Applied Natural Language Processing*, pp. 136–143.

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

Church, K. W., Gale, W. A., and Kruskal, J. B. (1991). Appendix A: the Good-Turing theorem. In *Computer Speech and Language* (Church and Gale, 1991), pp. 19–54.

Church, K. W., Hart, T., and Gao, J. (2007). Compressing trigram language models with Golomb coding. In *EMNLP/CoNLL 2007*, pp. 199–207.

Church, K. W. and Patil, R. (1982). Coping with syntactic ambiguity. *American Journal of Computational Linguistics*, *8*(3-4), 139–149.

Church, K. W. (1980). On memory limitations in natural language processing. Master's thesis, MIT. Distributed by the Indiana University Linguistics Club.

Church, K. W. and Hanks, P. (1989). Word association norms, mutual information, and lexicography. In *ACL-89*, Vancouver, B.C., pp. 76–83.

Church, K. W. and Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, *16*(1), 22–29.

Ciaramita, M. and Johnson, M. (2000). Explaining away ambiguity: learning verb selectional preference with Bayesian networks. In *COLING-00*, pp. 187–193.

Ciaramita, M. and Johnson, M. (2003). Supersense tagging of unknown nouns in WordNet. In *EMNLP-2003*, pp. 168–175.

Cieri, C., Miller, D., and Walker, K. (2004). The Fisher Corpus: a Resource for the Next Generations of Speech-to-Text. In *LREC-04*.

Clark, A. (2000). Inducing syntactic categories by context distribution clustering. In *CoNLL-00*.

Clark, A. (2001). The unsupervised induction of stochastic context-free grammars using distributional clustering. In *CoNLL-01*.

Clark, A. (2002). Memory-based learning of morphology with stochastic transducers. In *ACL-02*, Philadelphia, PA, pp. 513–520.

Clark, H. H. and Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*, 73–111.

Clark, H. H. and Sengal, C. J. (1979). In search of referents for nouns and pronouns. *Memory and Cognition*, *7*, 35–41.

Clark, H. H. (1994). Discourse in production. In Gernsbacher, M. A. (Ed.), *Handbook of Psycholinguistics*. Academic Press.

Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Clark, H. H. and Clark, E. V. (1977). *Psychology and Language*. Harcourt Brace Jovanovich.

Clark, H. H. and Marshall, C. (1981). Definite reference and mutual knowledge. In Joshi, A. K., Webber, B. L., and Sag, I. (Eds.), *Elements of discourse understanding*, pp. 10–63. Cambridge.

Clark, H. H. and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259–294.

Clark, H. H. and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Clark, J. and Yallop, C. (1995). *An Introduction to Phonetics and Phonology*. Blackwell, Oxford. 2nd ed.

Clark, S. and Curran, J. R. (2004a). The importance of supertagging for wide-coverage CCG parsing. In *COLING-04*, pp. 282–288.

Clark, S. and Curran, J. R. (2004b). Parsing the WSJ using CCG and Log-Linear Models. In *ACL-04*, pp. 104–111.

Clarke, J. and Lapata, M. (2006). Models for sentence compression: A comparison across domains, training requirements and evaluation measures. In *COLING/ACL 2006*, pp. 377–384.

Clarke, J. and Lapata, M. (2007). Modelling compression with discourse constraints. In *EMNLP/CoNLL 2007*, Prague, pp. 667–677.

Clarkson, P. R. and Rosenfeld, R. (1997). Statistical language modeling using the CMU-Cambridge toolkit. In *EUROSPEECH-97*, Vol. 1, pp. 2707–2710.

CMU (1993). The Carnegie Mellon Pronouncing Dictionary v0.1. Carnegie Mellon University.

Coccaro, N. and Jurafsky, D. (1998). Towards better integration of semantic predictors in statistical language modeling. In *ICSLP-98*, Sydney, Vol. 6, pp. 2403–2406.

Cohen, J., Kamm, T., and Andreou, A. (1995). Vocal tract normalization in speech recognition: compensating for systematic systematic speaker variability. *Journal of the Acoustical Society of America*, 97(5), 3246–3247.

Cohen, K. B. and Hunter, L. (2006). A critical review of PAS-Bio's argument structures for biomedical verbs. *BMC Bioinformatics*, 7(Suppl 3).

Cohen, K. B., Dolbey, A., Mensah, A. G., and Hunter, L. (2002). Contrast and variability in gene names. In *Proceedings of the ACL Workshop on Natural Language Processing in the Biomedical Domain*, pp. 14–20.

Cohen, K. B. and Hunter, L. (2004). Natural language processing and systems biology. In Dubitzky, W. and Azuaje, F. (Eds.), *Artificial Intelligence Methods and Tools for Systems Biology*, pp. 147–174. Springer.

Cohen, M. H. (1989). *Phonological Structures for Speech Recognition*. Ph.D. thesis, University of California, Berkeley.

Cohen, M. H., Giangola, J. P., and Balogh, J. (2004). *Voice User Interface Design*. Addison-Wesley, Boston.

Cohen, P. R. (1995). *Empirical Methods for Artificial Intelligence*. MIT Press.

Cohen, P. R., Johnston, M., McGee, D., Oviatt, S. L., Clow, J., and Smith, I. (1998). The efficiency of multimodal interaction: a case study. In *ICSLP-98*, Sydney, Vol. 2, pp. 249–252.

Cohen, P. R. and Oviatt, S. L. (1994). The role of voice in human-machine communication. In Roe, D. B. and Wilpon, J. G. (Eds.), *Voice Communication Between Humans and Machines*, pp. 34–75. National Academy Press, Washington, D.C.

Cohen, P. R. and Perrault, C. R. (1979). Elements of a plan-based theory of speech acts. *Cognitive Science*, 3(3), 177–212.

Cohen, W. W., Schapire, R. E., and Singer, Y. (1999). Learning to order things. *Journal of Artificial Intelligence Research*, 10, 243–270.

Coker, C., Umeda, N., and Browman, C. P. (1973). Automatic synthesis from ordinary english test. *IEEE Transactions on Audio and Electroacoustics*, 21(3), 293–298.

Colby, K. M., Weber, S., and Hilf, F. D. (1971). Artificial paranoia. *Artificial Intelligence*, 2(1), 1–25.

Cole, J. S. and Kisseberth, C. W. (1995). Restricting multi-level constraint evaluation. Rutgers Optimality Archive ROA-98.

Cole, R. A., Novick, D. G., Vermeulen, P. J. E., Sutton, S., Fanty, M., Wessels, L. F. A., de Villiers, J. H., Schalkwyk, J., Hansen, B., and Burnett, D. (1997). Experiments with a spoken dialogue system for taking the US census. *Speech Communication*, 23, 243–260.

Cole, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception and Psychophysics*, 13, 153–156.

Cole, R. A. and Jakimik, J. (1980). A model of speech perception. In Cole, R. A. (Ed.), *Perception and Production of Fluent Speech*, pp. 133–163. Lawrence Erlbaum.

Cole, R. A., Novick, D. G., Burnett, D., Hansen, B., Sutton, S., and Fanty, M. (1994). Towards automatic collection of the U.S. census. In *IEEE ICASSP-94*, Adelaide, Australia, Vol. I, pp. 93–96. IEEE.

Cole, R. A., Novick, D. G., Fanty, M., Sutton, S., Hansen, B., and Burnett, D. (1993). Rapid prototyping of spoken language systems: The Year 2000 Census Project. In *Proceedings of the International Symposium on Spoken Dialogue*, Waseda University, Tokyo, Japan.

Cole, R. A. and Rudnicky, A. I. (1983). What's new in speech perception? The research and ideas of William Chandler Bagley. *Psychological Review*, 90(1), 94–101.

Coleman, J. (2005). *Introducing Speech and Language Processing*. Cambridge University Press.

Coleman, J. and Pierrehumbert, J. B. (1997). Stochastic phonological grammars and acceptability. In *Proceedings of ACL SIGPHON*.

Collins, M. and Koo, T. (2005). Discriminative reranking for natural language parsing. *Computational Linguistics*, 31(1), 25–69.

Collins, M. (1996). A new statistical parser based on bigram lexical dependencies. In *ACL-96*, Santa Cruz, CA, pp. 184–191.

Collins, M. (1997). Three generative, lexicalised models for statistical parsing. In *ACL/EACL-97*, Madrid, Spain, pp. 16–23.

Collins, M. (1999). *Head-driven Statistical Models for Natural Language Parsing*. Ph.D. thesis, University of Pennsylvania, Philadelphia.

Collins, M. (2000). Discriminative reranking for natural language parsing. In *ICML 2000*, Stanford, CA, pp. 175–182.

Collins, M. (2003). Head-driven statistical models for natural language parsing. *Computational Linguistics*, 29(4), 589–637.

Collins, M., Hajič, J., Ramshaw, L. A., and Tillmann, C. (1999). A statistical parser for Czech. In *ACL-99*, College Park, MA, pp. 505–512.

Colmerauer, A. (1970). Les systèmes-q ou un formalisme pour analyser et synthétiser des phrase sur ordinateur. Internal publication 43, Département d'informatique de l'Université de Montréal.

Colmerauer, A. (1975). Les grammaires de métamorphose GIA. Internal publication, Groupe Intelligence artificielle, Faculté des Sciences de Luminy, Université Aix-Marseille II, France, Nov 1975. English version, Metamorphosis grammars. In L. Bolc, (Ed.), *Natural Language Communication with Computers, Lecture Notes in Computer Science 63*, Springer Verlag, Berlin, 1978, pp. 133–189.

Colmerauer, A. and Roussel, P. (1996). The birth of Prolog. In Bergin Jr., T. J. and Gibson, Jr., R. G. (Eds.), *History of Programming Languages – II*, pp. 331–352. ACM Press/Addison-Wesley, New York.

Comrie, B. (1989). *Language Universals and Linguistic Typology*. Blackwell, Oxford. Second edition.

Conkie, A. and Isard, S. (1996). Optimal coupling of diphones. In van Santen, J. P. H., Sproat, R., Olive, J. P., and Hirschberg, J. (Eds.), *Progress in Speech Synthesis*. Springer.

Connine, C. M. (1990). Effects of sentence context and lexical knowledge in speech processing. In Altmann, G. T. M. (Ed.), *Cognitive Models of Speech Processing*, pp. 281–294. MIT Press.

Connine, C. M. and Clifton, C. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 291–299.

Connolly, D., Burger, J. D., and Day, D. S. (1994). A machine learning approach to anaphoric reference. In *Proceedings of the International Conference on New Methods in Language Processing (NeMLaP)*.

Conroy, J. M., Schlesinger, J. D., and Goldstein, J. (2006). Classy tasked based summarization: Back to basics. In *DUC-06*.

Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, *19*(90), 297–301.

Cooper, F. S., Liberman, A. M., and Borst, J. M. (1951). The Interconversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech. *Proceedings of the National Academy of Sciences*, *37*(5), 318–325.

Copestake, A. (2002). *Implementing Typed Feature Structure Grammars*. CSLI, Stanford, CA.

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

Core, M., Ishizaki, M., Moore, J. D., Nakatani, C., Reithinger, N., Traum, D. R., and Tutiya, S. (1999). The report of the third workshop of the Discourse Resource Initiative, Chiba University and Kazusa Academia Hall. Tech. rep. No.3 CC-TR-99-1, Chiba Corpus Project, Chiba, Japan.

Corston-Oliver, S. H. (1998). Identifying the linguistic correlates of rhetorical relations. In *Workshop on Discourse Relations and Discourse Markers*, pp. 8–14.

Cottrell, G. W. (1985). *A Connectionist Approach to Word Sense Disambiguation*. Ph.D. thesis, University of Rochester, Rochester, NY. Revised version published in the same title by Pitman in 1989.

Cover, T. M. and King, R. C. (1978). A convergent gambling estimate of the entropy of English. *IEEE Transactions on Information Theory*, *24*(4), 413–421.

Cover, T. M. and Thomas, J. A. (1991). *Elements of information theory*. Wiley.

Cowie, J., Guthrie, J. A., and Guthrie, L. M. (1992). Lexical disambiguation using simulated annealing. In *COLING-92*, Nantes, France, pp. 359–365.

Cowper, E. A. (1976). *Constraints on Sentence Complexity: A Model for Syntactic Processing*. Ph.D. thesis, Brown University, Providence, RI.

Crawley, R. A., Stevenson, R. J., and Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, *19*, 245–264.

Crestani, F., Lemas, M., Rijsbergen, C. J. V., and Campbell, I. (1998). "Is This Document Relevant? ... Probably": a survey of probabilistic models in information retrieval. *ACM Computing Surveys*, *30*(4), 528–552.

Croft, W. (1990). *Typology and Universals*. Cambridge University Press.

Croft, W. (1995). Intonation units and grammatical structure. *Linguistics*, *33*, 839–882.

Crouch, C. J. and Yang, B. (1992). Experiments in automatic statistical thesaurus construction. In *SIGIR-92*, Copenhagen, Denmark, pp. 77–88. ACM.

Cruse, D. A. (2004). *Meaning in Language: an Introduction to Semantics and Pragmatics*. Oxford University Press, Oxford. Second edition.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press.

Cuayáhuitl, H., Renals, S., Lemon, O., and Shimodaira, H. (2007). Hierarchical dialogue optimization using semi-Markov decision processes. In *INTERSPEECH-07*.

Culicover, P. W. and Jackendoff, R. (2005). *Simpler Syntax*. Oxford University Press.

Cullingford, R. E. (1981). SAM. In Schank, R. C. and Riesbeck, C. K. (Eds.), *Inside Computer Understanding: Five Programs plus Miniatures*, pp. 75–119. Lawrence Erlbaum.

Culotta, A. and Sorensen, J. (2004). Dependency tree kernels for relation extraction. In *ACL-04*.

Culy, C. (1985). The complexity of the vocabulary of Bambara. *Linguistics and Philosophy*, *8*, 345–351.

Curran, J. R. (2003). *From Distributional to Semantic Similarity*. Ph.D. thesis, University of Edinburgh.

Curran, J. R. and Moens, M. (2002). Improvements in automatic thesaurus extraction. In *Proceedings of the ACL-02 workshop on Unsupervised Lexical Acquisition*, Philadelphia, PA, pp. 59–66.

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, *29*, 201–219.

Cutler, A. and Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.

Cutler, A. and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 113–121.

Cutting, D., Kupiec, J., Pedersen, J., and Sibun, P. (1992a). A practical part-of-speech tagger. In *Third Conference on Applied Natural Language Processing*, pp. 133–140.

Cutting, D., Karger, D. R., Pedersen, J., and Tukey, J. W. (1992b). Scatter/gather: A cluster-based approach to browsing large document collections. In *SIGIR-92*, Copenhagen, Denmark, pp. 318–329. ACM.

Daelemans, W. and van den Bosch, A. (1997). Language-independent data-oriented grapheme-to-phoneme conversion. In van Santen, J. P. H., Sproat, R., Olive, J. P., and Hirschberg, J. (Eds.), *Progress in Speech Synthesis*, pp. 77–89. Springer.

Daelemans, W., Zavrel, J., Berck, P., and Gillis, S. (1996). MBT: A memory based part of speech tagger-generator. In Ejerhed, E. and Dagan, I. (Eds.), *Proceedings of the Fourth Workshop on Very Large Corpora*, pp. 14–27.

Daelemans, W., Smedt, K. D., and Gazdar, G. (1992). Inheritance in natural language processing. *Computational Linguistics*, *18*(2), 205–218.

Dagan, I. (2000). Contextual word similarity. In Dale, R., Moisl, H., and Somers, H. (Eds.), *A Handbook of Natural Language Processing: Techniques and applications for the processing of language as text*. Marcel Dekker.

Dagan, I., Lee, L., and Pereira, F. C. N. (1999). Similarity-based models of cooccurrence probabilities. *Machine Learning*, *34*(1–3), 43–69.

Dagan, I., Marcus, S., and Markovitch, S. (1993). Contextual word similarity and estimation from sparse data. In *ACL-93*, Columbus, Ohio, pp. 164–171.

Dagan, I., Pereira, F. C. N., and Lee, L. (1994). Similarity-base estimation of word cooccurrence probabilities. In *ACL-94*, Las Cruces, NM, pp. 272–278.

Daly, N. A. and Zue, V. W. (1992). Statistical and linguistic analyses of F_0 in read and spontaneous speech. In *ICSLP-92*, Vol. 1, pp. 763–766.

Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, *7*(3), 171–176.

Damerau, F. J. and Mays, E. (1989). An examination of undetected typing errors. *Information Processing and Management*, *25*(6), 659–664.

Damper, R. I., Marchand, Y., Adamson, M. J., and Gustafson, K. (1999). Evaluating the pronunciation component of text-to-speech systems for english: A performance comparison of different approaches. *Computer Speech and Language*, *13*(2), 155–176.

Dang, H. T. (2006). Overview of DUC 2006. In *DUC-06*.

Danieli, M. and Gerbino, E. (1995). Metrics for evaluating dialogue strategies in a spoken language system. In *Proceedings of the 1995 AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, Stanford, CA, pp. 34–39. AAAI Press, Menlo Park, CA.

Darroch, J. N. and Ratcliff, D. (1972). Generalized iterative scaling for log-linear models. *The Annals of Mathematical Statistics*, *43*(5), 1470–1480.

Daumé III, H. and Marcu, D. (2002). A noisy-channel model for document compression. In *ACL-02*.

Daumé III, H. and Marcu, D. (2005). Induction of word and phrase alignments for automatic document summarization. *Computational Linguistics*, *31*(4), 505–530.

Daumé III, H. and Marcu, D. (2006). Bayesian query-focused summarization. In *COLING/ACL 2006*, Sydney, Australia.

David, Jr., E. E. and Selfridge, O. G. (1962). Eyes and ears for computers. *Proceedings of the IRE (Institute of Radio Engineers)*, *50*, 1093–1101.

Davidson, D. (1967). The logical form of action sentences. In Rescher, N. (Ed.), *The Logic of Decision and Action*. University of Pittsburgh Press.

Davis, E. (1990). *Representations of Commonsense Knowledge*. Morgan Kaufmann.

Davis, K. H., Biddulph, R., and Balashek, S. (1952). Automatic recognition of spoken digits. *Journal of the Acoustical Society of America*, *24*(6), 637–642.

Davis, S. and Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *28*(4), 357–366.

De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, M., and Baayen, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological, and central orthographic effects. *Brain and Language*, *81*, 555–567.

de Marcken, C. (1996). *Unsupervised Language Acquisition*. Ph.D. thesis, MIT.

de Tocqueville, A. (1840). *Democracy in America*. Doubleday, New York. The 1966 translation by George Lawrence.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society of Information Science*, *41*, 391–407.

Dejean, H. and Tjong Kim Sang, E. F. (2001). Introduction to the CoNLL-2001 shared task: Clause identification. In *Proceedings of CoNLL-2001*.

DeJong, G. F. (1982a). An overview of the FRUMP system. In Lehnert, W. G. and Ringle, M. H. (Eds.), *Strategies for Natural Language Processing*, pp. 149–176. Lawrence Erlbaum.

DeJong, G. F. (1982b). An overview of the FRUMP system. In Lehnert, W. G. and Ringle, M. H. (Eds.), *Strategies for Natural Language Processing*, pp. 149–176. Lawrence Erlbaum, New Jersey.

DeJong, G. F. and Waltz, D. L. (1983). Understanding novel language. *Computers and Mathematics with Applications*, *9*.

Delgado, R. L.-C. and Araki, M. (2005). *Spoken, Multilingual, and Multimodal Dialogue Systems*. Wiley.

Deligne, S., Yvon, F., and Bimbot, F. (1995). Variable-length sequence matching for phonetic transcription using joint multigrams. In *EUROSPEECH-95*, Madrid.

Della Pietra, S. A., Della Pietra, V. J., and Lafferty, J. D. (1997). Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(4), 380–393.

Demberg, V. (2006). Letter-to-phoneme conversion for a german text-to-speech system. Diplomarbeit Nr. 47, Universität Stuttgart.

Demetriou, G., Atwell, E., and Souter, C. (1997). Large-scale lexical semantics for speech recognition support. In *EUROSPEECH-97*, pp. 2755–2758.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the *EM* algorithm. *Journal of the Royal Statistical Society*, *39*(1), 1–21.

Denes, P. (1959). The design and operation of the mechanical speech recognizer at University College London. *Journal of the British Institution of Radio Engineers*, *19*(4), 219–234. Appears together with companion paper (Fry 1959).

Deng, L., Lennig, M., Seitz, F., and Mermelstein, P. (1990). Large vocabulary word recognition using context-dependent allophonic hidden Markov models. *Computer Speech and Language*, *4*, 345–357.

Deng, L. and Huang, X. (2004). Challenges in adopting speech recognition. *Communications of the ACM*, *47*(1), 69–75.

Deng, Y. and Byrne, W. (2005). HMM word and phrase alignment for statistical machine translation. In *HLT-EMNLP-05*.

Dermatas, E. and Kokkinakis, G. (1995). Automatic stochastic tagging of natural language texts. *Computational Linguistics*, *21*(2), 137–164.

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

Di Eugenio, B. (1990). Centering theory and the Italian pronominal system. In *COLING-90*, Helsinki, pp. 270–275.

Di Eugenio, B. (1996). The discourse functions of Italian subjects: A centering approach. In *COLING-96*, Copenhagen, pp. 352–357.

Diab, M. and Resnik, P. (2002). An unsupervised method for word sense tagging using parallel corpora. In *ACL-02*, pp. 255–262.

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, *10*(7), 1895–1924.

Digilakis, V. (1992). *Segment-based stochastic models of spectral dynamics for continuous speech recognition*. Ph.D. thesis, Boston University.

Dimitrova, L., Ide, N. M., Petkevic, V., Erjavec, T., Kaalep, H. J., and Tufis, D. (1998). Multext-East: parallel and comparable corpora and lexicons for six Central and Eastern European languages. In *COLING/ACL-98*, Montreal, Canada.

Divay, M. and Vitale, A. J. (1997). Algorithms for grapheme-phoneme translation for English and French: Applications for database searches and speech synthesis. *Computational Linguistics*, *23*(4), 495–523.

Dixon, N. and Maxey, H. (1968). Terminal analog synthesis of continuous speech using the diphone method of segment assembly. *IEEE Transactions on Audio and Electroacoustics*, *16*(1), 40–50.

Doddington, G. (2001). Speaker recognition based on idiolectal differences between speakers. In *EUROSPEECH-01*, Budapest, pp. 2521–2524.

Doddington, G. (2002). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *HLT-01*.

Dolan, W. B. (1994). Word sense ambiguation: Clustering related senses. In *COLING-94*, Kyoto, Japan, pp. 712–716. ACL.

Dolan, W. B., Quirk, C., and Brockett, C. (2004). Unsupervised construction of large paraphrase corpora: exploiting massively parallel news sources. In *COLING-04*.

Donovan, R. E. (1996). *Trainable Speech Synthesis*. Ph.D. thesis, Cambridge University Engineering Department.

Donovan, R. E. and Eide, E. M. (1998). The IBM trainable speech synthesis system. In *ICSLP-98*, Sydney.

Donovan, R. E. and Woodland, P. C. (1995). Improvements in an HMM-based speech synthesiser. In *EUROSPEECH-95*, Madrid, Vol. 1, pp. 573–576.

Dorr, B. (1992). The use of lexical semantics in interlingual machine translation. *Journal of Machine Translation*, *7*(3), 135–193.

Dorr, B. (1993). *Machine Translation*. MIT Press.

Dorr, B. (1994). Machine translation divergences: A formal description and proposed solution. *Computational Linguistics*, *20*(4), 597–633.

Dorr, B., Zajic, D., and Schwartz, R. (2003). Hedge trimmer: a parse-and-trim approach to headline generation. In *HLT-NAACL Workshop on Text Summarization*, pp. 1–8.

Dostert, L. (1955). The Georgetown-I.B.M. experiment. In *Machine Translation of Languages: Fourteen Essays*, pp. 124–135. MIT Press.

Doumpiotis, V., Tsakalidis, S., and Byrne, W. (2003a). Discriminative training for segmental minimum bayes-risk decoding. In *IEEE ICASSP-03*.

Doumpiotis, V., Tsakalidis, S., and Byrne, W. (2003b). Lattice segmentation and minimum bayes risk discriminative training. In *EUROSPEECH-03*.

Downing, P. (1977). On the creation and use of English compound nouns. *Language*, *53*(4), 810–842.

Dowty, D. R. (1979). *Word Meaning and Montague Grammar*. D. Reidel, Dordrecht.

Dowty, D. R., Wall, R. E., and Peters, S. (1981). *Introduction to Montague Semantics*. D. Reidel, Dordrecht.

Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., and Paolino, D. (1983). Outline of discourse transcription. In Edwards, J. A. and Lampert, M. D. (Eds.), *Talking Data: Transcription and Coding in Discourse Research*, pp. 45–89. Lawrence Erlbaum.

Duda, R. O., Hart, P. E., and Stork, D. G. (2000). *Pattern Classification*. Wiley-Interscience Publication.

Duda, R. O. and Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. John Wiley and Sons, New York.

Dudík, M. and Schapire, R. E. (2006). Maximum entropy distribution estimation with generalized regularization. In Lugosi, G. and Simon, H. U. (Eds.), *COLT 2006*, Berlin, pp. 123–138. Springer-Verlag.

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

Durbin, R., Eddy, S., Krogh, A., and Mitchison, G. (1998). *Biological sequence analysis*. Cambridge University Press.

Dutoit, T. (1997). *An Introduction to Text to Speech Synthesis*. Kluwer.

Džeroski, S., Erjavec, T., and Zavrel, J. (2000). Morphosyntactic Tagging of Slovene: Evaluating PoS Taggers and Tagsets. In *LREC-00*, Paris, pp. 1099–1104.

Earley, J. (1968). *An Efficient Context-Free Parsing Algorithm*. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.

Earley, J. (1970). An efficient context-free parsing algorithm. *Communications of the ACM*, *6*(8), 451–455. Reprinted in Grosz et al. (1986).

Echihabi, A., Hermjakob, U., Hovy, E. H., Marcu, D., Melz, E., and Ravichandran, D. (2005). How to select an answer string?. In Strzalkowski, T. and Harabagiu, S. (Eds.), *Advances in Textual Question Answering*. Kluwer.

Edmunson, H. (1969). New methods in automatic extracting. *Journal of the ACM*, *16*(2), 264–285.

Eide, E. M. and Gish, H. (1996). A parametric approach to vocal tract length normalization. In *IEEE ICASSP-96*, Atlanta, GA, pp. 346–348.

Eide, E. M., Bakis, R., Hamza, W., and Pitrelli, J. F. (2004). Towards synthesizing expressive speech. In Narayanan, S. S. and Alwan, A. (Eds.), *Text to Speech Synthesis: New paradigms and Advances*. Prentice Hall.

Eisner, J. (1996a). An empirical comparison of probability models for dependency grammar. Tech. rep. IRCS-96-11, Institute for Research in Cognitive Science, Univ. of Pennsylvania.

Eisner, J. (1996b). Three new probabilistic models for dependency parsing: An exploration. In *COLING-96*, Copenhagen, pp. 340–345.

Eisner, J. (1997). Efficient generation in primitive optimality theory. In *ACL/EACL-97*, Madrid, Spain, pp. 313–320.

Eisner, J. (2000a). Bilexical grammars and their cubic-time parsing algorithms. In Bunt, H. and Nijholt, A. (Eds.), *Advances in Probabilistic and Other Parsing Technologies*, pp. 29–62. Kluwer.

Eisner, J. (2000b). Directional constraint evaluation in Optimality Theory. In *COLING-00*, Saarbrücken, Germany, pp. 257–263.

Eisner, J. (2002a). An interactive spreadsheet for teaching the forward-backward algorithm. In *Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching NLP and CL*, pp. 10–18.

Eisner, J. (2002b). Comprehension and compilation in Optimality Theory. In *ACL-02*, Philadelphia, pp. 56–63.

Eisner, J. (2003). Learning non-isomorphic tree mappings for machine translation. In *ACL-03*.

Ejerhed, E. I. (1988). Finding clauses in unrestricted text by finitary and stochastic methods. In *Second Conference on Applied Natural Language Processing*, pp. 219–227.

Elhadad, M. (1990). Types in functional unification grammars. In *ACL-90*, Pittsburgh, PA, pp. 157–164.

Ellison, T. M. (1992). *The Machine Learning of Phonological Structure*. Ph.D. thesis, University of Western Australia.

Ellison, T. M. (1994). Phonological derivation in optimality theory. In *COLING-94*, Kyoto, pp. 1007–1013.

Erjavec, T. (2004). Multext-east version 3: Multilingual morphosyntactic specifications, lexicons and corpora. In *LREC-04*, pp. 1535–1538. ELRA.

Erkan, G. and Radev, D. (2004). Lexrank: Graph-based centrality as salience in text summarization. *Journal of Artificial Intelligence Research (JAIR)*, *22*, 457–479.

Etzioni, O., Cafarella, M., Downey, D., Popescu, A.-M., Shaked, T., Soderland, S., Weld, D. S., and Yates, A. (2005). Unsupervised named-entity extraction from the web: An experimental study. *Artificial Intelligence*, *165*(1), 91–134.

Evans, N. (2000). Word classes in the world's languages. In Booij, G., Lehmann, C., and Mugdan, J. (Eds.), *Morphology: a Handbook on Inflection and Word Formation*, pp. 708–732. Mouton, Berlin.

Evans, R. and Gazdar, G. (1996). DATR: A language for lexical knowledge representation. *Computational Linguistics*, *22*(2), 167–216.

Evermann, G. and Woodland, P. C. (2000). Large vocabulary decoding and confidence estimation using word posterior probabilities. In *IEEE ICASSP-00*, Istanbul, Vol. III, pp. 1655–1658.

Fackrell, J. and Skut, W. (2004). Improving pronunciation dictionary coverage of names by modelling spelling variation. In *Proceedings of the 5th Speech Synthesis Workshop*.

Fano, R. M. (1961). *Transmission of information; A statistical theory of communications*. MIT Press.

Fant, C. G. M. (3951). Speech communication research. *Ing. Vetenskaps Akad. Stockholm, Sweden*, *24*, 331–337.

Fant, G. (2004). *Speech Acoustics and Phonetics*. Kluwer.

Fant, G. M. (1960). *Acoustic Theory of Speech Production*. Mouton.

Fant, G. M. (1986). Glottal flow: Models and interaction. *Journal of Phonetics*, *14*, 393–399.

Fant, G. M. (1997). The voice source in connected speech. *Speech Communication*, *22*(2-3), 125–139.

Fass, D. (1988). *Collative Semantics: A Semantics for Natural Language*. Ph.D. thesis, New Mexico State University, Las Cruces, New Mexico. CRL Report No. MCCS-88-118.

Fass, D. (1991). met*: A method for discriminating metaphor and metonymy by computer. *Computational Linguistics*, *17*(1).

Fass, D. (1997). *Processing Metonymy and Metaphor*. Ablex Publishing, Greenwich, CT.

Fass, D., Martin, J. H., and Hinkelman, E. A. (Eds.). (1992). *Computational Intelligence: Special Issue on Non-Literal Language*, Vol. 8. Blackwell, Cambridge, MA.

Fauconnier, G. (1985). *Mental Spaces: Aspects of Meaning Construction in Natural Language*. MIT Press.

Fellbaum, C. (Ed.). (1998). *WordNet: An Electronic Lexical Database*. MIT Press.

Fensel, D., Hendler, J. A., Lieberman, H., and Wahlster, W. (Eds.). (2003). *Spinning the Semantic Web: Bring the World Wide Web to its Full Potential*. MIT Press, Cambridge, MA.

Ferrer, L., Shriberg, E., and Stolcke, A. (2003). A prosody-based approach to end-of-utterance detection that does not require speech recognition. In *IEEE ICASSP-03*.

Ferro, L., Gerber, L., Mani, I., Sundheim, B., and Wilson, G. (2005). Tides 2005 standard for the annotation of temporal expressions. Tech. rep., MITRE.

Fikes, R. E. and Nilsson, N. J. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, *2*, 189–208.

Filippova, K. and Strube, M. (2006). Using linguistically motivated features for paragraph boundary identification. In *EMNLP 2006*.

Fillmore, C. J. (1968). The case for case. In Bach, E. W. and Harms, R. T. (Eds.), *Universals in Linguistic Theory*, pp. 1–88. Holt, Rinehart & Winston.

Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, *VI*(2), 222–254.

Fillmore, C. J., Kay, P., and O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of Let Alone. *Language*, *64*(3), 510–538.

Finin, T. (1980). The semantic interpretation of nominal compounds. In *AAAI-80*, Stanford, CA, pp. 310–312.

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

Fisher, D., Soderland, S., McCarthy, J., Feng, F., and Lehnert, W. G. (1995). Description of the UMass system as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, San Francisco, pp. 127–140. Morgan Kaufmann.

Fisher, W. (1996). tsylb2 software and documentation..

Fitt, S. (2002). Unisyn lexicon. `http://www.cstr.ed.ac.uk/projects/unisyn/`.

Flanagan, J. L. (1972). *Speech Analysis, Synthesis, and Perception*. Springer.

Flanagan, J. L., Ishizaka, K., and Shipley, K. L. (1975). Synthesis of speech from a dynamic model of the vocal cords and vocal tract. *The Bell System Technical Journal*, *54*(3), 485–506.

Fodor, J. A. and Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, *4*, 414–420.

Ford, C., Fox, B., and Thompson, S. A. (1996). Practices in the construction of turns. *Pragmatics*, *6*, 427–454.

Ford, C. and Thompson, S. A. (1996). Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In Ochs, E., Schegloff, E. A., and Thompson, S. A. (Eds.), *Interaction and Grammar*, pp. 134–184. Cambridge University Press.

Ford, M. (1983). A method for obtaining measures of local parsing complexity through sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 203–218.

Forney Jr., G. D. (1973). The Viterbi algorithm. *Proceedings of the IEEE*, *61*(3), 268–278.

Fosler-Lussier, E. (1999). Multi-level decision trees for static and dynamic pronunciation models. In *EUROSPEECH-99*, Budapest.

Fosler-Lussier, E. and Morgan, N. (1999). Effects of speaking rate and word predictability on conversational pronunciations. *Speech Communication*, *29*(2-4), 137–158.

Foster, D. W. (1989). *Elegy by W.S.: A Study in Attribution*. Associated University Presses, Cranbury, NJ.

Foster, D. W. (1996). Primary culprit. *New York*, *29*(8), 50–57. February 26.

Foster, G. (2000). A maximum entropy/minimum divergence translation model. In *ACL-00*, Hong Kong.

Fox, B. and Jasperson, R. (1995). A syntactic exploration of repair in English conversation. In Davis, P. (Ed.), *Descriptive and Theoretical Modes in the Alternative Linguistics*, pp. 77–134. John Benjamins, Amsterdam. In press.

Fox Tree, J. E. and Clark, H. H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition*, *62*, 151–167.

Frakes, W. B. and Baeza-Yates, R. (1992). *Information Retrieval: Data Structures and Algorithms*. Prentice Hall.

Francis, H. S., Gregory, M. L., and Michaelis, L. A. (1999). Are lexical subjects deviant?. In *CLS-99*. University of Chicago.

Francis, W. N. (1979). A tagged corpus – problems and prospects. In Greenbaum, S., Leech, G., and Svartvik, J. (Eds.), *Studies in English linguistics for Randolph Quirk*, pp. 192–209. Longman.

Francis, W. N. and Kučera, H. (1982). *Frequency Analysis of English Usage*. Houghton Mifflin, Boston.

Frank, R. and Satta, G. (1998). Optimality theory and the generative complexity of constraint violability. *Computational Linguistics*, *24*(2), 307–315.

Frankel, J., Wester, M., and King, S. (2007). Articulatory feature recognition using dynamic bayesian networks. *Computer Speech and Language*, *21*(4), 620–640.

Franz, A. and Brants, T. (2006). All our n-gram are belong to you. `http://googleresearch.blogspot.com/2006/08/all-our-n-gram-are-belong-to-you.html`.

Franz, A. (1996). *Automatic Ambiguity Resolution in Natural Language Processing*. Springer-Verlag, Berlin.

Franz, A. (1997). Independence assumptions considered harmful. In *ACL/EACL-97*, Madrid, Spain, pp. 182–189.

Fraser, A. and Marcu, D. (2005). ISI's participation in the Romanian-English alignment task. In *Proceedings of the ACL Workshop on Building and Using Parallel Texts*, pp. 91–94.

Fraser, N. M. (1992). Assessment of interactive systems. In Gibbon, D., Moore, R., and Winski, R. (Eds.), *Handbook on Standards and Resources for Spoken Language Systems*, pp. 564–615. Mouton de Gruyter, Berlin.

Fraser, N. M. and Gilbert, G. N. (1991). Simulating speech systems. *Computer Speech and Language*, *5*, 81–99.

Fraser, N. M. and Hudson, R. A. (1992). Inheritance in word grammar. *Computational Linguistics*, *18*(2), 133–158.

Freitag, D. (1998). Multistrategy learning for information extraction. In *ICML 1998*, Madison, WI, pp. 161–169.

Freitag, D. and McCallum, A. (1999). Information extraction using HMMs and shrinkage. In *Proceedings of the AAAI-99 Workshop on Machine Learning for Information Retrieval*.

Friedl, J. E. F. (1997). *Master Regular Expressions*. O'Reilly.

Frisch, S. A., Large, N. R., and Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, *42*, 481–496.

Fromkin, V. and Ratner, N. B. (1998). Speech production. In Gleason, J. B. and Ratner, N. B. (Eds.), *Psycholinguistics*. Harcourt Brace, Fort Worth, TX.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, *27*, 765–768.

Fry, D. B. (1959). Theoretical aspects of mechanical speech recognition. *Journal of the British Institution of Radio Engineers*, *19*(4), 211–218. Appears together with companion paper (Denes 1959).

Fujisaki, H. and Ohno, S. (1997). Comparison and assessment of models in the study of fundamental frequency contours of speech. In *ESCA workshop on Intonation: Theory Models and Applications*.

Fung, P. and McKeown, K. R. (1997). A technical word and term translation aid using noisy parallel corpora across language groups. *Machine Translation*, *12*(1-2), 53–87.

Gabrilovich, E. and Markovitch, S. (2007). Computing Semantic Relatedness using Wikipedia-based Explicit Semantic Analysis. In *IJCAI-07*.

Gaizauskas, R., Wakao, T., Humphreys, K., Cunningham, H., and Wilks, Y. (1995). University of Sheffield: Description of the LaSIE system as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, San Francisco, pp. 207–220. Morgan Kaufmann.

Gale, W. A. and Church, K. W. (1990). Estimation procedures for language context: poor estimates are worse than none. In *COMPSTAT: Proceedings in Computational Statistics*, pp. 69–74.

Gale, W. A. and Church, K. W. (1991). A program for aligning sentences in bilingual corpora. In *ACL-91*, Berkeley, CA, pp. 177–184.

Gale, W. A. and Church, K. W. (1993). A program for aligning sentences in bilingual corpora. *Computational Linguistics*, *19*, 75–102.

Gale, W. A. and Church, K. W. (1994). What is wrong with adding one?. In Oostdijk, N. and de Haan, P. (Eds.), *Corpus-based Research into Language*, pp. 189–198. Rodopi, Amsterdam.

Gale, W. A., Church, K. W., and Yarowsky, D. (1992a). Work on statistical methods for word sense disambiguation. In Goldman, R. (Ed.), *Proceedings of the 1992 AAAI Fall Symposium on Probabilistic Approaches to Natural Language*.

Gale, W. A., Church, K. W., and Yarowsky, D. (1992b). Estimating upper and lower bounds on the performance of word-sense disambiguation programs. In *ACL-92*, Newark, DE, pp. 249–256.

Gale, W. A., Church, K. W., and Yarowsky, D. (1992c). One sense per discourse. In *Proceedings DARPA Speech and Natural Language Workshop*, pp. 233–237. Morgan Kaufmann.

Gale, W. A., Church, K. W., and Yarowsky, D. (1993). A method for disambiguating word senses in a large corpus. *Computers and the Humanities*, *26*, 415–439.

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

Galescu, L. and Allen, J. (2001). Bi-directional conversion between graphemes and phonemes using a joint N-gram model. In *Proceedings of the 4th ISCA Tutorial and Research Workshop on Speech Synthesis*.

Galley, M., Hopkins, M., Knight, K., and Marcu, D. (2004). What's in a translation rule?. In *HLT-NAACL-04*.

Galley, M. and McKeown, K. R. (2007). Lexicalized Markov grammars for sentence compression. In *NAACL-HLT 07*, Rochester, NY, pp. 180–187.

Garrett, M. F. (1975). The analysis of sentence production. In Bower, G. H. (Ed.), *The Psychology of Learning and Motivation*, Vol. 9. Academic.

Garside, R. (1987). The CLAWS word-tagging system. In Garside, R., Leech, G., and Sampson, G. (Eds.), *The Computational Analysis of English*, pp. 30–41. Longman, London.

Garside, R., Leech, G., and McEnery, A. (1997). *Corpus Annotation*. Longman, London and New York.

Gaussier, E. (1999). Unsupervised learning of derivational morphology from inflectional lexicons. In *ACL-99*.

Gaustad, T. (2001). Statistical corpus-based word sense disambiguation: Pseudowords vs. real ambiguous words. In *ACL/EACL 2001 – Student Research Workshop*, pp. 255–262.

Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry*, *12*(2), 155–184.

Gazdar, G. (1982). Phrase structure grammar. In Jacobson, P. and Pullum, G. K. (Eds.), *The Nature of Syntactic Representation*, pp. 131–186. Reidel, Dordrecht.

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Blackwell, Oxford.

Gazdar, G. and Mellish, C. (1989). *Natural Language Processing in LISP*. Addison Wesley.

Gazdar, G., Pullum, G. K., Carpenter, B., Klein, E., Hukari, T. E., and Levine, R. D. (1988). Category structures. *Computational Linguistics*, *14*(1), 1–19.

Ge, N., Hale, J., and Charniak, E. (1998). A statistical approach to anaphora resolution. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pp. 161–171.

Gee, J. P. and Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*, 411–458.

Geman, S. and Johnson, M. (2002). Dynamic programming for parsing and estimation of stochastic unification-based grammars. In *ACL-02*, pp. 279–286.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Genzel, D. and Charniak, E. (2002). Entropy rate constancy in text. In *ACL-02*.

Genzel, D. and Charniak, E. (2003). Variation of entropy and parse trees of sentences as a function of the sentence number. In *EMNLP 2003*.

Gerdemann, D. and van Noord, G. (2000). Approximation and exactness in finite state optimality theory. In *Proceedings of ACL SIGPHON*.

Germann, U., Jahr, M., Knight, K., Marcu, D., and Yamada, K. (2001). Fast decoding and optimal decoding for machine translation. In *ACL-01*, pp. 228–235.

Gershman, A. V. (1977). Conceptual analysis of noun groups in English. In *IJCAI-77*, Cambridge, MA, pp. 132–138.

Gibbon, D., Mertins, I., and Moore, R. (2000). *Handbook of Multimodal and Spoken Dialogue Systems: Resources, Terminology and Product Evaluation*. Kluwer, Dordrecht.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E. (2003). Sentence comprehension, linguistic complexity in. In Nadel, L. (Ed.), *Encyclopedia of Cognitive Science*, pp. 1137–1141. Nature Publishing Group, New York, NY.

Gil, D. (2000). Syntactic categories, cross-linguistic variation and universal grammar. In Vogel, P. M. and Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*, pp. 173–216. Mouton.

Gildea, D. and Hofmann, T. (1999). Topic-based language models using EM. In *EUROSPEECH-99*, Budapest, pp. 2167–2170.

Gildea, D. and Jurafsky, D. (1996). Learning bias and phonological rule induction. *Computational Linguistics*, *22*(4), 497–530.

Gildea, D. and Jurafsky, D. (2000). Automatic labeling of semantic roles. In *ACL-00*, Hong Kong, pp. 512–520.

Gildea, D. and Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational Linguistics*, *28*(3), 245–288.

Gillick, L. and Cox, S. (1989). Some statistical issues in the comparison of speech recognition algorithms. In *IEEE ICASSP-89*, pp. 532–535.

Girju, R., Badulescu, A., and Moldovan, D. (2006). Automatic discovery of part-whole relations. *Computational Linguistics*, *31*(1).

Girju, R., Badulescu, A., and Moldovan, D. (2003). Learning semantic constraints for the automatic discovery of part-whole relations. In *HLT-NAACL-03*, Edmonton, Canada, pp. 1–8.

Givón, T. (1990). *Syntax: A functional typological introduction*. John Benjamins, Amsterdam.

Glass, J. (2003). A probabilistic framework for segment-based speech recognition. *Computer Speech and Language,*, *17*(1–2), 137–152.

Glennie, A. (1960). On the syntax machine and the construction of a universal compiler. Tech. rep. No. 2, Contr. NR 049-141, Carnegie Mellon University (at the time Carnegie Institute of Technology), Pittsburgh, PA.

Godfrey, J., Holliman, E., and McDaniel, J. (1992). SWITCH-BOARD: Telephone speech corpus for research and development. In *IEEE ICASSP-92*, San Francisco, pp. 517–520.

Gold, B. and Morgan, N. (1999). *Speech and Audio Signal Processing*. Wiley Press.

Golding, A. R. and Roth, D. (1999). A winnow based approach to context-sensitive spelling correction. *Machine Learning*, *34*(1-3), 107–130. Special Issue on Machine Learning and Natural Language.

Golding, A. R. (1997). A bayesian hybrid method for context-sensitive spelling correction. In *Proceedings of the Third Workshop on Very Large Corpora*, Boston, MA, pp. 39–53.

Golding, A. R. and Schabes, Y. (1996). Combining trigram-based and feature-based methods for context-sensitive spelling COrrection. In *ACL-96*, Santa Cruz, CA, pp. 71–78.

Goldsmith, J. (1976). *Autosegmental Phonology*. Ph.D. thesis, MIT.

Goldsmith, J. (1993). Harmonic phonology. In Goldsmith, J. (Ed.), *The Last Phonological Rule*, pp. 21–60. University of Chicago Press, Chicago.

Goldsmith, J. (2001). Unsupervised learning of the morphology of a natural language. *Computational Linguistics*, *27*, 153–198.

Goldstein, J., Mittal, V., Carbonell, J., and Kantrowitz, M. (2000). Multi-document summarization by sentence extraction. In *Proceedings of the ANLP/NAACL Workshop on Automatic Summarization*.

Goldwater, S. and Griffiths, T. L. (2007). A fully Bayesian approach to unsupervised part-of-speech tagging. In *ACL-07*, Prague, Czech Republic.

Goldwater, S., Griffiths, T. L., and Johnson, M. (2006). Contextual dependencies in unsupervised word segmentation. In *COLING/ACL 2006*, Sydney, Australia.

Goldwater, S. and Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In *Stockholm Workshop on Variation within Optimality Theory*, pp. 111–120. Stockholm University Press.

Goldwater, S. and Johnson, M. (2005). Representational bias in unsupervised learning of syllable structure. In *Proceedings of the Conference on Computational Natural Language Learning (CoNLL-2005)*.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, *40*, 16–264.

Good, M. D., Whiteside, J. A., Wixon, D. R., and Jones, S. J. (1984). Building a user-derived interface. *Communications of the ACM*, *27*(10), 1032–1043.

Goodman, J. (1997). Probabilistic feature grammars. In *Proceedings of the International Workshop on Parsing Technology*.

Goodman, J. (2004). Exponential priors for maximum entropy models. In *ACL-04*.

Goodman, J. (2006). A bit of progress in language modeling: Extended version. Tech. rep. MSR-TR-2001-72, Machine Learning and Applied Statistics Group, Microsoft Research, Redmond, WA.

Goodwin, C. (1996). Transparent vision. In Ochs, E., Schegloff, E. A., and Thompson, S. A. (Eds.), *Interaction and Grammar*, pp. 370–404. Cambridge University Press.

Gordon, D. and Lakoff, G. (1971). Conversational postulates. In *CLS-71*, pp. 200–213. University of Chicago. Reprinted in Peter Cole and Jerry L. Morgan (Eds.), *Speech Acts: Syntax and Semantics Volume 3*, Academic, 1975.

Gordon, P. C., Grosz, B. J., and Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*(3), 311–347.

Gorin, A. L., Riccardi, G., and Wright, J. H. (1997). How may i help you?. *Speech Communication*, *23*, 113–127.

Götz, T., Meurers, W. D., and Gerdemann, D. (1997). The Con-Troll manual. Tech. rep., Seminar für Sprachwissenschaft, Universität Tübingen.

Gould, J. D., Conti, J., and Hovanyecz, T. (1983). Composing letters with a simulated listening typewriter. *Communications of the ACM*, *26*(4), 295–308.

Gould, J. D. and Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of the ACM*, *28*(3), 300–311.

Gould, S. J. (1980). *The Panda's Thumb*. Penguin Group, London.

Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realisation in four varieties of British English. *Journal of Phonetics*, *28*, 161–186.

Grabe, E. (2001). The IViE labelling guide..

Graff, D. (1997). The 1996 Broadcast News speech and language-model corpus. In *Proceedings DARPA Speech Recognition Workshop*, Chantilly, VA, pp. 11–14. Morgan Kaufmann.

Gray, R. M. (1984). Vector quantization. *IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-1*(2), 4–29.

Green, B. F., Wolf, A. K., Chomsky, C., and Laughery, K. (1961). Baseball: An automatic question answerer. In *Proceedings of the Western Joint Computer Conference 19*, pp. 219–224. Reprinted in Grosz et al. (1986).

Greenberg, S., Ellis, D., and Hollenback, J. (1996). Insights into spoken language gleaned from phonetic transcription of the Switchboard corpus. In *ICSLP-96*, Philadelphia, PA, pp. S24–27.

Greene, B. B. and Rubin, G. M. (1971). Automatic grammatical tagging of English. Department of Linguistics, Brown University, Providence, Rhode Island.

Grefenstette, G. (1994). *Explorations in Automatic Thesaurus Discovery*. Kluwer, Norwell, MA.

Grefenstette, G. (1998). Producing intelligent telegraphic text reduction to provide an audio scanning service for the blind. In *AAAI 1998 Spring Symposium on Intelligent Text Summarization*, pp. 102–108.

Grefenstette, G. (1999). Tokenization. In van Halteren, H. (Ed.), *Syntactic Wordclass Tagging*. Kluwer.

Gregory, M. L. and Altun, Y. (2004). Using conditional random fields to predict pitch accents in conversational speech. In *ACL-04*.

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., and Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. In *CLS-99*, pp. 151–166. University of Chicago, Chicago.

Grenager, T. and Manning, C. D. (2006). Unsupervised Discovery of a Statistical Verb Lexicon. In *EMNLP 2006*.

Grice, H. P. (1957). Meaning. *Philosophical Review*, *67*, 377–388. Reprinted in *Semantics*, edited by Danny D. Steinberg & Leon A. Jakobovits (1971), Cambridge University Press, pages 53–59.

Grice, H. P. (1975). Logic and conversation. In Cole, P. and Morgan, J. L. (Eds.), *Speech Acts: Syntax and Semantics Volume 3*, pp. 41–58. Academic Press.

Grice, H. P. (1978). Further notes on logic and conversation. In Cole, P. (Ed.), *Pragmatics: Syntax and Semantics Volume 9*, pp. 113–127. Academic Press.

Gries, S. T. and Stefanowitsch, A. (Eds.). (2006). *Corpus-Based Approaches to Metaphor and Metonymy*. Mouton de Gruyter.

Grishman, R. and Sundheim, B. (1995). Design of the MUC-6 evaluation. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, San Francisco, pp. 1–11. Morgan Kaufmann.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, *28*, 267–283.

Grosjean, F., Grosjean, L., and Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology*, *11*, 58–81.

Grosz, B. J. and Hirschberg, J. (1992). Some intonational characteristics of discourse structure. In *ICSLP-92*, Vol. 1, pp. 429–432.

Grosz, B. J. (1977a). The representation and use of focus in a system for understanding dialogs. In *IJCAI-77*, pp. 67–76. Morgan Kaufmann. Reprinted in Grosz et al. (1986).

Grosz, B. J. (1977b). *The Representation and Use of Focus in Dialogue Understanding*. Ph.D. thesis, University of California, Berkeley.

Grosz, B. J., Joshi, A. K., and Weinstein, S. (1983). Providing a unified account of definite noun phrases in English. In *ACL-83*, pp. 44–50.

Grosz, B. J., Joshi, A. K., and Weinstein, S. (1995a). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, *21*(2), 203–225.

Grosz, B. J., Joshi, A. K., and Weinstein, S. (1995b). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, *21*(2).

Grosz, B. J. and Sidner, C. L. (1980). Plans for discourse. In Cohen, P. R., Morgan, J., and Pollack, M. E. (Eds.), *Intentions in Communication*, pp. 417–444. MIT Press.

Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, *12*(3), 175–204.

Grosz, B. J., Sparck Jones, K., and Webber, B. L. (Eds.). (1986). *Readings in Natural Language Processing*. Morgan Kaufmann, Los Altos, CA.

Gruber, J. S. (1965). *Studies in Lexical Relations*. Ph.D. thesis, MIT.

Grudin, J. T. (1983). Error patterns in novice and skilled transcription typing. In Cooper, W. E. (Ed.), *Cognitive Aspects of Skilled Typewriting*, pp. 121–139. Springer-Verlag.

Guindon, R. (1988). A multidisciplinary perspective on dialogue structure in user-advisor dialogues. In Guindon, R. (Ed.), *Cognitive Science And Its Applications For Human-Computer Interaction*, pp. 163–200. Lawrence Erlbaum.

Gundel, J. K., Hedberg, N., and Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*(2), 274–307.

Gupta, V., Lennig, M., and Mermelstein, P. (1988). Fast search strategy in a large vocabulary word recognizer. *Journal of the Acoustical Society of America*, *84*(6), 2007–2017.

Gupta, V., Lennig, M., and Mermelstein, P. (1992). A language model for very large-vocabulary speech recognition. *Computer Speech and Language*, *6*, 331–344.

Gusfield, D. (1997). *Algorithms on strings, trees, and sequences: computer science and computational biology*. Cambridge University Press.

Guy, G. R. (1980). Variation in the group and the individual: The case of final stop deletion. In Labov, W. (Ed.), *Locating Language in Time and Space*, pp. 1–36. Academic.

Habash, N., Rambow, O., and Kiraz, G. A. (2005). Morphological analysis and generation for arabic dialects. In *ACL Workshop on Computational Approaches to Semitic Languages*, pp. 17–24.

Hachey, B. and Grover, C. (2005). Sentence extraction for legal text summarization. In *IJCAI-05*, pp. 1686–1687.

Hafer, M. A. and Weiss, S. F. (1974). Word segmentation by letter successor varieties. *Information Storage and Retrieval*, *10*(11-12), 371–385.

Haghighi, A. and Klein, D. (2006). Prototype-driven grammar induction. In *COLING/ACL 2006*, pp. 881–888.

Hain, T. (2002). Implicit pronunciation modelling in asr. In *Proceedings of ISCA Pronunciation Modeling Workshop*.

Hain, T., Woodland, P. C., Evermann, G., and Povey, D. (2001). New features in the CU-HTK system for transcription of conversational telephone speech. In *IEEE ICASSP-01*, Salt Lake City, Utah.

Hajič, J. (1998). *Building a Syntactically Annotated Corpus: The Prague Dependency Treebank*, pp. 106–132. Karolinum, Prague/Praha.

Hajič, J. (2000). Morphological tagging: Data vs. dictionaries. In *NAACL 2000*. Seattle.

Hajič, J. and Hladká, B. (1998). Tagging inflective languages: Prediction of morphological categories for a rich, structured tagset. In *COLING/ACL-98*, Montreal, Canada.

Hajič, J., Krbec, P., Květoň, P., Oliva, K., and Petkevič, V. (2001). Serial Combination of Rules and Statistics: A Case Study in Czech Tagging. In *ACL-01*, Toulouse, France.

Hakkani-Tür, D., Oflazer, K., and Tür, G. (2002). Statistical morphological disambiguation for agglutinative languages. *Journal of Computers and Humanities*, *36*(4).

Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, *30*(4), 609–642.

Hall, K. and Johnson, M. (2003). Language modeling using efficient best-first bottom-up parsing. In *IEEE ASRU-03*, pp. 507–512.

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. Edward Arnold, London.

Halliday, M. A. K. and Hasan, R. (1976). *Cohesion in English*. Longman, London. English Language Series, Title No. 9.

Hammond, M. (1997). Parsing in OT. Alternative title "Parsing syllables: Modeling OT computationally". Rutgers Optimality Archive ROA-222-1097.

Hamza, W., Bakis, R., Eide, E. M., Picheny, M. A., and Pitrelli, J. F. (2004). The IBM expressive speech synthesis system. In *ICSLP-04*, Jeju, Korea.

Hankamer, J. (1986). Finite state morphology and left to right phonology. In *Proceedings of the Fifth West Coast Conference on Formal Linguistics*, pp. 29–34.

Hankamer, J. and Black, H. A. (1991). Current approaches to computational morphology. Unpublished manuscript.

Harabagiu, S., Pasca, M., and Maiorano, S. (2000). Experiments with open-domain textual question answering. In *COLING-00*, Saarbrücken, Germany.

Harris, C. M. (1953). A study of the building blocks in speech. *Journal of the Acoustical Society of America*, *25*(5), 962–969.

Harris, R. A. (2005). *Voice Interaction Design: Crafting the New Conversational Speech Systems*. Morgan Kaufmann.

Harris, Z. S. (1946). From morpheme to utterance. *Language*, *22*(3), 161–183.

Harris, Z. S. (1954). Distributional structure. *Word*, *10*, 146–162. Reprinted in J. Fodor and J. Katz, The structure of language: Readings in the philosophy of language, Prentice-hall, 1964 and in Z. S. Harris, Papers in structural and transformational linguistics, Reidel, Dordrecht, 1970, 775–794.

Harris, Z. S. (1962). *String Analysis of Sentence Structure*. Mouton, The Hague.

Harris, Z. S. (1968). *Mathematical Structures of Language*. John Wiley.

Harris, Z. S. (1988). *Language and Information*. Columbia University Press.

Hasegawa-Johnson, M. and et al (2005). Landmark-based speech recognition: Report of the 2004 Johns Hopkins Summer Workshop. In *IEEE ICASSP-05*.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2001). *The Elements of Statistical Learning*. Springer.

Haviland, S. E. and Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behaviour*, *13*, 512–521.

Hayes, B. (2004). Phonological acquisition in optimality theory: the early stages. In Kager, R., Pater, J., and Zonneveld, W. (Eds.), *Constraints in Phonological Acquisition*. Cambridge University Press.

Hayes, B. and Wilson, C. (2007). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*. To appear.

Hayes, E. and Bayer, S. (1991). Metaphoric generalization through sort coercion. In *ACL-91*, Berkeley, CA, pp. 222–228. ACL.

Hearst, M. A. (1991). Noun homograph disambiguation. In *Proceedings of the 7th Annual Conference of the University of Waterloo Centre for the New OED and Text Research*, Oxford, pp. 1–19.

Hearst, M. A. (1992). Automatic acquisition of hyponyms from large text corpora. In *COLING-92*, Nantes, France.

Hearst, M. A. (1994). Multi-paragraph segmentation of expository text. In *ACL-94*, pp. 9–16.

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

Hearst, M. A. (1998). Automatic discovery of WordNet relations. In Fellbaum, C. (Ed.), *Wordnet: An Electronic Lexical Database*. MIT Press.

Heeman, P. A. (1999). POS tags and decision trees for language modeling. In *EMNLP/VLC-99*, College Park, MD, pp. 129–137.

Heeman, P. A. and Allen, J. (1999). Speech repairs, intonational phrases and discourse markers: Modeling speakers' utterances in spoken dialog. *Computational Linguistics*, *25*(4).

Heikkilä, J. (1995). A TWOL-based lexicon and feature system for English. In Karlsson, F., Voutilainen, A., Heikkilä, J., and Anttila, A. (Eds.), *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*, pp. 103–131. Mouton de Gruyter, Berlin.

Heim, I. and Kratzer, A. (1998). *Semantics in a Generative Grammar*. Blackwell Publishers, Malden, MA.

Heinz, J. M. and Stevens, K. N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America*, *33*, 589–596.

Hemphill, C. T., Godfrey, J., and Doddington, G. (1990). The ATIS spoken language systems pilot corpus. In *Proceedings DARPA Speech and Natural Language Workshop*, Hidden Valley, PA, pp. 96–101. Morgan Kaufmann.

Hermansky, H. (1990). Perceptual linear predictive (PLP) analysis of speech. *Journal of the Acoustical Society of America*, *87*(4), 1738–1752.

Hillard, D., Huang, Z., Ji, H., Grishman, R., Hakkani-Tür, D., Harper, M., Ostendorf, M., and Wang, W. (2006). Impact of automatic comma prediction on pos/name tagging of speech. In *Proceedings of IEEE/ACL 06 Workshop on Spoken Language Technology*, Aruba.

Hindle, D. (1983). Deterministic parsing of syntactic non-fluencies. In *ACL-83*, pp. 123–128.

Hindle, D. and Rooth, M. (1990). Structural ambiguity and lexical relations. In *Proceedings DARPA Speech and Natural Language Workshop*, Hidden Valley, PA, pp. 257–262. Morgan Kaufmann.

Hindle, D. and Rooth, M. (1991). Structural ambiguity and lexical relations. In *ACL-91*, Berkeley, CA, pp. 229–236.

Hindle, D. (1990). Noun classification from predicate-argument structures. In *ACL-90*, Pittsburgh, PA, pp. 268–275.

Hinkelman, E. A. and Allen, J. (1989). Two constraints on speech act ambiguity. In *ACL-89*, Vancouver, Canada, pp. 212–219.

Hintikka, J. (1969). Semantics for propositional attitudes. In Davis, J. W., Hockney, D. J., and Wilson, W. K. (Eds.), *Philosophical Logic*, pp. 21–45. D. Reidel, Dordrecht, Holland.

Hirschberg, J. and Litman, D. J. (1993). Empirical Studies on the Disambiguation of Cue Phrases. *Computational Linguistics*, *19*(3), 501–530.

Hirschberg, J. (1993). Pitch Accent in Context: Predicting Intonational Prominence from Text. *Artificial Intelligence*, *63*(1-2), 305–340.

Hirschberg, J. and Litman, D. J. (1993). Empirical studies on the disambiguation of cue phrases. *Computational Linguistics*, *19*(3), 501–530.

Hirschberg, J., Litman, D. J., and Swerts, M. (2001). Identifying user corrections automatically in spoken dialogue systems. In *NAACL 2001*.

Hirschberg, J. and Nakatani, C. (1996). A prosodic analysis of discourse segments in direction-giving monologues. In *ACL-96*, Santa Cruz, CA, pp. 286–293.

Hirschberg, J. and Pierrehumbert, J. B. (1986). The intonational structuring of discourse. In *ACL-86*, New York, pp. 136–144.

Hirschman, L. and Blaschke, C. (2006). Evaluation of text mining in biology. In Ananiadou, S. and McNaught, J. (Eds.), *Text Mining for Biology and Biomedicine*, chap. 9, pp. 213–245. Artech House, Norwood, MA.

Hirschman, L. and Pao, C. (1993). The cost of errors in a spoken language system. In *EUROSPEECH-93*, pp. 1419–1422.

Hirst, G. (1987). *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press.

Hirst, G. (1988). Resolving lexical ambiguity computationally with spreading activation and polaroid words. In Small, S. L., Cottrell, G. W., and Tanenhaus, M. K. (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence*, pp. 73–108. Morgan Kaufmann.

Hirst, G. and Budanitsky, A. (2005). Correcting real-word spelling errors by restoring lexical cohesion. *Natural Language Engineering*, *11*, 87–111.

Hirst, G. and Charniak, E. (1982). Word sense and case slot disambiguation. In *AAAI-82*, pp. 95–98.

Hjelmslev, L. (1969). *Prologomena to a Theory of Language*. University of Wisconsin Press. Translated by Francis J. Whitfield; original Danish edition 1943.

Hobbs, J. R. (1977). 38 examples of elusive antecedents from published texts. Tech. rep. 77-2, Department of Computer Science, City University of New York.

Hobbs, J. R. (1978). Resolving pronoun references. *Lingua*, *44*, 311–338. Reprinted in Grosz et al. (1986).

Hobbs, J. R. (1979a). Coherence and coreference. *Cognitive Science*, *3*, 67–90.

Hobbs, J. R. (1979b). Metaphor, metaphor schemata, and selective inferencing. Tech. rep. 204, SRI.

Hobbs, J. R. (1990). *Literature and Cognition*. CSLI Lecture Notes 21.

Hobbs, J. R., Appelt, D. E., Bear, J., Israel, D., Kameyama, M., Stickel, M. E., and Tyson, M. (1997). FASTUS: A cascaded finite-state transducer for extracting information from natural-language text. In Roche, E. and Schabes, Y. (Eds.), *Finite-State Language Processing*, pp. 383–406. MIT Press.

Hobbs, J. R. and Shieber, S. M. (1987). An algorithm for generating quantifier scopings. *Computational Linguistics*, *13*(1), 47–55.

Hobbs, J. R., Stickel, M. E., Appelt, D. E., and Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, *63*, 69–142.

Hockenmaier, J. and Steedman, M. (2002). Generative models for statistical parsing with Combinatory Categorial Grammar. In *ACL-02*, Philadelphia, PA.

Hofstadter, D. R. (1997). *Le ton beau de marot*. Basic Books.

Holmes, D. I. (1994). Authorship attribution. *Computers and the Humanities*, *28*, 87–106.

Honal, M. and Schultz, T. (2003). Correction of disfluencies in spontaneous speech using a noisy-channel approach. In *EUROSPEECH-03*.

Honal, M. and Schultz, T. (2005). Automatic disfluency removal on recognized spontaneous speech - rapid adaptation to speaker-dependent disfluencies. In *IEEE ICASSP-05*.

Hopcroft, J. E. and Ullman, J. D. (1979). *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, MA.

Hori, C. and Furui, S. (2004). Speech summarization: an approach through word extraction and a method for evaluation. *IEICE Transactions on Information and Systems*, *87*, 15–25.

Horning, J. J. (1969). *A study of grammatical inference*. Ph.D. thesis, Stanford University.

House, A. S., Williams, C. E., Hecker, M. H. L., and Kryter, K. D. (1965). Articulation-Testing Methods: Consonantal Differentiation with a Closed-Response Set. *Journal of the Acoustical Society of America*, *37*, 158–166.

Householder, F. W. (1995). Dionysius Thrax, the *technai*, and Sextus Empiricus. In Koerner, E. F. K. and Asher, R. E. (Eds.), *Concise History of the Language Sciences*, pp. 99–103. Elsevier Science, Oxford.

Hovy, E., Hermjakob, U., and Ravichandran, D. (2002). A question/answer typology with surface text patterns. In *HLT-01*.

Hovy, E. H. (1990). Parsimonious and profligate approaches to the question of discourse structure relations. In *Proceedings of the Fifth International Workshop on Natural Language Generation*, Dawson, PA, pp. 128–136.

Hovy, E. H. and Lin, C.-Y. (1999). Automated text summarization in SUMMARIST. In Mani, I. and Maybury, M. T. (Eds.), *Advances in Automatic Text Summarization*, pp. 81–94. MIT Press.

Hovy, E. H., Marcus, M. P., Palmer, M., Ramshaw, L. A., and Weischedel, R. (2006). Ontonotes: The 90% solution. In *HLT-NAACL-06*.

Howes, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America*, *29*, 296–305.

Huang, C., Chang, E., Zhou, J., and Lee, K.-F. (2000). Accent modeling based on pronunciation dictionary adaptation for large vocabulary Mandarin speech recognition. In *ICSLP-00*, Beijing, China.

Huang, L. and Chiang, D. (2005). Better k-best parsing. In *IWPT-05*, pp. 53–64.

Huang, X., Acero, A., and Hon, H.-W. (2001). *Spoken Language Processing: A Guide to Theory, Algorithm, and System Development*. Prentice Hall, Upper Saddle River, NJ.

Huddleston, R. and Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Hudson, R. A. (1984). *Word Grammar*. Blackwell, Oxford.

Huffman, D. A. (1954). The synthesis of sequential switching circuits. *Journal of the Franklin Institute*, *3*, 161–191. Continued in Volume 4.

Huffman, S. (1996). Learning information extraction patterns from examples. In Wertmer, S., Riloff, E., and Scheller, G. (Eds.), *Connectionist, Statistical, and Symbolic Approaches to Learning Natural Language Processing*, pp. 246–260. Springer, Berlin.

Huls, C., Bos, E., and Classen, W. (1995). Automatic referent resolution of deictic and anaphoric expressions. *Computational Linguistics*, *21*(1), 59–79.

Hunt, A. J. and Black, A. W. (1996a). Unit selection in a concatenative speech synthesis system using a large speech database. In *IEEE ICASSP-96*, Atlanta, GA, Vol. 1, pp. 373–376. IEEE.

Hunt, A. J. and Black, A. W. (1996b). Unit selection in a concatenative speech synthesis system using a large speech database. In *IEEE ICASSP-06*, Vol. 1, pp. 373–376.

Hutchins, J. (1997). From first conception to first demonstration: the nascent years of machine translation, 1947–1954. A chronology. *Machine Translation*, *12*, 192–252.

Hutchins, W. J. and Somers, H. L. (1992). *An Introduction to Machine Translation*. Academic Press.

Hutchins, W. J. (1986). *Machine Translation: Past, Present, Future*. Ellis Horwood, Chichester, England.

Huybregts, R. (1984). The weak inadequacy of context-free phrase structure grammars. In de Haan, G., Trommele, M., and Zonneveld, W. (Eds.), *Van Periferie naar Kern*. Foris, Dordrecht. Cited in Pullum (1991).

Ide, N. M. and Veronis, J. (Eds.). (1998a). *Computational Linguistics: Special Issue on Word Sense Disambiguation*, Vol. 24. MIT Press.

Ide, N. M. and Véronis, J. (1998b). Introduction to the special issue on word sense disambiguation. *Computational Linguistics*, *24*(1), 1–40.

Irons, E. T. (1961). A syntax directed compiler for ALGOL 60. *Communications of the ACM*, *4*, 51–55.

ISO8601 (2004). Data elements and interchange formats information interchange representation of dates and times. Tech. rep., International Organization for Standards (ISO).

Issar, S. and Ward, W. (1993). Cmu's robust spoken language understanding system. In *Eurospeech 93*, pp. 2147–2150.

Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *ASSP-32*, 67–72.

Iverson, E. and Helmreich, S. (1992). Metallel: An integrated approach to non-literal phrase interpretation. *Computational Intelligence*, *8*(3).

Iyer, R. M. and Ostendorf, M. (1999a). Modeling long distance dependencies in language: Topic mixtures versus dynamic cache model. *IEEE Transactions on Speech and Audio Processing*, *7*.

Iyer, R. M. and Ostendorf, M. (1999b). Relevance weighting for combining multi-domain data for n-gram language modeling. *Computer Speech and Language*, *13*(3), 267–282.

Iyer, R. M. and Ostendorf, M. (1997). Transforming out-of-domain estimates to improve in-domain language models. In *EUROSPEECH-97*, pp. 1975–1978.

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles*, *44*, 223–227.

Jaccard, P. (1912). The distribution of the flora of the alpine zone. *New Phytologist*, *11*, 37–50.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language*, *51*(3), 639–671.

Jackendoff, R. (1983). *Semantics and Cognition*. MIT Press.

Jackendoff, R. (1990). *Semantic Structures*. MIT Press.

Jackson, P. and Moulinier, I. (2002). *Natural language processing for online applications: text retrieval, extraction, and categorization*. John Benjamins Publishing Company.

Jacobs, P. (1985). *A Knowledge-Based Approach to Language Generation*. Ph.D. thesis, University of California, Berkeley, CA. Available as University of California at Berkeley Computer Science Division Tech. rep. #86/254.

Jacobs, P. (1987). Knowledge-based natural language generation. *Artificial Intelligence*, *33*, 325–378.

Jacobs, P. S. and Rau, L. F. (1990). SCISOR: A system for extracting information from on-line news. *Communications of the ACM*, *33*(11), 88–97.

Jacquemin, C. (1997). Guessing morphology from terms and corpora. In *SIGIR 1997*, Philadelphia, PA, pp. 156–165.

Jakobson, R. (1939). Observations sur le classement phonologique des consonnes. In Blancquaert, E. and Pée, W. (Eds.), *ICPhS-39*, Ghent, pp. 34–41.

Janssen, T. M. V. (1997). Compositionality. In van Benthem, J. and ter Meulen, A. (Eds.), *Handbook of Logic and Language*, chap. 7, pp. 417–473. North-Holland, Amsterdam.

Jarosz, G. (2006). Richness of the base and probabilistic unsupervised learning in optimality theory. In *Proceedings of ACL SIGPHON*, New York, NY, pp. 50–59.

Jarosz, G. (2008). Restrictiveness and phonological grammar and lexicon learning. In *CLS 43*. In press.

Järvinen, T. and Tapanainen, P. (1997). A dependency parser for English. Tech. rep. TR-1, Department of General Linguistics, University of Helsinki, Helsinki.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens 'yeah' and 'mm hm'. *Papers in Linguistics*, *17*(2), 197–216.

Jeffreys, H. (1948). *Theory of Probability*. Clarendon Press, Oxford. 2nd edn Section 3.23.

Jekat, S., Klein, A., Maier, E., Maleck, I., Mast, M., and Quantz, J. (1995). Dialogue acts in verbmobil.. Verbmobil–Report–65–95.

Jelinek, F. (1969). A fast sequential decoding algorithm using a stack. *IBM Journal of Research and Development*, *13*, 675–685.

Jelinek, F. (1976). Continuous speech recognition by statistical methods. *Proceedings of the IEEE*, *64*(4), 532–557.

Jelinek, F. (1988). Address to the first workshop on the evaluation of natural language processing systems. December 7, 1988.

Jelinek, F. (1990). Self-organized language modeling for speech recognition. In Waibel, A. and Lee, K.-F. (Eds.), *Readings in Speech Recognition*, pp. 450–506. Morgan Kaufmann, Los Altos. Originally distributed as IBM technical report in 1985.

Jelinek, F. (1997). *Statistical Methods for Speech Recognition*. MIT Press.

Jelinek, F. and Lafferty, J. D. (1991). Computation of the probability of initial substring generation by stochastic context-free grammars. *Computational Linguistics*, *17*(3), 315–323.

Jelinek, F., Lafferty, J. D., Magerman, D. M., Mercer, R. L., Ratnaparkhi, A., and Roukos, S. (1994). Decision tree parsing using a hidden derivation model. In *ARPA Human Language Technologies Workshop*, Plainsboro, N.J., pp. 272–277. Morgan Kaufmann.

Jelinek, F. and Mercer, R. L. (1980). Interpolated estimation of Markov source parameters from sparse data. In Gelsema, E. S. and Kanal, L. N. (Eds.), *Proceedings, Workshop on Pattern Recognition in Practice*, pp. 381–397. North Holland, Amsterdam.

Jelinek, F., Mercer, R. L., and Bahl, L. R. (1975). Design of a linguistic statistical decoder for the recognition of continuous speech. *IEEE Transactions on Information Theory*, *IT-21*(3), 250–256.

Jiang, J. J. and Conrath, D. W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. In *ROCLING X*, Taiwan.

Jilka, M., Mohler, G., and Dogil, G. (1999). Rules for the generation of ToBI-based American English intonation. *Speech Communication*, *28*(2), 83–108.

Jiménez, V. M. and Marzal, A. (2000). Computation of the *n* best parse trees for weighted and stochastic context-free grammars. In *Advances in Pattern Recognition: Proceedings of the Joint IAPR International Workshops, SSPR 2000 and SPR 2000*, Alicante, Spain, pp. 183–192. Springer.

Jing, H. (2000). Sentence reduction for automatic text summarization. In *ANLP 2000*, Seattle, WA, pp. 310–315.

Jing, H. (2002). Using hidden Markov modeling to decompose human-written summaries. *Computational Linguistics*, *28*(4), 527–543.

Johnson, C. D. (1972). *Formal Aspects of Phonological Description*. Mouton, The Hague. Monographs on Linguistic Analysis No. 3.

Johnson, C. (1999). Syntactic and semantic principles of FrameNet annotation, version 1. Tech. rep. TR-99-018, ICSI, Berkeley, CA.

Johnson, K. (2003). *Acoustic and Auditory Phonetics*. Blackwell, Oxford. 2nd ed.

Johnson, M. (1984). A discovery procedure for certain phonological rules. In *COLING-84*, Stanford, CA, pp. 344–347.

Johnson, M. (1988). *Attribute-Value Logic and the Theory of Grammar*. CSLI Lecture Notes. Chicago University Press, Chicago.

Johnson, M. (1990). Expressing disjunctive and negative feature constraints with classical first-order logic. In *ACL-90*, Pittsburgh, PA, pp. 173–179.

Johnson, M. (1998a). Finite-state approximation of constraint-based grammars using left-corner grammar transforms. In *COLING/ACL-98*, Montreal, pp. 619–623.

Johnson, M. (1998b). PCFG models of linguistic tree representations. *Computational Linguistics*, *24*(4), 613–632.

Johnson, M. (2001). Joint and conditional estimation of tagging and parsing models. In *ACL-01*, pp. 314–321.

Johnson, M., Geman, S., Canon, S., Chi, Z., and Riezler, S. (1999). Estimators for stochastic "unification-based" grammars. In *ACL-99*, pp. 535–541.

Johnson, S. C. and Lesk, M. E. (1978). Language development tools. *Bell System Technical Journal*, *57*(6), 2155–2175.

Johnson, W. E. (1932). Probability: deductive and inductive problems (appendix to). *Mind*, *41*(164), 421–423.

Johnson-Laird, P. N. (1983). *Mental Models*. Harvard University Press, Cambridge, MA.

Johnston, M., Ehlen, P., Gibbon, D., and Liu, Z. (2007). The multimodal presentation dashboard. In *NAACL HLT 2007 Workshop 'Bridging the Gap'*.

Jones, D. A., Gibson, E., Shen, W., Granoien, N., Herzog, M., Reynolds, D. A., and Weinstein, C. (2005). Measuring human readability of machine generated text: Three case studies in speech recognition and machine translation. In *IEEE ICASSP-05*, pp. 18–23.

Jones, D. A., Wolf, F., Gibson, E., Williams, E., Fedorenko, E., Reynolds, D. A., and Zissman, M. (2003). Measuring the readability of automatic speech-to-text transcripts. In *EUROSPEECH-03*, pp. 1585–1588.

Jones, M. A. and McCoy, K. (1992). Transparently-motivated metaphor generation. In Dale, R., Hovy, E. H., Rösner, D., and Stock, O. (Eds.), *Aspects of Automated Natural Language Generation*, Lecture Notes in Artificial Intelligence 587, pp. 183–198. Springer Verlag, Berlin.

Jones, M. P. and Martin, J. H. (1997). Contextual spelling correction using latent semantic analysis. In *Proceedings of the 5th Conference on Applied Natural Language Processing (ANLP'97)*, Washington, D.C., pp. 166–173.

Joos, M. (1950). Description of language design. *Journal of the Acoustical Society of America*, *22*, 701–708.

Joshi, A. K., Vijay-Shanker, K., and Weir, D. J. (1991). The convergence of mildly context-sensitive grammatical formalisms. In Sells, P., Shieber, S., and Wasow, T. (Eds.), *Foundational issues in natural language processing*, pp. 31–81. MIT Press.

Joshi, A. K. (1985). Tree adjoining grammars: how much context-sensitivity is required to provide reasonable structural descriptions?. In Dowty, D. R., Karttunen, L., and Zwicky, A. (Eds.), *Natural Language Parsing*, pp. 206–250. Cambridge University Press.

Joshi, A. K. and Hopely, P. (1999). A parser from antiquity. In Kornai, A. (Ed.), *Extended Finite State Models of Language*, pp. 6–15. Cambridge University Press.

Joshi, A. K. and Kuhn, S. (1979). Centered logic: The role of entity centered sentence representation in natural language inferencing. In *IJCAI-79*, pp. 435–439.

Joshi, A. K. and Srinivas, B. (1994). Disambiguation of super parts of speech (or supertags): Almost parsing. In *COLING-94*, Kyoto, pp. 154–160.

Joshi, A. K. and Weinstein, S. (1981). Control of inference: Role of some aspects of discourse structure – centering. In *IJCAI-81*, pp. 385–387.

Jun, S.-A. (Ed.). (2005). *Prosodic Typology and Transcription: A Unified Approach*. Oxford University Press.

Juneja, A. and Espy-Wilson, C. (2003). Speech segmentation using probabilistic phonetic feature hierarchy and support vector machines. In *IJCNN 2003*.

Junqua, J. C. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers. *Journal of the Acoustical Society of America*, *93*(1), 510–524.

Juola, P. (1999). Measuring linguistic complexity. Presented at the 4th Conference on Conceptual Structure, Discourse, and Language (CSDL-4), Georgia.

Jurafsky, D., Ward, W., Jianping, Z., Herold, K., Xiuyang, Y., and Sen, Z. (2001). What kind of pronunciation variation is hard for triphones to model?. In *IEEE ICASSP-01*, Salt Lake City, Utah, pp. I.577–580.

Jurafsky, D. (1992). *An On-line Computational Model of Human Sentence Interpretation: A Theory of the Representation and Use of Linguistic Knowledge*. Ph.D. thesis, University of California, Berkeley, CA. Available as University of California at Berkeley Computer Science Division Tech. rep. #92/676.

Jurafsky, D., Bell, A., Fosler-Lussier, E., Girand, C., and Raymond, W. D. (1998). Reduction of English function words in Switchboard. In *ICSLP-98*, Sydney, Vol. 7, pp. 3111–3114.

Jurafsky, D., Bell, A., Gregory, M. L., and Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J. L. and Hopper, P. (Eds.), *Frequency and the Emergence of Linguistic Structure*, pp. 229–254. Benjamins, Amsterdam.

Jurafsky, D., Wooters, C., Tajchman, G., Segal, J., Stolcke, A., Fosler, E., and Morgan, N. (1994). The Berkeley restaurant project. In *ICSLP-94*, Yokohama, Japan, pp. 2139–2142.

Jurafsky, D., Wooters, C., Tajchman, G., Segal, J., Stolcke, A., Fosler, E., and Morgan, N. (1995). Using a stochastic context-free grammar as a language model for speech recognition. In *IEEE ICASSP-95*, pp. 189–192. IEEE.

Kager, R. (2000). *Optimality Theory*. Cambridge University Press.

Kahn, D. (1976). *Syllable-based Generalizations in English Phonology*. Ph.D. thesis, MIT.

Kameyama, M. (1986). A property-sharing constraint in centering. In *ACL-86*, New York, pp. 200–206.

Kamm, C. A. (1994). User interfaces for voice applications. In Roe, D. B. and Wilpon, J. G. (Eds.), *Voice Communication Between Humans and Machines*, pp. 422–442. National Academy Press, Washington, D.C.

Kan, M. Y., Klavans, J. L., and McKeown, K. R. (1998). Linear segmentation and segment significance. In *Proc. 6th Workshop on Very Large Corpora (WVLC-98)*, Montreal, Canada, pp. 197–205.

Kaplan, R. M., Riezler, S., King, T. H., Maxwell, J. T., Vasserman, A., and Crouch, R. (2004). Speed and accuracy in shallow and deep stochastic parsing. In *HLT-NAACL-04*.

Kaplan, R. M. (1973). A general syntactic processor. In Rustin, R. (Ed.), *Natural Language Processing*, pp. 193–241. Algorithmics Press, New York.

Kaplan, R. M. (1987). Three seductions of computational psycholinguistics. In Whitelock, P., Wood, M. M., Somers, H. L.,

Johnson, R., and Bennett, P. (Eds.), *Linguistic Theory and Computer Applications*, pp. 149–188. Academic Press, London.

Kaplan, R. M. and Bresnan, J. (1982). Lexical-functional grammar: A formal system for grammatical representation. In Bresnan, J. (Ed.), *The Mental Representation of Grammatical Relations*, pp. 173–281. MIT Press.

Kaplan, R. M. and Kay, M. (1981). Phonological rules and finite-state transducers. Paper presented at the Annual meeting of the Linguistics Society of America. New York.

Kaplan, R. M. and Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, *20*(3), 331–378.

Karamanis, N. (2003). *Entity Coherence for Descriptive Text Structuring*. Ph.D. thesis, University of Edinburgh.

Karamanis, N. (2006). Evaluating centering for sentence ordering in two new domains. In *HLT-NAACL-06*.

Karamanis, N. (2007). Supplementing entity coherence with local rhetorical relations for information ordering. *Journal of Logic, Language and Information*. To appear.

Karamanis, N. and Manurung, H. M. (2002). Stochastic text structuring using the principle of continuity. In *INLG 2002*, pp. 81–88.

Karamanis, N., Poesio, M., Mellish, C., and Oberlander, J. (2004). Evaluating centering-based metrics of coherence for text structuring using a reliably annotated corpus. In *ACL-04*.

Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. *Journal of Linguistics*, *43*, 365–392.

Karlsson, F., Voutilainen, A., Heikkilä, J., and Anttila, A. (1995a). *Constraint Grammar — A language-independent system for parsing unrestricted text*. Mouton de Gruyter, Berlin.

Karlsson, F., Voutilainen, A., Heikkilä, J., and Anttila, A. (Eds.). (1995b). *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Mouton de Gruyter, Berlin.

Karttunen, L., Chanod, J., Grefenstette, G., and Schiller, A. (1996). Regular expressions for language engineering. *Natural Language Engineering*, *2*(4), 305–238.

Karttunen, L. (1983). KIMMO: A general morphological processor. In *Texas Linguistics Forum 22*, pp. 165–186.

Karttunen, L. (1993). Finite-state constraints. In Goldsmith, J. (Ed.), *The Last Phonological Rule*, pp. 173–194. University of Chicago Press.

Karttunen, L. (1998). The proper treatment of optimality in computational phonology. In *Proceedings of FSMNLP'98: International Workshop on Finite-State Methods in Natural Language Processing*, Bilkent University. Ankara, Turkey, pp. 1–12.

Karttunen, L. (1999). Comments on Joshi. In Kornai, A. (Ed.), *Extended Finite State Models of Language*, pp. 16–18. Cambridge University Press.

Karttunen, L. and Kay, M. (1985). Structure sharing with binary trees. In *ACL-85*, Chicago, pp. 133–136. ACL.

Kasami, T. (1965). An efficient recognition and syntax analysis algorithm for context-free languages. Tech. rep. AFCRL-65-758, Air Force Cambridge Research Laboratory, Bedford, MA.

Kashyap, R. L. and Oommen, B. J. (1983). Spelling correction using probabilistic methods. *Pattern Recognition Letters*, *2*, 147–154.

Kasper, R. T. and Rounds, W. C. (1986). A logical semantics for feature structures. In *ACL-86*, New York, pp. 257–266.

Kataja, L. and Koskenniemi, K. (1988). Finite state description of Semitic morphology. In *COLING-88*, Budapest, pp. 313–315.

Katz, J. J. and Fodor, J. A. (1963). The structure of a semantic theory. *Language*, *39*, 170–210.

Katz, S. M. (1987). Estimation of probabilities from sparse data for the language model component of a speech recogniser. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *35*(3), 400–401.

Kawahara, T., Hasegawa, M., Shitaoka, K., Kitade, T., and Nanjo, H. (2004). Automatic indexing of lecture presentations using unsupervised learning of presumed discourse markers. *Speech and Audio Processing, IEEE Transactions on*, *12*(4), 409–419.

Kawamoto, A. H. (1988). Distributed representations of ambiguous words and their resolution in connectionist networks. In Small, S. L., Cottrell, G. W., and Tanenhaus, M. (Eds.), *Lexical Ambiguity Resolution*, pp. 195–228. Morgan Kaufman.

Kay, M. (1967). Experiments with a powerful parser. In *Proc. 2eme Conference Internationale sur le Traitement Automatique des Langues*, Grenoble.

Kay, M. (1973). The MIND system. In Rustin, R. (Ed.), *Natural Language Processing*, pp. 155–188. Algorithmics Press, New York.

Kay, M. (1979). Functional grammar. In *BLS-79*, Berkeley, CA, pp. 142–158.

Kay, M. (1982). Algorithm schemata and data structures in syntactic processing. In Allén, S. (Ed.), *Text Processing: Text Analysis and Generation, Text Typology and Attribution*, pp. 327–358. Almqvist and Wiksell, Stockholm.

Kay, M. (1984). Functional unification grammar: A formalism for machine translation. In *COLING-84*, Stanford, CA, pp. 75–78.

Kay, M. (1985). Parsing in functional unification grammar. In Dowty, D. R., Karttunen, L., and Zwicky, A. (Eds.), *Natural Language Parsing*, pp. 251–278. Cambridge University Press.

Kay, M. (1987). Nonconcatenative finite-state morphology. In *EACL-87*, Copenhagen, Denmark, pp. 2–10.

Kay, M. and Röscheisen, M. (1988). Text-translation alignment. Tech. rep. P90-00143, Xerox Palo Alto Research Center, Palo Alto, CA.

Kay, M. and Röscheisen, M. (1993). Text-translation alignment. *Computational Linguistics*, *19*, 121–142.

Kay, P. and Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: The What's X Doing Y? construction. *Language*, *75*(1), 1–33.

Kazakov, D. (1997). Unsupervised learning of naïve morphology with genetic algorithms. In *ECML/Mlnet Workshop on Empirical Learning of Natural Language Processing Tasks*, Prague, pp. 105–111.

Keating, P. A., Byrd, D., Flemming, E., and Todaka, Y. (1994). Phonetic analysis of word and segment variation using the TIMIT corpus of American English. *Speech Communication*, *14*, 131–142.

Kehler, A. (1993). The effect of establishing coherence in ellipsis and anaphora resolution. In *ACL-93*, Columbus, Ohio, pp. 62–69.

Kehler, A. (1994a). Common topics and coherent situations: Interpreting ellipsis in the context of discourse inference. In *ACL-94*, Las Cruces, New Mexico, pp. 50–57.

Kehler, A. (1994b). Temporal relations: Reference or discourse coherence?. In *ACL-94*, Las Cruces, New Mexico, pp. 319–321.

Kehler, A. (1997a). Current theories of centering for pronoun interpretation: A critical evaluation. *Computational Linguistics*, *23*(3), 467–475.

Kehler, A. (1997b). Probabilistic coreference in information extraction. In *EMNLP 1997*, Providence, RI, pp. 163–173.

Kehler, A. (2000). *Coherence, Reference, and the Theory of Grammar*. CSLI Publications.

Kehler, A., Appelt, D. E., Taylor, L., and Simma, A. (2004). The (non)utility of predicate-argument frequencies for pronoun interpretation. In *HLT-NAACL-04*.

Keller, F. (2004). The entropy rate principle as a predictor of processing effort: An evaluation against eye-tracking data. In *EMNLP 2004*, Barcelona, pp. 317–324.

Keller, F. and Lapata, M. (2003). Using the web to obtain frequencies for unseen bigrams. *Computational Linguistics*, *29*, 459–484.

Kelley, L. G. (1979). *The True Interpreter: A History of Translation Theory and Practice in the West*. St. Martin's Press, New York.

Kelly, E. F. and Stone, P. J. (1975). *Computer Recognition of English Word Senses*. North-Holland, Amsterdam.

Kennedy, C. and Boguraev, B. (1996). Anaphora for everyone: Pronominal anaphora resolution without a parser. In *COLING-96*, Copenhagen, pp. 113–118.

Kernighan, M. D., Church, K. W., and Gale, W. A. (1990). A spelling correction program base on a noisy channel model. In *COLING-90*, Helsinki, Vol. II, pp. 205–211.

Kibble, R. and Power, R. (2000). An integrated framework for text planning and pronominalisation. In *INLG 2000*, pp. 77–84.

Kiefer, B., Krieger, H.-U., Carroll, J., and Malouf, R. (1999a). A bag of useful techniques for efficient and robust parsing. In *ACL-99*, pp. 535–541.

Kiefer, B., Krieger, H.-U., Carroll, J., and Malouf, R. (1999b). A bag of useful techniques for efficient and robust parsing. In *ACL-99*, College Park, MD, pp. 473–480.

Kilgarriff, A. (2001). English lexical sample task description. In *Proceedings of Senseval-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems*, Toulouse, France, pp. 17–20.

Kilgarriff, A. and Palmer, M. (Eds.). (2000). *Computing and the Humanities: Special Issue on SENSEVAL*, Vol. 34. Kluwer.

Kilgarriff, A. and Rosenzweig, J. (2000). Framework and results for English SENSEVAL. *Computers and the Humanities*, *34*(1-2).

Kim, D., Gales, M., Hain, T., and Woodland, P. C. (2004). Using vtln for broadcast news transcription. In *ICSLP-04*, Jeju, South Korea.

Kim, J. and Woodland, P. C. (2001). The use of prosody in a combined system for punctuation generation and speech recognition. In *EUROSPEECH-01*, pp. 2757–2760.

King, J. and Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602.

King, P. (1989). *A Logical Formalism for Head-Driven Phrase Structure Grammar*. Ph.D. thesis, University of Manchester.

Kinoshita, S., Cohen, K. B., Ogren, P. V., and Hunter, L. (2005). BioCreAtIvE Task1A: entity identification with a stochastic tagger.. *BMC Bioinformatics*, *6*(1).

Kintsch, W. (1974). *The Representation of Meaning in Memory*. Wiley, New York.

Kipper, K., Dang, H. T., and Palmer, M. (2000). Class-based construction of a verb lexicon. In *AAAI-00*, Austin, TX.

Kiraz, G. A. (1997). Compiling regular formalisms with rule features into finite-state automata. In *ACL/EACL-97*, Madrid, Spain, pp. 329–336.

Kiraz, G. A. (2000). Multitiered nonlinear morphology using multitape finite automata: A case study on syriac and arabic. *Computational Linguistics*, *26*(1), 77–105.

Kiraz, G. A. (2001). *Computational Nonlinear Morphology with Emphasis on Semitic Languages*. Cambridge University Press.

Kiraz, G. A. and Möbius, B. (1998). Multilingual syllabification using weighted finite-state transducers. In *Proceedings of 3rd ESCA Workshop on Speech Synthesis*, Jenolan Caves, pp. 59–64.

Kirchhoff, K., Fink, G. A., and Sagerer, G. (2002). Combining acoustic and articulatory feature information for robust speech recognition. *Speech Communication*, *37*, 303–319.

Kisseberth, C. W. (1969). On the abstractness of phonology: The evidence from Yawelmani. *Papers in Linguistics*, *1*, 248–282.

Kisseberth, C. W. (1970). On the functional unity of phonological rules. *Linguistic Inquiry*, *1*(3), 291–306.

Kita, K., Fukui, Y., Nagata, M., and Morimoto, T. (1996). Automatic acquisition of probabilistic dialogue models. In *ICSLP-96*, Philadelphia, PA, Vol. 1, pp. 196–199.

Klatt, D. H. (1975). Voice onset time, friction, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, *18*, 686–706.

Klatt, D. H. (1982). The Klattalk text-to-speech conversion system. In *IEEE ICASSP-82*, pp. 1589–1592.

Klatt, D. H. (1977). Review of the ARPA speech understanding project. *Journal of the Acoustical Society of America*, *62*(6), 1345–1366.

Klatt, D. H. (1979). Synthesis by rule of segmental durations in English sentences. In Lindblom, B. E. F. and Öhman, S. (Eds.), *Frontiers of Speech Communication Research*, pp. 287–299. Academic.

Klavans, J. L. (Ed.). (1995). *Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity and Generativity*. AAAI Press, Menlo Park, CA. AAAI Technical Report SS-95-01.

Kleene, S. C. (1951). Representation of events in nerve nets and finite automata. Tech. rep. RM-704, RAND Corporation. RAND Research Memorandum.

Kleene, S. C. (1956). Representation of events in nerve nets and finite automata. In Shannon, C. and McCarthy, J. (Eds.), *Automata Studies*, pp. 3–41. Princeton University Press, Princeton, NJ.

Klein, D. (2005). *The unsupervised learning of Natural Language Structure*. Ph.D. thesis, Stanford University.

Klein, D. and Manning, C. D. (2001). Parsing and hypergraphs. In *The Seventh Internation Workshop on Parsing Technologies*.

Klein, D. and Manning, C. D. (2002). A generative constituent-context model for improved grammar induction. In *ACL-02*.

Klein, D. and Manning, C. D. (2003a). A* parsing: Fast exact Viterbi parse selection. In *HLT-NAACL-03*.

Klein, D. and Manning, C. D. (2003b). Accurate unlexicalized parsing. In *HLT-NAACL-03*.

Klein, D. and Manning, C. D. (2003c). Fast exact inference with a factored model for natural language parsing. In Becker, S., Thrun, S., and Obermayer, K. (Eds.), *Advances in Neural Information Processing Systems 15*. MIT Press.

Klein, D. and Manning, C. D. (2004). Corpus-based induction of syntactic structure: Models of dependency and constituency. In *ACL-04*.

Klein, S. and Simmons, R. F. (1963). A computational approach to grammatical coding of English words. *Journal of the Association for Computing Machinery*, *10*(3), 334–347.

Klimt, B. and Yang, Y. (2004). The Enron corpus: A new dataset for email classification research. In *Proceedings of the European Conference on Machine Learning*, pp. 217–226. Springer.

Klovstad, J. W. and Mondshein, L. F. (1975). The CASPERS linguistic analysis system. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *ASSP-23*(1), 118–123.

Kneser, R. (1996). Statistical language modeling using a variable context length. In *ICSLP-96*, Philadelphia, PA, Vol. 1, pp. 494–497.

Kneser, R. and Ney, H. (1993). Improved clustering techniques for class-based statistical language modelling. In *EUROSPEECH-93*, pp. 973–976.

Kneser, R. and Ney, H. (1995). Improved backing-off for m-gram language modeling. In *IEEE ICASSP-95*, Vol. 1, pp. 181–184.

Knight, K. (1989). Unification: A multidisciplinary survey. *ACM Computing Surveys*, *21*(1), 93–124.

Knight, K. (1999a). Decoding complexity in word-replacement translation models. *Computational Linguistics*, *25*(4), 607–615.

Knight, K. (1999b). A statistical MT tutorial workbook. Manuscript prepared for the 1999 JHU Summer Workshop.

Knight, K. and Marcu, D. (2000). Statistics-based summarization - step one: Sentence compression. In *AAAI-00*, pp. 703–710.

Knill, K. and Young, S. J. (1997). Hidden Markov Models in speech and language processing. In Young, S. J. and Bloothooft, G. (Eds.), *Corpus-based Methods in Language and Speech Processing*, pp. 27–68. Kluwer, Dordrecht.

Knott, A. and Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, *18*(1), 35–62.

Knuth, D. E. (1968). Semantics of context-free languages. *Mathematical Systems Theory*, *2*(2), 127–145.

Knuth, D. E. (1973). *Sorting and Searching: The Art of Computer Programming Volume 3*. Addison-Wesley, Reading, MA.

Koehn, P. and Hoang, H. (2007). Factored translation models. In *EMNLP/CoNLL 2007*.

Koehn, P. (2003a). *Noun Phrase Translation*. Ph.D. thesis, University of Southern California.

Koehn, P. (2003b). Pharaoh: a beam search decoder for phrase-based statistical machine translation models: User manual and description..

Koehn, P. (2004). Pharaoh: a beam search decoder for phrase-based statistical machine translation models. In *Proceedings of AMTA 2004*.

Koehn, P., Abney, S. P., Hirschberg, J., and Collins, M. (2000). Improving intonational phrasing with syntactic information. In *IEEE ICASSP-00*.

Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., and Herbst, E. (2006). Moses: Open Source Toolkit for Statistical Machine Translation. In *ACL-07*, Prague. Demonstration session.

Koehn, P., Och, F. J., and Marcu, D. (2003). Statistical phrase-based translation. In *HLT-NAACL-03*, pp. 48–54.

Koenig, W., Dunn, H. K., Y., L., and Lacy (1946). The sound spectrograph. *Journal of the Acoustical Society of America*, *18*, 19–49.

Koerner, E. F. K. and Asher, R. E. (Eds.). (1995). *Concise History of the Language Sciences*. Elsevier Science, Oxford.

Kogan, Y., Collier, N., Pakhomov, S., and Krauthammer, M. (2005). Towards semantic role labeling & IE in the medical literature. In *AMIA 2005 Symposium Proceedings*, pp. 410–414.

Kogure, K. (1990). Strategic lazy incremental copy graph unification. In *COLING-90*, Helsinki, pp. 223–228.

Kompe, R., Kießling, A., Kuhn, T., Mast, M., Niemann, H., Nöth, E., Ott, K., and Batliner, A. (1993). Prosody takes over: A prosodically guided dialog system. In *EUROSPEECH-93*, Berlin, Vol. 3, pp. 2003–2006.

Koppel, M., Argamon, S., and Shimoni, A. R. (2002). Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, *17*(4), 401–412.

Kornai, A. (1991). *Formal Phonology*. Ph.D. thesis, Stanford University, Stanford, CA.

Koskenniemi, K. (1983). Two-level morphology: A general computational model of word-form recognition and production. Tech. rep. Publication No. 11, Department of General Linguistics, University of Helsinki.

Koskenniemi, K. and Church, K. W. (1988). Complexity, two-level morphology, and Finnish. In *COLING-88*, Budapest, pp. 335–339.

Kozima, H. (1993). Text segmentation based on similarity between words. In *ACL-93*, pp. 286–288.

Krieger, H.-U. and Nerbonne, J. (1993). Feature-based inheritance networks for computational lexicons. In Briscoe, T., de Paiva, V., and Copestake, A. (Eds.), *Inheritance, Defaults, and the Lexicon*, pp. 90–136. Cambridge University Press.

Krieger, H.-U. and Schäfer, U. (1994). TDL — A type description language for HPSG. Part 1: Overview. Tech. rep. RR-94-37, DFKI, Saarbrücken.

Krovetz, R. (1993). Viewing morphology as an inference process. In *SIGIR-93*, pp. 191–202. ACM.

Krovetz, R. (1998). More than one sense per discourse. In *Proceedings of the ACL-SIGLEX SENSEVAL Workshop*.

Krovetz, R. and Croft, W. B. (1992). Lexical ambiguity and information retrieval. *ACM Transactions on Information Systems*, *10*(2), 115–141.

Kruskal, J. B. (1983). An overview of sequence comparison. In Sankoff, D. and Kruskal, J. B. (Eds.), *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, pp. 1–44. Addison-Wesley, Reading, MA.

Kudo, T. and Matsumoto, Y. (2002). Japanese dependency analysis using cascaded chunking. In *CoNLL-02*, pp. 63–69.

Kuhl, P. K., F.-M., T., and Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*, 9096–9101.

Kuhn, R. and De Mori, R. (1990). A cache-based natural language model for speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(6), 570–583.

Kukich, K. (1992). Techniques for automatically correcting words in text. *ACM Computing Surveys*, *24*(4), 377–439.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, *22*, 79–86.

Kumar, S. and Byrne, W. (2002). Risk based lattice cutting for segmental minimum Bayes-risk decoding. In *ICSLP-02*, Denver, CO.

Kumar, S. and Byrne, W. (2005). Local phrase reordering models for statistical machine translation. In *HLT-EMNLP-05*.

Kuno, S. (1965). The predictive analyzer and a path elimination technique. *Communications of the ACM*, *8*(7), 453–462.

Kuno, S. and Oettinger, A. G. (1963). Multiple-path syntactic analyzer. In Popplewell, C. M. (Ed.), *Information Processing 1962: Proceedings of the IFIP Congress 1962*, Munich, pp. 306–312. North-Holland. Reprinted in Grosz et al. (1986).

Kupiec, J. (1992). Robust part-of-speech tagging using a hidden Markov model. *Computer Speech and Language*, *6*, 225–242.

Kupiec, J., Pedersen, J., and Chen, F. (1995). A trainable document summarizer. In *SIGIR 1995*, pp. 68–73.

Kučera, H. (1992). The mathematics of language. In *The American Heritage Dictionary of the English Language*, pp. xxxi–xxxiii. Houghton Mifflin, Boston.

Kučera, H. and Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press, Providence, RI.

Labov, W. (1966). *The Social Stratification of English in New York City*. Center for Applied Linguistics, Washington, D.C.

Labov, W. (1972). The internal evolution of linguistic rules. In Stockwell, R. P. and Macaulay, R. K. S. (Eds.), *Linguistic Change and Generative Theory*, pp. 101–171. Indiana University Press, Bloomington.

Labov, W. (1975). *The quantitative study of linguistic structure*. Pennsylvania Working Papers on Linguistic Change and Variation v.1 no. 3. U.S. Regional Survey, Philadelphia, PA.

Labov, W. (1994). *Principles of Linguistic Change: Internal Factors*. Blackwell, Oxford.

Labov, W. and Fanshel, D. (1977). *Therapeutic Discourse*. Academic Press.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge Studies in Linguistics. Cambridge University Press.

Ladefoged, P. (1993). *A Course in Phonetics*. Harcourt Brace Jovanovich. Third Edition.

Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. University of Chicago, Chicago, IL. Second Edition.

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

Lafferty, J. D., McCallum, A., and Pereira, F. C. N. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *ICML 2001*, Stanford, CA.

Lakoff, G. (1965). *On the Nature of Syntactic Irregularity*. Ph.D. thesis, Indiana University. Published as *Irregularity in Syntax*. Holt, Rinehart, and Winston, New York, 1970.

Lakoff, G. (1972). Linguistics and natural logic. In Davidson, D. and Harman, G. (Eds.), *Semantics for Natural Language*, pp. 545–665. D. Reidel, Dordrecht.

Lakoff, G. (1993). Cognitive phonology. In Goldsmith, J. (Ed.), *The Last Phonological Rule*, pp. 117–145. University of Chicago Press, Chicago.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press, Chicago, IL.

Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, *65*(3), 154–170.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge University Press.

Landauer, T. K. (Ed.). (1995). *The Trouble With Computers: Usefulness, Usability, and Productivity*. MIT Press.

Landauer, T. K. and Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Landauer, T. K., Laham, D., Rehder, B., and Schreiner, M. E. (1997). How well can passage meaning be derived without using word order: A comparison of latent semantic analysis and humans. In *COGSCI-97*, Stanford, CA, pp. 412–417. Lawrence Erlbaum.

Landes, S., Leacock, C., and Tengi, R. I. (1998). Building semantic concordances. In Fellbaum, C. (Ed.), *WordNet: An Electronic Lexical Database*, pp. 199–216. MIT Press.

Langendoen, D. T. (1975). Finite-state parsing of phrase-structure languages and the status of readjustment rules in the grammar. *Linguistic Inquiry*, *6*(4), 533–554.

Langendoen, D. T. and Langsam, Y. (1987). On the design of finite transducers for parsing phrase-structure languages. In Manaster-Ramer, A. (Ed.), *Mathematics of Language*, pp. 191–235. John Benjamins, Amsterdam.

Lapata, M. (2003). Probabilistic text structuring: experiments with sentence ordering. In *ACL-03*, Sapporo, Japan, pp. 545–552.

Lapata, M. (2006). Automatic evaluation of information ordering. *Computational Linguistics*, *32*(4), 471–484.

Lappin, S. and Leass, H. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics*, *20*(4), 535–561.

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

Lascarides, A. and Asher, N. (1993). Temporal interpretation, discourse relations, and common sense entailment. *Linguistics and Philosophy*, *16*(5), 437–493.

Lascarides, A. and Copestake, A. (1997). Default representation in constraint-based frameworks. *Computational Linguistics*, *25*(1), 55–106.

Lauer, M. (1995). Corpus statistics meet the noun compound. In *ACL-95*, Cambridge, MA, pp. 47–54.

Lawrence, W. (1953). The synthesis of speech from signals which have a low information rate.. In Jackson, W. (Ed.), *Communication Theory*, pp. 460–469. Butterworth.

LDC (1993). *LDC Catalog: CSR-I (WSJ0) Complete*. University of Pennsylvania. www.ldc.upenn.edu/Catalog/LDC93S6A.html.

LDC (1995). COMLEX English Pronunciation Dictionary Version 0.2 (COMLEX 0.2). Linguistic Data Consortium.

LDC (1998). *LDC Catalog: Hub4 project*. University of Pennsylvania. www.ldc.upenn.edu/Catalog/LDC98S71.html or www.ldc.upenn.edu/Catalog/Hub4.html.

Leacock, C. and Chodorow, M. S. (1998). Combining local context and WordNet similarity for word sense identification. In Fellbaum, C. (Ed.), *Wordnet: An Electronic Lexical Database*, pp. 265–283. MIT Press.

Leacock, C., Towell, G., and Voorhees, E. M. (1993). Corpus-based statistical sense resolution. In *Proceedings of the ARPA Human Language Technology Workshop*, pp. 260–265.

Lease, M., Johnson, M., and Charniak, E. (2006). Recognizing disfluencies in conversational speech. *IEEE Transactions on Audio, Speech and Language Processing*, *14*(5), 1566–1573.

Lee, L. and Rose, R. C. (1996). Speaker normalisation using efficient frequency warping procedures. In *ICASSP96*, pp. 353–356.

Lee, L. (1999). Measures of distributional similarity. In *ACL-99*, pp. 25–32.

Lee, L. (2001). On the effectiveness of the skew divergence for statistical language analysis. In *Artificial Intelligence and Statistics*, pp. 65–72.

Lee, S., Bresch, E., Adams, J., Kazemzadeh, A., and Narayanan, S. S. (2006). A study of emotional speech articulation using a fast magnetic resonance imaging technique. In *ICSLP-06*.

Leech, G., Garside, R., and Bryant, M. (1994). CLAWS4: The tagging of the British National Corpus. In *COLING-94*, Kyoto, pp. 622–628.

Lees, R. (1970). Problems in the grammatical analysis of English nominal compounds. In Bierwitsch, M. and Heidolph, K. E. (Eds.), *Progress in Linguistics*, pp. 174–187. Mouton, The Hague.

Leggetter, C. J. and Woodland, P. C. (1995). Maximum likelihood linear regression for speaker adaptation of *HMM*s. *Computer Speech and Language,*, *9*(2), 171–186.

Lehiste, I. (Ed.). (1967). *Readings in Acoustic Phonetics*. MIT Press.

Lehnert, W. G. (1977). A conceptual theory of question answering. In *IJCAI-77*, pp. 158–164. Morgan Kaufmann.

Lehnert, W. G., Cardie, C., Fisher, D., Riloff, E., and Williams, R. (1991). Description of the CIRCUS system as used for MUC-3. In Sundheim, B. (Ed.), *Proceedings of the Third Message Understanding Conference*, pp. 223–233.

Lehrer, A. (1974). *Semantic Fields and Lexical Structure*. North-Holland, Amsterdam.

Lehrer, A. and Kittay, E. (Eds.). (1992). *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization*. Lawrence Erlbaum.

Lemon, O., Georgila, K., Henderson, J., and Stuttle, M. (2006). An ISU dialogue system exhibiting reinforcement learning of dialogue policies: generic slot-filling in the TALK in-car system. In *EACL-06*.

Lerner, A. J. (1978). *The Street Where I Live*. Da Capo Press, New York.

Lesk, M. E. (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In *Proceedings of the Fifth International Conference on Systems Documentation*, Toronto, CA, pp. 24–26. ACM.

Levelt, W. J. M. (1970). A scaling approach to the study of syntactic relations. In d'Arcais, G. B. F. and Levelt, W. J. M. (Eds.), *Advances in psycholinguistics*, pp. 109–121. North-Holland, Amsterdam.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory*, *10*(8), 707–710. Original in *Doklady Akademii Nauk SSSR* 163(4): 845–848 (1965).

Levesque, H. J., Cohen, P. R., and Nunes, J. H. T. (1990). On acting together. In *AAAI-90*, Boston, MA, pp. 94–99. Morgan Kaufmann.

Levi, J. (1978). *The Syntax and Semantics of Complex Nominals*. Academic Press.

Levin, B. (1993). *English Verb Classes And Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago.

Levin, B. and Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge University Press.

Levin, E., Pieraccini, R., and Eckert, W. (2000). A stochastic model of human-machine interaction for learning dialog strategies. *IEEE Transactions on Speech and Audio Processing*, *8*, 11–23.

Levin, L., Gates, D., Lavie, A., and Waibel, A. (1998). An interlingua based on domain actions for machine translation of task-oriented dialogues. In *ICSLP-98*, Sydney, pp. 1155–1158.

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Levinson, S. E. (1995). Structural methods in automatic speech recognition. *Proceedings of the IEEE*, *73*(11), 1625–1650.

Levitt, S. D. and Dubner, S. J. (2005). *Freakonomics*. Morrow.

Levow, G.-A. (1998). Characterizing and recognizing spoken corrections in human-computer dialogue. In *COLING-ACL*, pp. 736–742.

Levy, R. (2007). Expectation-based syntactic comprehension. *Cognition*. In press.

Lewin, I., Becket, R., Boye, J., Carter, D., Rayner, M., and Wirén, M. (1999). Language processing for spoken dialogue systems: is shallow parsing enough?. In *Accessing Information in Spoken Audio: Proceedings of ESCA ETRW Workshop, Cambridge, 19 & 20th April 1999*, pp. 37–42.

Lewis, H. and Papadimitriou, C. (1988). *Elements of the Theory of Computation*. Prentice-Hall. Second edition.

Li, A., Zheng, F., Byrne, W., Fung, P., Kamm, T., Yi, L., Song, Z., Ruhi, U., Venkataramani, V., and Chen, X. (2000). CASS: A phonetically transcribed corpus of Mandarin spontaneous speech. In *ICSLP-00*, Beijing, China, pp. 485–488.

Li, C. N. and Thompson, S. A. (1981). *Mandarin Chinese: A Functional Reference Grammar*. University of California Press.

Li, H. and Abe, N. (1998). Generalizing case frames using a thesaurus and the MDL principle. *Computational Linguistics*, *24*(2), 217–244.

Li, X. and Roth, D. (2002). Learning question classifiers. In *COLING-02*, pp. 556–562.

Li, X. and Roth, D. (2005). Learning question classifiers: The role of semantic information. *Journal of Natural Language Engineering*, *11*(4).

Liberman, A. M., Delattre, P. C., and Cooper, F. S. (1952). The role of selected stimulus variables in the perception of the unvoiced stop consonants. *American Journal of Psychology*, *65*, 497–516.

Liberman, M. Y. and Church, K. W. (1992). Text analysis and word pronunciation in text-to-speech synthesis. In Furui, S. and Sondhi, M. M. (Eds.), *Advances in Speech Signal Processing*, pp. 791–832. Marcel Dekker.

Liberman, M. Y. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, *8*, 249–336.

Liberman, M. Y. and Sproat, R. (1992). The stress and structure of modified noun phrases in English. In Sag, I. A. and Szabolcsi, A. (Eds.), *Lexical Matters*, pp. 131–181. CSLI, Stanford University.

Lidstone, G. J. (1920). Note on the general case of the Bayes-Laplace formula for inductive or a posteriori probabilities. *Transactions of the Faculty of Actuaries*, *8*, 182–192.

Light, M. and Greiff, W. (2002). Statistical models for the induction and use of selectional preferences. *Cognitive Science*, *87*, 1–13.

Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. In *ACL 2004 Workshop on Text Summarization Branches Out*.

Lin, C.-Y. and Hovy, E. (2000). The automated acquisition of topic signatures for text summarization. In *COLING-00*, pp. 495–501.

Lin, C.-Y. and Hovy, E. H. (2003). Automatic evaluation of summaries using N-gram co-occurrence statistics. In *HLT-NAACL-03*, Edmonton, Canada.

Lin, D. (1995). A dependency-based method for evaluating broad-coverage parsers. In *IJCAI-95*, Montreal, pp. 1420–1425.

Lin, D. (1998a). Automatic retrieval and clustering of similar words. In *COLING/ACL-98*, Montreal, pp. 768–774.

Lin, D. (1998b). An information-theoretic definition of similarity. In *ICML 1998*, San Francisco, pp. 296–304.

Lin, D. (2007). Dependency-based word similarity demo. http://www.cs.ualberta.ca/~lindek/demos.htm.

Lin, D. and Pantel, P. (2001). Discovery of inference rules for question-answering. *Natural Language Engineering*, *7*(4), 343–360.

Lin, D. and Pantel, P. (2002). Concept discovery from text. In *COLING-02*, pp. 1–7.

Lin, D., Zhao, S., Qin, L., and Zhou, M. (2003). Identifying synonyms among distributionally similar words. In *IJCAI-03*, pp. 1492–1493.

Lin, J. and Demner-Fushman, D. (2005). Evaluating summaries and answers: Two sides of the same coin?. In *ACL 2005 Workshop on Measures for MT and Summarization*.

Lin, J. (2007). An exploration of the principles underlying redundancy-based factoid question answering. *ACM Transactions on Information Systems*.

Lindblom, B. E. F. (1990). Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, W. J. and Marchal, A. (Eds.), *Speech Production and Speech Modelling*, pp. 403–439. Kluwer.

Lindsey, R. (1963). Inferential memory as the basis of machines which understand natural language. In Feigenbaum, E. and Feldman, J. (Eds.), *Computers and Thought*, pp. 217–233. McGraw Hill.

Litman, D. J. (1985). *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues*. Ph.D. thesis, University of Rochester, Rochester, NY.

Litman, D. J. and Allen, J. (1987). A plan recognition model for subdialogues in conversation. *Cognitive Science*, *11*, 163–200.

Litman, D. J. and Pan, S. (2002). Designing and evaluating an adaptive spoken dialogue system. *User Modeling and User-Adapted Interaction*, *12*(2-3), 111–137.

Litman, D. J. and Silliman, S. (2004). Itspoke: An intelligent tutoring spoken dialogue system. In *HLT-NAACL-04*.

Litman, D. J., Swerts, M., and Hirschberg, J. (2000). Predicting automatic speech recognition performance using prosodic cues. In *NAACL 2000*.

Litman, D. J., Walker, M. A., and Kearns, M. (1999). Automatic detection of poor speech recognition at the dialogue level. In *ACL-99*, College Park, MA, pp. 309–316.

Liu, Y., Chawla, N. V., Harper, M. P., Shriberg, E., and Stolcke, A. (2006a). A study in machine learning from imbalanced data for sentence boundary detection in speech. *Computer Speech & Language*, *20*(4), 468–494.

Liu, Y., Shriberg, E., Stolcke, A., Hillard, D., Ostendorf, M., and Harper, M. (2006b). Enriching speech recognition with automatic detection of sentence boundaries and disfluencies. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(5), 1526–1540.

Liu, Y., Shriberg, E., Stolcke, A., Peskin, B., Ang, J., Hillard, D., Ostendorf, M., Tomalin, M., Woodland, P. C., and Harper, M. P. (2005). Structural metadata research in the ears program. In *IEEE ICASSP-05*.

Liu, Y. (2004). Word fragment identification using acoustic-prosodic features in conversational speech. In *HLT-NAACL-03 student research workshop*, pp. 37–42.

Liu, Y., Liu, Q., and Lin, S. (2005). Log-linear models for word alignment. In *ACL-05*, pp. 459–466.

Livescu, K., Glass, J., and Bilmes, J. (2003). Hidden feature modeling for speech recognition using dynamic bayesian networks. In *EUROSPEECH-03*.

Livescu, K. (2005). *Feature-Based Pronuncaition Modeling for Automatic Speech Recognition*. Ph.D. thesis, Massachusetts Institute of Technology.

Livescu, K. and Glass, J. (2004a). Feature-based pronunciation modeling for speech recognition. In *HLT-NAACL-04*, Boston, MA.

Livescu, K. and Glass, J. (2004b). Feature-based pronunciation modeling with trainable asynchrony probabilities. In *ICSLP-04*, Jeju, South Korea.

Lochbaum, K. E. (1998). A collaborative planning model of intentional structure. *Computational Linguistics*, *24*(4), 525–572.

Lochbaum, K. E., Grosz, B. J., and Sidner, C. L. (1990). Models of plans to support communication: An initial report. In *AAAI-90*, Boston, MA, pp. 485–490. Morgan Kaufmann.

Lochbaum, K. E., Grosz, B. J., and Sidner, C. L. (2000). Discourse structure and intention recognition. In Dale, R., Somers, H. L., and Moisl, H. (Eds.), *Handbook of Natural Language Processing*. Marcel Dekker.

Longacre, R. E. (1983). *The Grammar of Discourse*. Plenum Press.

Lowe, J. B., Baker, C. F., and Fillmore, C. J. (1997). A frame-semantic approach to semantic annotation. In *Proceedings of ACL SIGLEX Workshop on Tagging Text with Lexical Semantics*, Washington, D.C., pp. 18–24.

Lowerre, B. T. (1968). *The Harpy Speech Recognition System*. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.

Lu, Z., Cohen, B. K., and Hunter, L. (2006). Finding GeneRIFs via Gene Ontology annotations.. In *PSB 2006*, pp. 52–63.

Lucassen, J. and Mercer, R. L. (1984). An information theoretic approach to the automatic determination of phonemic baseforms. In *IEEE ICASSP-84*, Vol. 9, pp. 304–307.

Luce, P. A., Pisoni, D. B., and Goldfinger, S. D. (1990). Similarity neighborhoods of spoken words. In Altmann, G. T. M. (Ed.), *Cognitive Models of Speech Processing*, pp. 122–147. MIT Press.

Luhn, H. P. (1957). A statistical approach to the mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, *1*(4), 309–317.

Luhn, H. P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, *2*(2), 159–165.

Luk, R. W. P. and Damper, R. I. (1996). Stochastic phonographic transduction for english. *Computer Speech and Language*, *10*(2), 133–153.

Lyons, J. (1977). *Semantics*. Cambridge University Press.

Lyons, R. G. (2004). *Understanding Digital Signal Processing*. Prentice Hall. 2nd. ed.

Macleod, C., Grishman, R., and Meyers, A. (1998). COMLEX Syntax Reference Manual Version 3.0. Linguistic Data Consortium.

MacWhinney, B. (1977). Starting points. *Language*, *53*, 152–168.

MacWhinney, B. (1982). Basic syntactic processes. In Kuczaj, S. (Ed.), *Language Acquisition: Volume 1, Syntax and Semantics*, pp. 73–136. Lawrence Erlbaum.

MacWhinney, B. and Csaba Pléh (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, *29*, 95–141.

Madhu, S. and Lytel, D. (1965). A figure of merit technique for the resolution of non-grammatical ambiguity. *Mechanical Translation*, *8*(2), 9–13.

Magerman, D. M. (1995a). Statistical decision-tree models for parsing. In *ACL-95*, pp. 276–283.

Magerman, D. M. (1995b). Statistical decision-tree models for parsing. In *ACL-95*.

Magerman, D. M. and Marcus, M. P. (1991). Pearl: A probabilistic chart parser. In *EACL-91*, Berlin, Germany.

Main, M. G. and Benson, D. B. (1983). Denotational semantics for natural language question-answering programs. *American Journal of Computational Linguistics*, *9*(1), 11–21.

Makhoul, J., Baron, A., Bulyko, I., Nguyen, L., Ramshaw, L. A., Stallard, D., Schwartz, R., and Xiang, B. (2005). The effects of speech recognition and punctuation on information extraction performance. In *INTERSPEECH-05*, Lisbon, Portugal, pp. 57–60.

Makkai, A. (1972). *Idiom Structure in English*. Mouton, The Hague.

Malouf, R., Carroll, J., and Copestake, A. (2000). Efficient feature structure operations without compilation. *Natural Language Engineering*, *6*(1).

Malouf, R. (2002). A comparison of algorithms for maximum entropy parameter estimation. In *CoNLL-2002*, pp. 49–55.

Mangu, L. and Brill, E. (1997). Automatic rule acquisition for spelling correction. In *ICML 1997*, Nashville, TN, pp. 187–194.

Mangu, L., Brill, E., and Stolcke, A. (2000). Finding consensus in speech recognition: Word error minimization and other applications of confusion networks. *Computer Speech and Language*, *14*(4), 373–400.

Mani, I. (2001). *Automatic Summarization*. John Benjamins.

Mani, I. and Bloedorn, E. (1999). Summarizing similarities and differences among related documents. *Information Retrieval*, *1*(1-2), 35–67.

Mani, I., Gates, B., and Bloedorn, E. (1999). Improving summaries by revising them. In *ACL-99*, pp. 558–565.

Mani, I. and Maybury, M. T. (1999). *Advances in Automatic Text Summarization*. MIT Press.

Mann, W. C. and Thompson, S. A. (1987). Rhetorical structure theory: A theory of text organization. Tech. rep. RS-87-190, Information Sciences Institute.

Manning, C. D. (1998). Rethinking text segmentation models: An information extraction case study. Tech. rep. SULTRY-98-07-01, University of Sydney.

Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Manning, C. D. and Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.

Marchand, Y. and Damper, R. I. (2000). A multi-strategy approach to improving pronunciation by analogy. *Computational Linguistics*, *26*(2), 195–219.

Marcu, D. (1995). Discourse trees are good indicators of importance in text. In Mani, I. and Maybury, M. T. (Eds.), *Advances in Automatic Text Summarization*, Cambridge, MA, pp. 123–136. MIT Press.

Marcu, D. (1999). The automatic construction of large-scale corpora for summarization research. In *SIGIR 1999*, Berkeley, CA, pp. 137–144.

Marcu, D. (2000a). The rhetorical parsing of unrestricted texts: A surface-based approach. *Computational Linguistics*, *26*(3), 395–448.

Marcu, D. (Ed.). (2000b). *The Theory and Practice of Discourse Parsing and Summarization*. MIT Press.

Marcu, D. and Echihabi, A. (2002). An unsupervised approach to recognizing discourse relations. In *ACL-02*, pp. 368–375.

Marcu, D. and Wong, W. (2002). A phrase-based, joint probability model for statistical machine translation. In *EMNLP 2002*, pp. 133–139.

Marcus, M. P. (1990). Summary of session 9: Automatic acquisition of linguistic structure. In *Proceedings DARPA Speech and Natural Language Workshop*, Hidden Valley, PA, pp. 249–250. Morgan Kaufmann.

Marcus, M. P., Kim, G., Marcinkiewicz, M. A., MacIntyre, R., Bies, A., Ferguson, M., Katz, K., and Schasberger, B. (1994). The Penn Treebank: Annotating predicate argument structure. In *ARPA Human Language Technology Workshop*, Plainsboro, NJ, pp. 114–119. Morgan Kaufmann.

Marcus, M. P., Santorini, B., and Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn treebank. *Computational Linguistics*, *19*(2), 313–330.

Markov, A. A. (1913). Essai d'une recherche statistique sur le texte du roman "Eugene Onegin" illustrant la liaison des epreuve en chain ('Example of a statistical investigation of the text of "Eugene Onegin" illustrating the dependence between samples in chain'). *Izvistia Imperatorskoi Akademii Nauk (Bulletin de l'Académie Impériale des Sciences de St.-Pétersbourg)*, *7*, 153–162. English translation by Morris Halle, 1956.

Markov, A. A. (2006). Classical text in translation: A. A. Markov, an example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains. *Science in Context*, *19*(4), 591–600. Translated by David Link.

Marshall, I. (1983). Choice of grammatical word-class without GLobal syntactic analysis: Tagging words in the LOB corpus. *Computers and the Humanities*, *17*, 139–150.

Marshall, I. (1987). Tag selection using probabilistic methods. In Garside, R., Leech, G., and Sampson, G. (Eds.), *The Computational Analysis of English*, pp. 42–56. Longman, London.

Marslen-Wilson, W. and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*, 522–523.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., and Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*(1), 3–33.

Martin, J. H. (1986). The acquisition of polysemy. In *ICML 1986*, Irvine, CA, pp. 198–204.

Martin, J. H. (1990). *A Computational Model of Metaphor Interpretation*. Perspectives in Artificial Intelligence. Academic Press.

Martin, J. H. (1996). Computational approaches to figurative language. *Metaphor and Symbolic Activity*, *11*(1), 85–100.

Martin, J. H. (2006). A rational analysis of the context effect on metaphor processing. In Gries, S. T. and Stefanowitsch, A. (Eds.), *Corpus-Based Approaches to Metaphor and Metonymy*. Mouton de Gruyter.

Massaro, D. W. (1998). *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*. MIT Press.

Massaro, D. W. and Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 753–771.

Masterman, M. (1957). The thesaurus in syntax and semantics. *Mechanical Translation*, *4*(1), 1–2.

Matsuzaki, T., Miyao, Y., and ichi Tsujii, J. (2007). Efficient hpsg parsing with supertagging and cfg-filtering. In *IJCAI-07*.

Matthews, A. and Chodorow, M. S. (1988). Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location, and depth of embedding. *Journal of Memory and Language*, *27*, 245–260.

Mays, E., Damerau, F. J., and Mercer, R. L. (1991). Context based spelling correction. *Information Processing and Management*, *27*(5), 517–522.

McCallum, A., Freitag, D., and Pereira, F. C. N. (2000). Maximum Entropy Markov Models for Information Extraction and Segmentation. In *ICML 2000*, pp. 591–598.

McCallum, A. (2005). Information extraction: Distilling structured data from unstructured text. *ACM Queue*, 48–57.

McCallum, A., Freitag, D., and Pereira, F. C. N. (2000). Maximum entropy Markov models for information extraction and segmentation. In *ICML 2000*, pp. 591–598.

McCarthy, D. (2006). Relating WordNet senses for word sense disambiguation. In *Proceedings of ACL Workshop on Making Sense of Sense*.

McCarthy, D. and Carroll, J. (2003). Disambiguating nouns, verbs, and adjectives using automatically acquired selectional preferences. *Computational Linguistics*, *29*(4), 639–654.

McCarthy, D., Koeling, R., Weeds, J., and Carroll, J. (2004). Finding predominant word senses in untagged text. In *ACL-04*, pp. 279–286.

McCarthy, J. J. (1981). A prosodic theory of non-concatenative morphology. *Linguistic Inquiry*, *12*, 373–418.

McCarthy, J. F. and Lehnert, W. G. (1995). Using decision trees for coreference resolution. In *IJCAI-95*, Montreal, Canada, pp. 1050–1055.

McCawley, J. D. (1968). The role of semantics in a grammar. In Bach, E. W. and Harms, R. T. (Eds.), *Universals in Linguistic Theory*, pp. 124–169. Holt, Rinehart & Winston.

McCawley, J. D. (1978). Where you can shove infixes. In Bell, A. and Hooper, J. B. (Eds.), *Syllables and Segments*, pp. 213–221. North-Holland, Amsterdam.

McCawley, J. D. (1993). *Everything that Linguists have Always Wanted to Know about Logic* (2nd edition). University of Chicago Press, Chicago, IL.

McCawley, J. D. (1998). *The Syntactic Phenomena of English*. University of Chicago Press, Chicago.

McClelland, J. L. and Elman, J. L. (1986). Interactive processes in speech perception: The TRACE model. In McClelland, J. L., Rumelhart, D. E., and the PDP Research Group (Eds.), *Parallel Distributed Processing Volume 2: Psychological and Biological Models*, pp. 58–121. MIT Press.

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *5*, 115–133. Reprinted in *Neurocomputing: Foundations of Research, ed. by J. A. Anderson and E Rosenfeld. MIT Press 1988*.

McDermott, E. and Hazen, T. (2004). Minimum Classification Error training of landmark models for real-time continuous speech recognition. In *IEEE ICASSP-04*.

McDonald, D. B. (1982). *Understanding Noun Compounds*. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA. CMU Technical Report CS-82-102.

McDonald, R., Pereira, F. C. N., Ribarov, K., and Hajič, J. (2005). Non-projective dependency parsing using spanning tree algorithms. In *HLT-EMNLP-05*.

McDonald, R. (2006). Discriminative sentence compression with soft syntactic constraints. In *EACL-06*.

McDonald, R., Crammer, K., and Pereira, F. C. N. (2005). On-line large-margin training of dependency parsers. In *ACL-05*, Ann Arbor, pp. 91–98.

McGuiness, D. L. and van Harmelen, F. (2004). OWL web ontology overview. Tech. rep. 20040210, World Wide Web Consortium.

McGurk, H. and Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

McKeown, K. R., Passonneau, R., Elson, D., Nenkova, A., and Hirschberg, J. (2005). Do summaries help? A task-based evaluation of multi-document summarization. In *SIGIR 2005*, Salvador, Brazil.

McLuhan, M. (1964). *Understanding media: The extensions of man*. New American Library, New York.

McNamee, P. and Mayfield, J. (2002). Entity extraction without language-specific resources. In *Proceedings of the Conference on Computational Natural Language Learning (CoNLL-2002)*, Taipei, Taiwan.

McRoy, S. (1992). Using multiple knowledge sources for word sense discrimination. *Computational Linguistics*, *18*(1), 1–30.

McTear, M. F. (2002). Spoken dialogue technology: Enabling the conversational interface. *ACM Computing Surveys*, *34*(1), 90–169.

McTear, M. F. (2004). *Spoken Dialogue Technology*. Springer Verlag, London.

Mealy, G. H. (1955). A method for synthesizing sequential circuits. *Bell System Technical Journal*, *34*(5), 1045–1079.

Megyesi, B. (1999). Improving brill's pos tagger for an agglutinative language. In *EMNLP/VLC-99*, College Park, MA.

Melamed, I. D. (2003). Multitext grammars and synchronous parsers. In *HLT-NAACL-03*.

Mellish, C., Knott, A., Oberlander, J., , and O'Donnell, M. (1998). Experiments using stochastic search for text planning. In *INLG 1998*, pp. 98–107.

Mel'čuk, I. A. (1979). *Studies in dependency syntax*. Karoma Publishers, Ann Arbor.

Merialdo, B. (1994). Tagging English text with a probabilistic model. *Computational Linguistics*, *20*(2), 155–172.

Merlo, P. and Stevenson, S. (2001). Automatic verb classification based on statistical distribution of argument structure. *Computational Linguistics*, *27*(3), 373–408.

Merlo, P., Stevenson, S., Tsang, V., and Allaria, G. (2001). A multilingual paradigm for automatic verb classification. In *ACL-02*, pp. 207–214.

Merton, R. K. (1961). Singletons and multiples in scientific discovery. *American Philosophical Society Proceedings*, *105*(5), 470–486.

Meurers, W. D. and Minnen, G. (1997). A computational treatment of lexical rules in HPSG as covariation in lexical entries. *Computational Linguistics*, *23*(4), 543–568.

Mihalcea, R. and Moldovan, D. (2001). Automatic generation of a coarse grained WordNet. In *NAACL Workshop on WordNet and Other Lexical Resources*.

Mihalcea, R. and Moldovan, D. (1999). An automatic method for generating sense tagged corpora. In *Proceedings of AAAI*, pp. 461–466.

Mikheev, A. (2003). Text segmentation. In Mitkov, R. (Ed.), *Oxford Handbook of Computational Linguistics*. Oxford University Press, Oxford.

Mikheev, A., Moens, M., and Grover, C. (1999). Named entity recognition without gazetteers. In *EACL-99*, pp. 1–8.

Miller, C. A. (1998). Pronunciation modeling in speech synthesis. Tech. rep. IRCS 98–09, University of Pennsylvania Institute for Research in Cognitive Science, Philadephia, PA.

Miller, G. A. and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, *27*, 338–352.

Miller, G. A. and Beebe-Center, J. G. (1958). Some psychological methods for evaluating the quality of translations. *Mechanical Translation*, *3*, 73–80.

Miller, G. A. and Charles, W. G. (1991). Contextual correlates of semantics similarity. *Language and Cognitive Processes*, *6*(1), 1–28.

Miller, G. A. and Chomsky, N. (1963). Finitary models of language users. In Luce, R. D., Bush, R. R., and Galanter, E. (Eds.), *Handbook of Mathematical Psychology*, Vol. II, pp. 419–491. John Wiley.

Miller, G. A., Leacock, C., Tengi, R. I., and Bunker, R. T. (1993). A semantic concordance. In *Proceedings ARPA Workshop on Human Language Technology*, pp. 303–308.

Miller, G. A. and Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *American Journal of Psychology*, *63*, 176–185.

Miller, J. L. (1994). On the internal structure of phonetic categories: a progress report. *Cognition*, *50*, 271–275.

Miller, S., Bobrow, R. J., Ingria, R., and Schwartz, R. (1994). Hidden understanding models of natural language. In *ACL-94*, Las Cruces, NM, pp. 25–32.

Miller, S., Fox, H., Ramshaw, L. A., and Weischedel, R. (2000). A novel use of statistical parsing to extract information from text. In *NAACL 2000*, Seattle, WA, pp. 226–233.

Miller, S., Stallard, D., Bobrow, R. J., and Schwartz, R. (1996). A fully statistical approach to natural language interfaces. In *ACL-96*, Santa Cruz, CA, pp. 55–61.

Miltsakaki, E., Prasad, R., Joshi, A. K., and Webber, B. L. (2004a). Annotating discourse connectives and their arguments. In *Proceedings of the NAACL/HLT Workshop: Frontiers in Corpus Annotation*.

Miltsakaki, E., Prasad, R., Joshi, A. K., and Webber, B. L. (2004b). The Penn Discourse Treebank. In *LREC-04*.

Mishne, G. and de Rijke, M. (2006). MoodViews: Tools for blog mood analysis. In Nicolov, N., Salvetti, F., Liberman, M. Y., and Martin, J. H. (Eds.), *Computational Approaches to Analyzing Weblogs: Papers from the 2006 Spring Symposium*, Stanford, Ca. AAAI.

Mitkov, R. (2002). *Anaphora Resolution*. Longman.

Mitkov, R. and Boguraev, B. (Eds.). (1997). *Proceedings of the ACL-97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*, Madrid, Spain.

Miyao, Y. and Tsujii, J. (2005). Probabilistic disambiguation models for wide-coverage hpsg parsing. In *ACL-05*, pp. 83–90.

Mohammad, S. and Hirst, G. (2005). Distributional measures as proxies for semantic relatedness. Submitted.

Mohri, M. and Nederhof, M. J. (2001). Regular approximation of context-free grammars through transformation. In Junqua, J.-C. and van Noord, G. (Eds.), *Robustness in Language and Speech Technology*, pp. 153–163. Kluwer.

Mohri, M. (1996). On some applications of finite-state automata theory to natural language processing. *Natural Language Engineering*, *2*(1), 61–80.

Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, *23*(2), 269–312.

Mohri, M. (2000). Minimization algorithms for sequential transducers. *Theoretical Computer Science*, *234*, 177–201.

Mohri, M. and Sproat, R. (1998). On a common fallacy in computational linguistics. In Suominen, M., Arppe, A., Airola, A., Heinämäki, O., Miestamo, M., Määttä, U., Niemi, J., Pitkänen, K. K., and Sinnemäki, K. (Eds.), *A Man of Measure: Festschrift in Honour of Fred Karlsson on this 60th Birthday*, pp. 432–439. SKY Journal of Linguistics, Volume 19, 2006.

Möller, S. (2002). A new taxonomy for the quality of telephone services based on spoken dialogue systems. In *In Proceedings of the 3rd SIGdial Workshop on Discourse and Dialogue*, pp. 142–153.

Möller, S. (2004). *Quality of Telephone-Based Spoken Dialogue Systems*. Springer.

Montague, R. (1973). The proper treatment of quantification in ordinary English. In Thomason, R. (Ed.), *Formal Philosophy: Selected Papers of Richard Montague*, pp. 247–270. Yale University Press, New Haven, CT.

Monz, C. (2004). Minimal span weighting retrieval for question answering. In *SIGIR Workshop on Information Retrieval for Question Answering*, pp. 23–30.

Moore, E. F. (1956). Gedanken-experiments on sequential machines. In Shannon, C. and McCarthy, J. (Eds.), *Automata Studies*, pp. 129–153. Princeton University Press, Princeton, NJ.

Moore, R. C. (1977). Reasoning about knowledge and action. In *IJCAI-77*, pp. 223–227.

Moore, R. C. (2002). Fast and accurate sentence alignment of bilingual corpora. In *Machine Translation: From Research to Real Users (Proceedings, 5th Conference of the Association for Machine Translation in the Americas, Tiburon, California)*, pp. 135–244.

Moore, R. C. (2004). On log-likelihood-ratios and the significance of rare events. In *EMNLP 2004*, Barcelona, pp. 333–340.

Moore, R. C. (2005). A discriminative framework for bilingual word alignment. In *HLT-EMNLP-05*, pp. 81–88.

Morgan, A. A., Wellner, B., Colombe, J. B., Arens, R., Colosimo, M. E., and Hirschman, L. (2007). Evaluating human gene and protein mention normalization to unique identifiers. In *Pacific Symposium on Biocomputing*, pp. 281–291.

Morgan, N. and Fosler-Lussier, E. (1989). Combining multiple estimators of speaking rate. In *IEEE ICASSP-89*.

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

Morris, W. (Ed.). (1985). *American Heritage Dictionary* (2nd College Edition edition). Houghton Mifflin.

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., and Baayen, R. H. (2004a). Morphological family size in a morphologically rich language: The case of finnish compared to dutch and hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1271–1278.

Moscoso del Prado Martín, F., Kostic, A., and Baayen, R. H. (2004b). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, *94*(1), 1–18.

Moshier, M. A. (1988). *Extensions to Unification Grammar for the Description of Programming Languages*. Ph.D. thesis, University of Michigan, Ann Arbor, MI.

Mosteller, F. and Wallace, D. L. (1964). *Inference and Disputed Authorship: The Federalist*. Springer-Verlag. A second edition appeared in 1984 as *Applied Bayesian and Classical Inference*.

Müller, K. (2001). Automatic detection of syllable boundaries combining the advantages of treebank and bracketed corpora training. In *ACL-01*, Toulouse, France.

Müller, K. (2002). Probabilistic context-free grammars for phonology. In *Proceedings of ACL SIGPHON*, Philadelphia, PA, pp. 70–80.

Müller, K., Möbius, B., and Prescher, D. (2000). Inducing probabilistic syllable classes using multivariate clustering. In *ACL-00*, pp. 225–232.

Munoz, M., Punyakanok, V., Roth, D., and Zimak, D. (1999). A learning approach to shallow parsing. In *EMNLP/VLC-99*, College Park, MD, pp. 168–178.

Munteanu, C. and Penn, G. (2004). Optimizing typed feature structure grammar parsing through non-statistical indexing. In *ACL-04*, Barcelona, Spain, pp. 223–230.

Murata, T. (1989). Petri nets: Properties, analysis, and applications. *Proceedings of the IEEE*, *77*(4), 541–576.

Murveit, H., Butzberger, J. W., Digalakis, V. V., and Weintraub, M. (1993). Large-vocabulary dictation using SRI's decipher speech recognition system: Progressive-search techniques. In *IEEE ICASSP-93*, Vol. 2, pp. 319–322. IEEE.

Nádas, A. (1983). A decision theoretic formulation of a training problem in speech recognition and a comparison of training by unconditional versus conditional maximum likelihood. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *31*(4), 814–817.

Nádas, A. (1984). Estimation of probabilities in the language model of the IBM speech recognition system. *IEEE Transactions on Acoustics, Speech, Signal Processing*, *32*(4), 859–861.

Nagata, M. and Morimoto, T. (1994). First steps toward statistical modeling of dialogue to predict the speech act type of the next utterance. *Speech Communication*, *15*, 193–203.

Nakajima, S. and Hamada, H. (1988). Automatic generation of synthesis units based on context oriented clustering. In *IEEE ICASSP-88*, pp. 659–662.

Nakatani, C. and Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. *Journal of the Acoustical Society of America*, *95*(3), 1603–1616.

Nakov, P. I. and Hearst, M. A. (2003). Category-based pseudowords. In *HLT-NAACL-03*, Edmonton, Canada.

Nakov, P. I. and Hearst, M. A. (2005). A study of using search engine page hits as a proxy for n-gram frequencies. In *Proceedings of RANLP-05 (Recent Advances in Natural Language Processing)*, Borovets, Bulgaria.

Narayanan, S. S. and Alwan, A. (Eds.). (2004). *Text to Speech Synthesis: New paradigms and advances*. Prentice Hall.

Narayanan, S. (1997a). *Knowledge-based Action Representations for Metaphor and Aspect (KARMA)*. Ph.D. thesis, University of California, Berkeley.

Narayanan, S. (1997b). Talking the talk *is* like walking the walk: A computational model of verbal aspect. In *COGSCI-97*, Stanford, CA, pp. 548–553.

Narayanan, S., Fillmore, C. J., Baker, C. F., and Petruck, M. R. L. (1999). Framenet meets the semantic web: A DAML+OIL frame representation. In *AAAI-02*.

Naur, P., Backus, J. W., Bauer, F. L., Green, J., Katz, C., McCarthy, J., Perlis, A. J., Rutishauser, H., Samelson, K., Vauquois, B., Wegstein, J. H., van Wijngaarden, A., and Woodger, M. (1960). Report on the algorithmic language ALGOL 60. *Communications of the ACM*, *3*(5), 299–314. Revised in CACM 6:1, 1-17, 1963.

Navigli, R. (2006). Meaningful clustering of senses helps boost word sense disambiguation performance. In *COLING/ACL 2006*, pp. 105–112.

Nederhof, M. J. (2000). Practical experiments with regular approximation of context-free languages. *Computational Linguistics*, *26*(1), 17–44.

Needleman, S. B. and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino-acid sequence of two proteins. *Journal of Molecular Biology*, *48*, 443–453.

Nenkova, A. and Vanderwende, L. (2005). The Impact of Frequency on Summarization. Tech. rep. MSR-TR-2005-101, Microsoft Research, Redmond, WA.

Nenkova, A., Brenier, J. M., Kothari, A., Calhoun, S., Whitton, L., Beaver, D., and Jurafsky, D. (2007). To memorize or to predict: Prominence labeling in conversational speech. In *NAACL-HLT 07*.

Nenkova, A. and McKeown, K. R. (2003). References to named entities: a corpus study. In *HLT-NAACL-03*, pp. 70–72.

Nenkova, A., Passonneau, R., and McKeown, K. R. (2007). The pyramid method: Incorporating human content selection variation in summarization evaluation. *ACM TSLP*, *4*(2).

Nespor, M. and Vogel, I. (1986). *Prosodic phonology*. Foris, Dordrecht.

Nesson, R., Shieber, S. M., and Rush, A. (2006). Induction of probabilistic synchronous tree-insertion grammars for machine translation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas (AMTA 2006)*, Boston, MA.

Neu, H. (1980). Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In Labov, W. (Ed.), *Locating Language in Time and Space*, pp. 37–54. Academic Press.

Newell, A., Langer, S., and Hickey, M. (1998). The rôle of natural language processing in alternative and augmentative communication. *Natural Language Engineering*, *4*(1), 1–16.

Newman, S. (1944). *Yokuts Language of California*. Viking Fund Publications in Anthropology 2, New York.

Ney, H., Essen, U., and Kneser, R. (1994). On structuring probabilistic dependencies in stochastic language modelling. *Computer Speech and Language*, *8*, 1–38.

Ney, H., Haeb-Umbach, R., Tran, B.-H., and Oerder, M. (1992). Improvements in beam search for 10000-word continuous speech recognition. In *IEEE ICASSP-92*, San Francisco, CA, pp. I.9–12. IEEE.

Ney, H. (1991). Dynamic programming parsing for context-free grammars in continuous speech recognition. *IEEE Transactions on Signal Processing*, *39*(2), 336–340.

Ng, H. T. and Lee, H. B. (1996). Integrating multiple knowledge sources to disambiguate word senses: An exemplar-based approach. In *ACL-96*, Santa Cruz, CA, pp. 40–47. ACL.

Ng, H. T. and Zelle, J. (1997). Corpus-based approaches to semantic interpretation in NLP. *AI Magazine*, *18*(4), 45–64.

Ng, S.-K. (2006). Integrating text mining with data mining. In Ananiadou, S. and McNaught, J. (Eds.), *Text mining for biology and biomedicine*. Artech House Publishers.

Ng, V. (2004). Learning noun phrase anaphoricity to improve coreference resolution: Issues in representation and optimization. In *ACL-04*.

Ng, V. (2005). Machine learning for coreference resolution: From local classification to global ranking. In *ACL-05*.

Ng, V. and Cardie, C. (2002a). Identifying anaphoric and non-anaphoric noun phrases to improve coreference resolution. In *COLING-02*.

Ng, V. and Cardie, C. (2002b). Improving machine learning approaches to coreference resolution. In *ACL-02*.

Nguyen, L. and Schwartz, R. (1999). Single-tree method for grammar-directed search. In *IEEE ICASSP-99*, pp. 613–616. IEEE.

Nichols, J. (1986). Head-marking and dependent-marking grammar. *Language*, *62*(1), 56–119.

Nicolov, N. and Glance, N. (Eds.). (2007). *Proceedings of the First International Conference on Weblogs and Social Media (ICWSM)*, Boulder, CO.

Nicolov, N., Salvetti, F., Liberman, M. Y., and Martin, J. H. (Eds.). (2006). *Computational Approaches to Analyzing Weblogs: Papers from the 2006 Spring Symposium*, Stanford, Ca. AAAI.

Nida, E. A. (1975). *Componential Analysis of Meaning: An Introduction to Semantic Structures*. Mouton, The Hague.

Niekrasz, J. and Purver, M. (2006). A multimodal discourse ontology for meeting understanding. In Renals, S. and Bengio, S. (Eds.), *Machine Learning for Multimodal Interaction: Second International Workshop MLMI 2005, Revised Selected Papers*, No. 3689 in Lecture Notes in Computer Science, pp. 162–173. Springer-Verlag.

Nielsen, J. (1992). The usability engineering life cycle. *IEEE Computer*, *25*(3), 12–22.

Niesler, T. R., Whittaker, E. W. D., and Woodland, P. C. (1998). Comparison of part-of-speech and automatically derived category-based language models for speech recognition. In *IEEE ICASSP-98*, Vol. 1, pp. 177–180.

Niesler, T. R. and Woodland, P. C. (1996). A variable-length category-based n-gram language model. In *IEEE ICASSP-96*, Atlanta, GA, Vol. I, pp. 164–167. IEEE.

Niesler, T. R. and Woodland, P. C. (1999). Modelling word-pair relations in a category-based language model. In *IEEE ICASSP-99*, pp. 795–798. IEEE.

Nilsson, N. J. (1980). *Principles of Artificial Intelligence*. Morgan Kaufmann, Los Altos, CA.

Ninomiya, T., Tsuruoka, Y., Miyao, Y., Taura, K., and Tsujii, J. (2006). Fast and scalable hpsg parsing. *Traitement automatique des langues (TAL)*, *46*(2).

Nirenburg, S., Carbonell, J., Tomita, M., and Goodman, K. (1992). *Machine Translation: A Knowledge-based Approach*. Morgan Kaufmann.

Nirenburg, S., Somers, H. L., and Wilks, Y. (Eds.). (2002). *Readings in Machine Translation*. MIT Press.

Nissim, M., Dingare, S., Carletta, J., and Steedman, M. (2004). An annotation scheme for information status in dialogue. In *LREC-04*, Lisbon.

NIST (1990). TIMIT Acoustic-Phonetic Continuous Speech Corpus. National Institute of Standards and Technology Speech Disc 1-1.1. NIST Order No. PB91-505065.

NIST (2005). Speech recognition scoring toolkit (sctk) version 2.1. Available at http://www.nist.gov/speech/tools/.

NIST (2007a). The ace 2007 (ace07) evaluation plan: Evaluation of the detection and recognition of ace entities, values, temporal expressions, relations, and events. Unpublished.

NIST (2007b). Matched Pairs Sentence-Segment Word Error (MAPSSWE) Test. http://www.nist.gov/speech/tests/sigtests/mapsswe.htm.

Nivre, J. (2007). Incremental non-projective dependency parsing. In *NAACL-HLT 07*.

Nivre, J., Hall, J., and Nilsson, J. (2006). Maltparser: A data-driven parser-generator for dependency parsing. In *LREC-06*, pp. 2216–2219.

Niyogi, P., Burges, C., and Ramesh, P. (1998). Distinctive feature detection using support vector machines. In *IEEE ICASSP-98*.

Norman, D. A. (1988). *The Design of Everyday Things*. Basic Books.

Norman, D. A. and Rumelhart, D. E. (1975). *Explorations in Cognition*. Freeman, San Francisco, CA.

Normandin, Y. (1996). Maximum mutual information estimation of hidden Markov models. In Lee, C.-H., Soong, F. K., and Paliwal, K. K. (Eds.), *Automatic Speech and Speaker Recognition*, pp. 57–82. Kluwer.

Norvig, P. (1987). *A Unified Theory of Inference for Text Understanding*. Ph.D. thesis, University of California, Berkeley, CA. Available as University of California at Berkeley Computer Science Division Tech. rep. #87/339.

Norvig, P. (1991). Techniques for automatic memoization with applications to context-free parsing. *Computational Linguistics*, *17*(1), 91–98.

Norvig, P. (2005). The gettysburgh powerpoint presentation. http://norvig.com/Gettysburg/.

Norvig, P. (2007). How to write a spelling corrector. http://www.norvig.com/spell-correct.html.

Och, F. J. (2003). Minimum error rate training in statistical machine translation. In *ACL-03*, pp. 160–167.

Och, F. J. and Ney, H. (2002). Discriminative training and maximum entropy models for statistical machine translation. In *ACL-02*, pp. 295–302.

Och, F. J. and Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, *29*(1), 19–51.

Och, F. J. and Ney, H. (2004). The alignment template approach to statistical machine translation. *Computational Linguistics*, *30*(4), 417–449.

Och, F. J., Ueffing, N., and Ney, H. (2001). An efficient a* search algorithm for statistical machine translation. In *Proceedings of the ACL Workshop on Data-Driven methods in Machine Translation*, pp. 1–8.

Och, F. J. (1998). *Ein beispielsbasierter und statistischer Ansatz zum maschinellen Lernen von natürlichsprachlicher Übersetzung*. Ph.D. thesis, Universität Erlangen-Nürnberg, Germany. Diplomarbeit (diploma thesis).

Odden, D. (2005). *Introducing Phonology*. Cambridge University Press.

Odell, J. J. (1995). *The Use of Context in Large Vocabulary Speech Recognition*. Ph.D. thesis, Queen's College, University of Cambridge.

Odell, M. K. and Russell, R. C. (1918/1922). U.S. Patents 1261167 (1918), 1435663 (1922). Cited in Knuth (1973).

Oden, G. C. and Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review*, *85*, 172–191.

Oehrle, R. T., Bach, E., and Wheeler, D. W. (Eds.). (1988). *Categorial Grammars and Natural Language Structures*. D. Reidel, Dordrecht.

Oflazer, K. (1993). Two-level description of Turkish morphology. In *Proceedings, Sixth Conference of the European Chapter of the ACL*.

Ogburn, W. F. and Thomas, D. S. (1922). Are inventions inevitable? A note on social evolution. *Political Science Quarterly*, *37*, 83–98.

Ogren, P. V., Cohen, K. B., Acquaah-Mensah, G. K., Eberlein, J., and Hunter, L. (2004). The compositional structure of Gene Ontology terms. In *Pac Symp Biocomput*, pp. 214–225.

Olive, J. P. (1977). Rule synthesis of speech from dyadic units. In *ICASSP77*, pp. 568–570.

Olive, J. P. and Liberman, M. Y. (1979). A set of concatenative units for speech synthesis. *Journal of the Acoustical Society of America*, *65*, S130.

Olive, J. P., van Santen, J. P. H., Möbius, B., and Shih, C. (1998). Synthesis. In Sproat, R. (Ed.), *Multilingual Text-To-Speech Synthesis: The Bell Labs Approach*, pp. 191–228. Kluwer, Dordrecht.

Oncina, J., García, P., and Vidal, E. (1993). Learning subsequential transducers for pattern recognition tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*, 448–458.

Oppenheim, A. V., Schafer, R. W., and Stockham, T. G. J. (1968). Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE*, *56*(8), 1264–1291.

Oravecz, C. and Dienes, P. (2002). Efficient stochastic part-of-speech tagging for Hungarian. In *LREC-02*, Las Palmas, Canary Islands, Spain, pp. 710–717.

Ortmanns, S., Ney, H., and Aubert, X. (1997). A word graph algorithm for large vocabulary continuous speech recognition. *Computer Speech and Language*, , *11*, 43–72.

Ortony, A. (Ed.). (1993). *Metaphor* (2nd edition). Cambridge University Press, Cambridge.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The Measurement of Meaning*. University of Illinois Press, Urbana, IL.

O'Shaughnessy, D. (2000). *Speech Communications: Human and Machine*. IEEE Press, New York. 2nd. ed.

Ostendorf, M., Digilakis, V., and Kimball, O. (1996). From HMMs to segment models: A unified view of stochastic modeling for speech recognition. *IEEE Transactions on Speech and Audio*, *4*(5), 360–378.

Ostendorf, M. and Ross, K. (1997). Multi-level recognition of intonation labels. In Sagisaka, Y., Campbell, N., and Higuchi, N. (Eds.), *Computing Prosody: Computational Models for Processing Spontaneous Speech*, chap. 19, pp. 291–308. Springer.

Ostendorf, M. and Veilleux, N. (1994). A hierarchical stochastic model for automatic prediction of prosodic boundary location. *Computational Linguistics*, *20*(1).

Oviatt, S. L., Cohen, P. R., Wang, M. Q., and Gaston, J. (1993). A simulation-based research strategy for designing complex NL sysems. In *Proceedings DARPA Speech and Natural Language Workshop*, Princeton, NJ, pp. 370–375. Morgan Kaufmann.

Oviatt, S. L., MacEachern, M., and Levow, G.-A. (1998). Predicting hyperarticulate speech during human-computer error resolution. *Speech Communication*, *24*, 87–110.

Packard, D. W. (1973). Computer-assisted morphological analysis of ancient Greek. In Zampolli, A. and Calzolari, N. (Eds.), *Computational and Mathematical Linguistics: Proceedings of the International Conference on Computational Linguistics*, Pisa, pp. 343–355. Leo S. Olschki.

Palmer, D. D. (2000). Tokenisation and sentence segmentation. In Dale, R., Somers, H. L., and Moisl, H. (Eds.), *Handbook of Natural Language Processing*. Marcel Dekker.

Palmer, M., Dang, H. T., and Fellbaum, C. (2006). Making fine-grained and coarse-grained sense distinctions, both manually and automatically. *Natural Language Engineering*, *13*(2), 137–163.

Palmer, M., Babko-Malaya, O., and Dang, H. T. (2004). Different sense granularities for different applications. In *HLT-NAACL Workshop on Scalable Natural Language Understanding*, Boston, MA, pp. 49–56.

Palmer, M., Fellbaum, C., Cotton, S., Delfs, L., and Dang, H. T. (2001a). English tasks: All-words and verb lexical sample. In *Proceedings of SENSEVAL-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems*, Toulouse, France.

Palmer, M., Fellbaum, C., Cotton, S., Delfs, L., and Dang, H. T. (2001b). English tasks: All-words and verb lexical sample. In *Proceedings of Senseval-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems*, Toulouse, France, pp. 21–24.

Palmer, M. and Finin, T. (1990). Workshop on the evaluation of natural language processing systems. *Computational Linguistics*, *16*(3), 175–181.

Palmer, M., Kingsbury, P., and Gildea, D. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, *31*(1), 71–106.

Palmer, M., Ng, H. T., and Dang, H. T. (2006). Evaluation of wsd systems. In Agirre, E. and Edmonds, P. (Eds.), *Word Sense Disambiguation: Algorithms and Applications*. Kluwer.

Pan, S. and Hirschberg, J. (2000). Modeling local context for pitch accent prediction. In *ACL-00*, Hong Kong, pp. 233–240.

Pan, S. and McKeown, K. R. (1999). Word informativeness and automatic pitch accent modeling. In *EMNLP/VLC-99*.

Pang, B., Knight, K., and Marcu, D. (2003). Syntax-based alignment of multiple translations: extracting paraphrases and generating new sentences. In *HLT-NAACL-03*, pp. 102–109.

Pantel, P. and Ravichandran, D. (2004). Automatically labeling semantic classes. In *HLT-NAACL-04*, Boston, MA.

Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In *ACL-02*, Philadelphia, PA.

Parsons, T. (1990). *Events in the Semantics of English*. MIT Press.

Partee, B. H. (Ed.). (1976). *Montague Grammar*. Academic Press.

Partee, B. H., ter Meulen, A., and Wall, R. E. (1990). *Mathematical Methods in Linguistics*. Kluwer, Dordrecht.

Pasca, M. (2003). *Open-Domain Question Answering from Large Text Collections*. CSLI.

Passonneau, R. (2006). Measuring agreement on set-valued items (masi) for semantic and pragmatic annotation. In *LREC-06*.

Passonneau, R., Nenkova, A., McKeown, K. R., and Sigleman, S. (2005). Applying the pyramid method in DUC 2005. In *In Proceedings of the Document Understanding Conference (DUC'05)*.

Passonneau, R. and Litman, D. J. (1993). Intention-based segmentation: Human reliability and correlation with linguistic cues. In *ACL-93*, Columbus, Ohio, pp. 148–155.

Pater, J. (2007). Gradual learning and convergence. *Linguistic Inquiry*. In press.

Pater, J., Potts, C., and Bhatt, R. (2007). Harmonic grammar with linear programming. unpublished manuscript.

Patwardhan, S., Banerjee, S., and Pedersen, T. (2003). Using measures of semantic relatedness for word sense disambiguation. In *Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics*, pp. 241–257. Springer.

Paul, D. B. (1991). Algorithms for an optimal A^* search and linearizing the search in the stack decoder. In *IEEE ICASSP-91*, Vol. 1, pp. 693–696. IEEE.

Pearl, J. (1984). *Heuristics*. Addison-Wesley, Reading, MA.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman.

Pedersen, T. and Bruce, R. (1997). Distinguishing word senses in untagged text. In *EMNLP 1997*, Providence, RI.

Pedersen, T., Patwardhan, S., and Michelizzi, J. (2004). WordNet::Similarity – Measuring the relatedness of concepts. In *HLT-NAACL-04*.

Peirce, C. S. (1955). Abduction and induction. In Buchler, J. (Ed.), *Philosophical Writings of Peirce*, pp. 150–156. Dover Books, New York.

Penn, G. and Munteanu, C. (2003). A tabulation-based parsing method that reduces copying. In *Proceedings of the 41st Annual Meeting of the ACL*, Sapporo, Japan.

Percival, W. K. (1976). On the historical source of immediate constituent analysis. In McCawley, J. D. (Ed.), *Syntax and Semantics Volume 7, Notes from the Linguistic Underground*, pp. 229–242. Academic Press.

Pereira, F. C. N. (1985). A structure-sharing representation for unification-based grammar formalisms. In *ACL-85*, Chicago, pp. 137–144.

Pereira, F. C. N., Riley, M. D., and Sproat, R. (1994). Weighted rational transductions and their applications to human language processing. In *ARPA Human Language Technology Workshop*, Plainsboro, NJ, pp. 262–267. Morgan Kaufmann.

Pereira, F. C. N. and Shieber, S. M. (1984). The semantics of grammar formalisms seen as computer languages. In *COLING-84*, Stanford, CA, pp. 123–129.

Pereira, F. C. N. and Shieber, S. M. (1987). *Prolog and Natural-Language Analysis*, Vol. 10 of *CSLI Lecture Notes*. Chicago University Press, Chicago.

Pereira, F. C. N., Tishby, N., and Lee, L. (1993). Distributional clustering of English words. In *ACL-93*, Columbus, Ohio, pp. 183–190.

Pereira, F. C. N. and Warren, D. H. D. (1980). Definite clause grammars for language analysis— a survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence*, *13*(3), 231–278.

Pereira, F. C. N. and Wright, R. N. (1997). Finite-state approximation of phrase-structure grammars. In Roche, E. and Schabes, Y. (Eds.), *Finite-State Language Processing*, pp. 149–174. MIT Press.

Perrault, C. R. and Allen, J. (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, *6*(3-4), 167–182.

Peshkin, L. and Pfefer, A. (2003). Bayesian information extraction network. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*.

Peters, W., Peters, I., , and Vossen, P. (1998). Automatic sense clustering in EuroWordNet. In *LREC-98*, Granada, Spain, pp. 447–454.

Peterson, G. E. and Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.

Peterson, G. E., Wang, W. S.-Y., and Sivertsen, E. (1958). Segmentation techniques in speech synthesis. *Journal of the Acoustical Society of America*, *30*(8), 739–742.

Peterson, J. L. (1986). A note on undetected typing errors. *Communications of the ACM*, *29*(7), 633–637.

Petrov, S., Barrett, L., Thibaux, R., and Klein, D. (2006). Learning accurate, compact, and interpretable tree annotation. In *COLING/ACL 2006*, Sydney, Australia, pp. 433–440.

Pevzner, L. and Hearst, M. A. (2002). A critique and improvement of an evaluation metric for text segmentation. *Computational Linguistics*, *28*(1), 19–36.

Pieraccini, R., Levin, E., and Lee, C.-H. (1991). Stochastic representation of conceptual structure in the ATIS task. In *Proceedings DARPA Speech and Natural Language Workshop*, Pacific Grove, CA, pp. 121–124. Morgan Kaufmann.

Pierce, J. R., Carroll, J. B., and et al (1966). *Language and Machines: Computers in Translation and Linguistics*. ALPAC report. National Academy of Sciences, National Research Council, Washington, DC.

Pierre, I. (1984). Another look at nominal compounds. In *COLING-84*, Stanford, CA, pp. 509–516.

Pierrehumbert, J. B. and Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In Cohen, P. R., Morgan, J., and Pollack, M. (Eds.), *Intentions in Communication*, pp. 271–311. MIT Press.

Pierrehumbert, J. B. (1980). *The Phonology and Phonetics of English Intonation*. Ph.D. thesis, MIT.

Pito, R. (1993). Tgrepdoc man page..

Pitrelli, J. F., Beckman, M. E., and Hirschberg, J. (1994). Evaluation of prosodic transcription labeling reliability in the ToBI framework. In *ICSLP-94*, Vol. 1, pp. 123–126.

Plotkin, J. B. and Nowak, M. A. (2000). Language evolution and information theory. *Journal of Theoretical Biology*, *205*(1), 147–159.

Poesio, M., Stevenson, R., Di Eugenio, B., and Hitzeman, J. (2004). Centering: A parametric theory and its instantiations. *Computational Linguistics*, *30*(3), 309–363.

Poesio, M. and Vieira, R. (1998). A corpus-based investigation of definite description use. *Computational Linguistics*, *24*(2), 183–216.

Polanyi, L., Culy, C., van den Berg, M., Thione, G. L., and Ahn, D. (2004a). A Rule Based Approach to Discourse Parsing. In *Proceedings of SIGDIAL*.

Polanyi, L., Culy, C., van den Berg, M., Thione, G. L., and Ahn, D. (2004b). Sentential Structure and Discourse Parsing. In *Discourse Annotation Workshop, ACL04*.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, *12*.

Polifroni, J., Hirschman, L., Seneff, S., and Zue, V. W. (1992). Experiments in evaluating interactive spoken language systems. In *Proceedings DARPA Speech and Natural Language Workshop*, Harriman, NY, pp. 28–33. Morgan Kaufmann.

Pollard, C. (1984). *Generalized phrase structure grammars, head grammars, and natural language*. Ph.D. thesis, Stanford University.

Pollard, C. and Moshier, M. A. (1990). Unifying partial descriptions of sets. In Hanson, P. P. (Ed.), *Information, Language, and Cognition*, pp. 285–322. University of British Columbia Press, Vancouver.

Pollard, C. and Sag, I. A. (1987). *Information-Based Syntax and Semantics: Volume 1: Fundamentals*. University of Chicago Press, Chicago.

Pollard, C. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Pollock, J. J. and Zamora, A. (1975). Automatic abstracting research at Chemical Abstracts Service. *Journal of Chemical Information and Computer Sciences*, *15*(4), 226–232.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130–127.

Power, R. (1979). The organization of purposeful dialogs. *Linguistics*, *17*, 105–152.

Power, R., Scott, D., and Bouayad-Agha, N. (2003). Document structure. *Computational Linguistics*, *29*(2), 211–260.

Pradhan, S., Hacioglu, K., Ward, W., Martin, J. H., and Jurafsky, D. (2003). Semantic role parsing: Adding semantic structure to unstructured text. In *Proceedings of the International Conference on Data Mining (ICDM-2003)*.

Pradhan, S., Ward, W., Hacioglu, K., Martin, J. H., and Jurafsky, D. (2005). Semantic role labeling using different syntactic views. In *ACL-05*, Ann Arbor, MI.

Price, P. J., Fisher, W., Bernstein, J., and Pallet, D. (1988). The DARPA 1000-word resource management database for continuous speech recognition. In *IEEE ICASSP-88*, New York, Vol. 1, pp. 651–654.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., and Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*(6).

Prince, A. and Smolensky, P. (1993). Optimality theory: Constraint interaction in generative grammar.. Appeared as Tech. rep. CU-CS-696-93, Department of Computer Science, University of Colorado at Boulder, and TEch. rep. TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, April 1993.

Prince, A. and Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Blackwell.

Prince, A. and Tesar, B. (2004). Learning phonotactic distributions. In Kager, R., Pater, J., and Zonneveld, W. (Eds.), *Constraints in Phonological Acquisition*, pp. 245–291. Cambridge University Press.

Prince, E. (1981). Toward a taxonomy of given-new information. In Cole, P. (Ed.), *Radical Pragmatics*, pp. 223–255. Academic Press.

Prince, E. (1992). The ZPG letter: Subjects, definiteness, and information-status. In Thompson, S. and Mann, W. (Eds.), *Discourse Description: Diverse Analyses of a Fundraising Text*, pp. 295–325. John Benjamins, Philadelphia/Amsterdam.

Prüst, H. (1992). *On Discourse Structuring, VP Anaphora, and Gapping*. Ph.D. thesis, University of Amsterdam.

Pullum, G. K. and Gazdar, G. (1982). Natural languages and context-free languages. *Linguistics and Philosophy*, *4*, 471–504.

Pullum, G. K. (1991). *The Great Eskimo Vocabulary Hoax*. University of Chicago, Chicago, IL.

Pullum, G. K. and Ladusaw, W. A. (1996). *Phonetic Symbol Guide*. University of Chicago, Chicago, IL. Second Edition.

Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press.

Pustejovsky, J. and Bergler, S. (Eds.). (1992). *Lexical Semantics and Knowledge Representation*. Lecture Notes in Artificial Intelligence. Springer Verlag, Berlin.

Pustejovsky, J. and Boguraev, B. (Eds.). (1996). *Lexical Semantics: The Problem of Polysemy*. Oxford University Press, Oxford.

Pustejovsky, J., Castao, J., Ingria, R., Saur, R., Gaizauskas, R., Setzer, A., and Katz, G. (2003a). TimeML: robust specification of event and temporal expressions in text. In *Proceedings of the Fifth International Workshop on Computational Semantics (IWCS-5)*.

Pustejovsky, J., Hanks, P., Saurí, R., See, A., Gaizauskas, R., Setzer, A., Radev, D., Sundheim, B., Day, D., Ferro, L., and Lazo, M. (2003b). The TIMEBANK corpus. In *Proceedings of the Corpus Linguistics 2003 conference*, pp. 647–656. UCREL Technical Paper number 16.

Pustejovsky, J., Hanks, P., Saur, R., See, A., Gaizauskas, R., Setzer, A., Radev, D., Sundheim, B., Day, D., Ferro, L., and Lazo, M. (2003c). The TIMEBANK corpus. In *Proceedings of Corpus Linguistics 2003 Conference*, pp. 647–656.

Pustejovsky, J., Ingria, R., Sauri, R., Castano, J., Littman, J., Gaizauskas, R., Setzer, A., Katz, G., and Mani, I. (2005). *The Specification Language TimeML*, chap. 27. Oxford, Oxford, England.

Qu, Y., Shanahan, J., and Wiebe, J. (Eds.). (2004). *Exploring Attitude and Affect in Text: Papers from the 2004 Spring Symposium*, Stanford, Ca. AAAI.

Quillian, M. R. (1968). Semantic memory. In Minsky, M. (Ed.), *Semantic Information Processing*, pp. 227–270. MIT Press.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106.

Quirk, C., Brockett, C., and Dolan, W. B. (2004). Monolingual machine translation for paraphrase generation. In *EMNLP 2004*, pp. 142–149.

Quirk, C., Menezes, A., and Cherry, C. (2005). Dependency treelet translation: Syntactically informed phrasal smt. In *ACL-05*.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Longman, London.

Rabin, M. O. and Scott, D. (1959). Finite automata and their decision problems. *IBM Journal of Research and Development*, *3*(2), 114–125.

Rabiner, L. R. (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286.

Rabiner, L. R. and Juang, B. H. (1993). *Fundamentals of Speech Recognition*. Prentice Hall.

Radev, D., Blair-Goldensohn, S., and Zhang, Z. (2001). Experiments in single and multi-document summarization using MEAD. In *DUC-01*, New Orleans, LA.

Radev, D., Jing, H., and Budzikowska, M. (2000). Summarization of multiple documents: clustering, sentence extraction, and evaluation. In *ANLP-NAACL Workshop on Automatic Summarization*, Seattle, WA.

Radford, A. (1988). *Transformational Grammar: A First Course*. Cambridge University Press.

Radford, A. (1997). *Syntactic Theory and the Structure of English: A Minimalist Approach*. Cambridge University Press.

Ramshaw, L. A. and Marcus, M. P. (1995). Text chunking using transformation-based learning. In *Proceedings of the Third Annual Workshop on Very Large Corpora*, pp. 82–94.

Rand, D. and Sankoff, D. (1990). Goldvarb: A variable rule application for the macintosh. Available at http://www.crm.umontreal.ca/ sankoff/GoldVarb_Eng.html.

Raphael, B. (1968). SIR: A computer program for semantic information retrieval. In Minsky, M. (Ed.), *Semantic Information Processing*, pp. 33–145. MIT Press.

Ratnaparkhi, A. (1996). A maximum entropy part-of-speech tagger. In *EMNLP 1996*, Philadelphia, PA, pp. 133–142.

Ratnaparkhi, A., Reynar, J. C., and Roukos, S. (1994). A Maximum Entropy model for prepositional phrase attachment. In *ARPA Human Language Technologies Workshop*, Plainsboro, N.J., pp. 250–255.

Rau, L. F., Jacobs, P. S., and Zernik, U. (1989). Information extraction and text summarization using linguistic knowledge acquisition. *Information Processing and Management*, *25*(4), 419–428.

Ravichandran, D. and Hovy, E. H. (2002). Learning surface text patterns for a question answering system. In *ACL-02*, Philadelphia, PA, pp. 41–47.

Ravishankar, M. K. (1996). *Efficient Algorithms for Speech Recognition*. Ph.D. thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh. Available as CMU CS tech report CMU-CS-96-143.

Rayner, M. and Hockey, B. A. (2003). Transparent combination of rule-based and data-driven approaches in a speech understanding architecture. In *EACL-03*, Budapest, Hungary.

Rayner, M. and Hockey, B. A. (2004). Side effect free dialogue management in a voice enabled procedure browser. In *ICSLP-04*, pp. 2833–2836.

Rayner, M., Hockey, B. A., and Bouillon, P. (2006). *Putting Linguistics into Speech Recognition*. CSLI.

Rebholz-Schuhmann, D., Marcel, S., Albert, S., Tolle, R., Casari, G., and Kirsch, H. (2004). Automatic extraction of mutations from medline and cross-validation with omim. *Nucleic Acids Research*, *32*(1), 135–142.

Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press.

Regier, T. (1996). *The Human Semantic Potential*. MIT Press.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. Macmillan, New York.

Reichert, T. A., Cohen, D. N., and Wong, A. K. C. (1973). An application of information theory to genetic mutations and the matching of polypeptide sequences. *Journal of Theoretical Biology*, *42*, 245–261.

Reichman, R. (1985). *Getting Computers to Talk Like You and Me*. MIT Press.

Reimer, U. and Hahn, U. (1988). Text condensation as knowledge base abstraction. In *CAIA-88*, pp. 14–18.

Reiter, E. and Dale, R. (2000). *Building Natural Language Generation Systems*. Cambridge University Press.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, *13*, 81–132.

Reithinger, N., Engel, R., Kipp, M., and Klesen, M. (1996). Predicting dialogue acts for a speech-to-speech translation system. In *ICSLP-96*, Philadelphia, PA, Vol. 2, pp. 654–657.

Resnik, P. (1992). Probabilistic tree-adjoining grammar as a framework for statistical natural language processing. In *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France, pp. 418–424.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxanomy. In *International Joint Conference for Artificial Intelligence (IJCAI-95)*, pp. 448–453.

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition*, *61*, 127–159.

Resnik, P. (1997). Selectional preference and sense disambiguation. In *Proceedings of ACL SIGLEX Workshop on Tagging Text with Lexical Semantics*, Washington, D.C., pp. 52–57.

Resnik, P. (1998). WordNet and class-based probabilities. In Fellbaum, C. (Ed.), *WordNet: An Electronic Lexical Database*. MIT Press.

Resnik, P. (2006). Word sense disambiguation in nlp appplications. In Agirre, E. and Edmonds, P. (Eds.), *Word Sense Disambiguation: Algorithms and Applications*. Kluwer.

Reynar, J. C. (1994). An automatic method of finding topic boundaries. In *ACL-94*, pp. 27–30.

Reynar, J. C. (1999). Statistical models for topic segmentation. In *ACL/EACL-97*, pp. 357–364.

Reynolds, D. A. and Rose, R. C. (1995). Robust text-independent speaker identification using gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, *3*(1), 72–83.

Rhodes, F. L. (1929). *Beginnings of telephony*. Harper and Brothers, New York.

Rhodes, R. A. (1992). Flapping in American English. In Dressler, W. U., Prinzhorn, M., and Rennison, J. (Eds.), *Proceedings of the 7th International Phonology Meeting*, pp. 217–232. Rosenberg and Sellier, Turin.

Riesbeck, C. K. (1975). Conceptual analysis. In Schank, R. C. (Ed.), *Conceptual Information Processing*, pp. 83–156. American Elsevier, New York.

Riesbeck, C. K. (1986). From conceptual analyzer to direct memory access parsing: An overview. In *Advances in Cognitive Science 1*, pp. 236–258. Ellis Horwood, Chichester.

Riezler, S., King, T. H., Crouch, R., and Zaenen, A. (2003). Statistical sentence condensation using ambiguity packing and stochastic disambiguation methods for Lexical-Functional Grammar. In *HLT-NAACL-03*, Edmonton, Canada.

Riezler, S., Prescher, D., Kuhn, J., and Johnson, M. (2000). Lexicalized stochastic modeling of constraint-based grammars using log-linear measures and em training. In *ACL-00*, Hong Kong.

Riezler, S., King, T. H., Kaplan, R. M., Crouch, R., III, J. T. M., and Johnson, M. (2002). Parsing the wall street journal using a lexical-functional grammar and discriminative estimation techniques. In *ACL-02*, Philadelphia, PA.

Riggle, J. (2005). Contenders and learning. In *WCCFL 23*, pp. 101–114.

Riley, M. D. (1992). Tree-based modelling for speech synthesis. In Bailly, G. and Beniot, C. (Eds.), *Talking Machines: Theories, Models and Designs*. North Holland, Amsterdam.

Riloff, E. (1993). Automatically constructing a dictionary for information extraction tasks. In *AAAI-93*, Washington, D.C., pp. 811–816.

Riloff, E. (1996). Automatically generating extraction patterns from untagged text. In *AAAI-96*, pp. 117–124.

Riloff, E. and Jones, R. (1999). Learning dictionaries for information extraction by multi-level bootstrapping. In *AAAI-99*, pp. 474–479.

Rindflesch, T. C., Tanabe, L., Weinstein, J. N., and Hunter, L. (2000). EDGAR: Extraction of drugs, genes and relations from the biomedical literature. In *Pacific Symposium on Biocomputing*, pp. 515–524.

Rivest, R. L. (1987). Learning decision lists. *Machine Learning*, *2*(3), 229–246.

Roark, B. (2001). Probabilistic top-down parsing and language modeling. *Computational Linguistics*, *27*(2), 249–276.

Roark, B. and Sproat, R. (2007). *Computational Approaches to Morphology and Syntax*. Oxford University Press.

Robins, R. H. (1967). *A Short History of Linguistics*. Indiana University Press, Bloomington.

Robinson, J. A. (1965). A machine-oriented logic based on the resolution principle. *Journal of the Association for Computing Machinery*, *12*, 23–41.

Robinson, J. J. (1975). Performance grammars. In Reddy, D. R. (Ed.), *Speech Recognition: Invited Paper Presented at the 1974 IEEE Symposium*, pp. 401–427. Academic Press.

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

Rocchio, J. J. (1971). Relevance feedback in information retrieval. In *The SMART Retrieval System: Experiments in Automatic Indexing*, pp. 324–336. Prentice Hall.

Roche, E. and Schabes, Y. (1997a). Deterministic part-of-speech tagging with finite-state transducers. In Roche, E. and Schabes, Y. (Eds.), *Finite-State Language Processing*, pp. 205–239. MIT Press.

Roche, E. and Schabes, Y. (1997b). Introduction. In Roche, E. and Schabes, Y. (Eds.), *Finite-State Language Processing*, pp. 1–65. MIT Press.

Rohde, D. L. T. (2005). Tgrep2 user manual..

Rosario, B. and Hearst, M. A. (2004). Classifying semantic relations in bioscience texts. In *Proceedings of ACL 2004*, pp. 430–437.

Rosario, B. and Hearst, M. A. (2005). Multi-way Relation Classification: Application to Protein-Protein Interactions. In *HLT-EMNLP-05*.

Rosenberg, A. E. (1971). Effect of Glottal Pulse Shape on the Quality of Natural Vowels. *The Journal of the Acoustical Society of America*, *49*, 583–590.

Rosenfeld, R. (1996). A maximum entropy approach to adaptive statistical language modeling. *Computer Speech and Language*, *10*, 187–228.

Roth, D. and tau Yih, W. (2001). Relational learning via propositional algorithms: An information extraction case study. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1257–1263.

Roth, D. and Zelenko, D. (1998). Part of speech tagging using a network of linear separators. In *COLING/ACL-98*, Montreal, Canada, pp. 1136–1142.

Rounds, W. C. and Kasper, R. T. (1986). A complete logical calculus for record structures representing linguistic information. In *Proceedings of the 1st Annual IEEE Symposium on Logic in Computer Science*, pp. 38–43.

Roy, N., Pineau, J., and Thrun, S. (2000). Spoken dialog management for robots. In *ACL-00*, Hong Kong.

Rubenstein, H. and Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM*, *8*(10), 627–633.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., and Scheffczyk, J. (2006). FrameNet ii: Extended theory and practice. Version 1.3, http://www.icsi.berkeley.edu/framenet/.

Russell, S. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Russell, S. and Norvig, P. (2002). *Artificial Intelligence: A Modern Approach*. Prentice Hall. Second edition.

Russell, S. W. (1976). Computer understanding of metaphorically used verbs. *American Journal of Computational Linguistics*, *2*. Microfiche 44.

Ryder, M. E. (1994). *Ordered Chaos: The Interpretation of English Noun-Noun Compounds*. University of California Press, Berkeley.

Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Sadek, D. and De Mori, R. (1998). Dialogue systems. In De Mori, R. (Ed.), *Spoken Dialogues With Computers*. Academic Press, London.

Sadek, M. D. (1991). Dialogue acts are rational plans. In *ESCA/ETR Workshop on the Structure of Multimodal Dialogue*, pp. 19–48.

Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996a). Statistical learning by 8-month old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.

Sag, I. A. and Liberman, M. Y. (1975). The intonational disambiguation of indirect speech acts. In *CLS-75*, pp. 487–498. University of Chicago.

Sag, I. A. and Wasow, T. (Eds.). (1999). *Syntactic Theory: A Formal Introduction*. CSLI Publications, Stanford, CA.

Sag, I. A., Wasow, T., and Bender, E. M. (Eds.). (2003). *Syntactic Theory: A Formal Introduction*. CSLI Publications, Stanford, CA.

Sagisaka, Y. (1988). Speech synthesis by rule using an optimal selection of non-uniform synthesis units. In *IEEE ICASSP-88*, pp. 679–682.

Sagisaka, Y., Kaiki, N., Iwahashi, N., and Mimura, K. (1992). Atr – v-talk speech synthesis system. In *ICSLP-92*, Banff, Canada, pp. 483–486.

Sagisaka, Y., Campbell, N., and Higuchi, N. (Eds.). (1997). *Computing Prosody: Computational Models for Processing Spontaneous Speech*. Springer.

Saint-Dizier, P. and Viegas, E. (Eds.). (1995). *Computational Lexical Semantics*. Cambridge University Press.

Sakoe, H. and Chiba, S. (1971). A dynamic programming approach to continuous speech recognition. In *Proceedings of the Seventh International Congress on Acoustics*, Budapest, Vol. 3, pp. 65–69. Akadémiai Kiadó.

Sakoe, H. and Chiba, S. (1984). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *ASSP-26*(1), 43–49.

Salasoo, A. and Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*, *24*, 210–231.

Salomaa, A. (1969). Probabilistic and weighted grammars. *Information and Control*, *15*, 529–544.

Salomaa, A. (1973). *Formal Languages*. Academic.

Salton, G. (1971). *The SMART Retrieval System: Experiments in Automatic Document Processing*. Prentice Hall.

Salton, G. and Buckley, C. (1990). Improving retrieval performance by relevance feedback. *Information Processing and Management*, *41*, 288–297.

Salton, G. and McGill, M. J. (1983). *Introduction to Modern Information Retrieval*. McGraw-Hill, New York, NY.

Sampson, G. (1987). Alternative grammatical coding systems. In Garside, R., Leech, G., and Sampson, G. (Eds.), *The Computational Analysis of English*, pp. 165–183. Longman, London and New York.

Sampson, G. (1996). *Evolutionary Language Understanding*. Cassell, London.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, *110*, 474–494.

Samuel, K., Carberry, S., and Vijay-Shanker, K. (1998). Computing dialogue acts from features with transformation-based learning. In Chu-Carroll, J. and Green, N. (Eds.), *Applying Machine Learning to Discourse Processing. Papers from the 1998 AAAI Spring Symposium*, pp. 90–97. Technical Report SS-98-01.

Samuelsson, C. (1993). Morphological tagging based entirely on Bayesian inference. In *9th Nordic Conference on Computational Linguistics NODALIDA-93*. Stockholm.

Samuelsson, C. and Reichl, W. (1999). A class-based language model for large-vocabulary speech recognition extracted from part-of-speech statistics. In *IEEE ICASSP-99*, pp. 537–540. IEEE.

San-Segundo, R., Montero, J. M., Ferreiros, J., Còrdoba, R., and Pardo, J. M. (2001). Designing confirmation mechanisms and error recovery techniques in a railway information system for Spanish. In *In Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue, Aalborg, Denmark*.

Sanders, T. J. M., Spooren, W. P. M., and Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1–35.

Sanderson, M. (1994). Word sense disambiguation and information retrieval. In *SIGIR-94*, Dublin, Ireland, pp. 142–151. ACM.

Sanfilippo, A. (1993). LKB encoding of lexical knowledge. In Briscoe, T., de Paiva, V., and Copestake, A. (Eds.), *Inheritance, Defaults, and the Lexicon*, pp. 190–222. Cambridge University Press.

Sang, E. F. T. K. (2002). Introduction to the conll-2002 shared task: Language-independent named entity recognition. In *Proceedings of CoNLL-2002*, pp. 155–158. Taipei, Taiwan.

Sang, E. F. T. K. and De Meulder, F. (2003). Introduction to the conll-2003 shared task: Language-independent named entity recognition. In Daelemans, W. and Osborne, M. (Eds.), *Proceedings of CoNLL-2003*, pp. 142–147. Edmonton, Canada.

Sankoff, D. (1972). Matching sequences under deletion-insertion constraints. *Proceedings of the Natural Academy of Sciences of the U.S.A.*, *69*, 4–6.

Santorini, B. (1990). Part-of-speech tagging guidelines for the Penn Treebank project. 3rd revision, 2nd printing.

Saraclar, M., Nock, H., and Khudanpur, S. (2000). Pronunciation modeling by sharing gaussian densities across phonetic models. *Computer Speech and Language*, *14*(2), 137–160.

Sassano, M. and Utsuro, T. (2000). Named entity chunking techniques in supervised learning for japanese named entity recognition. In *COLING-00*, Saarbrcken, Germany, pp. 705–711.

Scha, R. and Polanyi, L. (1988). An augmented context free grammar for discourse. In *COLING-88*, Budapest, pp. 573–577.

Schabes, Y. (1990). *Mathematical and Computational Aspects of Lexicalized Grammars*. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

Schabes, Y. (1992). Stochastic lexicalized tree-adjoining grammars. In *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France, pp. 426–433.

Schabes, Y., Abeillé, A., and Joshi, A. K. (1988). Parsing strategies with 'lexicalized' grammars: Applications to Tree Adjoining Grammars. In *COLING-88*, Budapest, pp. 578–583.

Schachter, P. (1985). Parts-of-speech systems. In Shopen, T. (Ed.), *Language Typology and Syntactic Description, Volume 1*, pp. 3–61. Cambridge University Press.

Schank, R. C. (1972). Conceptual dependency: A theory of natural language processing. *Cognitive Psychology*, *3*, 552–631.

Schank, R. C. and Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum.

Schank, R. C. and Albelson, R. P. (1977). *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum.

Schank, R. C. and Riesbeck, C. K. (Eds.). (1981). *Inside Computer Understanding: Five Programs plus Miniatures*. Lawrence Erlbaum.

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, *70*, 1075–1095.

Schegloff, E. A. (1979). Identification and recognition in telephone conversation openings. In Psathas, G. (Ed.), *Everyday Language: Studies in Ethnomethodology*, pp. 23–78. Irvington.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In Tannen, D. (Ed.), *Analyzing Discourse: Text and Talk*, pp. 71–93. Georgetown University Press, Washington, D.C.

Schone, P. and Jurafsky, D. (2000). Knowlege-free induction of morphology using latent semantic analysis. In *Proceedings of the Conference on Computational Natural Language Learning (CoNLL-2000)*.

Schone, P. and Jurafsky, D. (2001). Knowledge-free induction of inflectional morphologies. In *NAACL 2001*.

Schönfinkel, M. (1924). Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, *92*, 305–316. English Translation appears in *From Frege to Gödel: A Source Book in Mathematical Logic*, Harvard University Press, Cambridge, MA, 1967.

Schroder, M. (2006). Expressing degree of activation in synthetic speech. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(4), 1128–1136.

Schubert, L. K. and Pelletier, F. J. (1982). From English to logic: Context-free computation of 'conventional' logical translation. *American Journal of Computational Linguistics*, *8*(1), 27–44.

Schulte im Walde, S. (2000). Clustering verbs semantically according to their alternation behaviour. In *COLING-00*, Saarbrücken, Germany, pp. 747–753.

Schütze, H. (1992a). Context space. In Goldman, R. (Ed.), *Proceedings of the 1992 AAAI Fall Symposium on Probabilistic Approaches to Natural Language*.

Schütze, H. (1992b). Dimensions of meaning. In *Proceedings of Supercomputing '92*, pp. 787–796. IEEE, IEEE Press.

Schütze, H. (1995). Distributional part-of-speech tagging. In *EACL-95*.

Schütze, H. (1997). *Ambiguity Resolution in Language Learning: Computational and Cognitive Models*. CSLI Publications, Stanford, CA.

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

Schütze, H. and Pedersen, J. (1995). Information retrieval based on word senses. In *Proceedings of the Fourth Annual Symposium on Document Analysis and Information Retrieval*, Las Vegas, pp. 161–175.

Schütze, H. and Singer, Y. (1994). Part-of-speech tagging using a variable memory Markov model. In *ACL-94*, Las Cruces, NM, pp. 181–187.

Schützenberger, M. P. (1977). Sur une variante des fonctions sequentielles. *Theoretical Computer Science*, *4*, 47–57.

Schwartz, A. S. and Hearst, M. A. (2003). A simple algorithm for identifying abbreviation definitions in biomedical text. In *Pacific Symposium on Biocomputing*, Vol. 8, pp. 451–462.

Schwartz, R. and Austin, S. (1991). A comparison of several approximate algorithms for finding multiple (*N*-BEST) sentence hypotheses. In *icassp91*, Toronto, Vol. 1, pp. 701–704. IEEE.

Schwartz, R. and Chow, Y.-L. (1990). The N-best algorithm: An efficient and exact procedure for finding the N most likely sentence hypotheses. In *IEEE ICASSP-90*, Vol. 1, pp. 81–84. IEEE.

Schwartz, R., Chow, Y.-L., Kimball, O., Roukos, S., Krasnwer, M., and Makhoul, J. (1985). Context-dependent modeling for acoustic-phonetic recognition of continuous speech. In *IEEE ICASSP-85*, Vol. 3, pp. 1205–1208. IEEE.

Scott, M. and Shillcock, R. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648–652.

Searle, J. R. (1975a). Indirect speech acts. In Cole, P. and Morgan, J. L. (Eds.), *Speech Acts: Syntax and Semantics Volume 3*, pp. 59–82. Academic Press.

Searle, J. R. (1975b). A taxonomy of illocutionary acts. In Gunderson, K. (Ed.), *Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science*, Vol. VII, pp. 344–369. University of Minnesota Press, Amsterdam. Also appears in John R. Searle, *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge University Press, 1979.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417–457.

Sejnowski, T. J. and Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex Systems*, *1*(1), 145–168.

Sekine, S. and Collins, M. (1997). The evalb software. http://cs.nyu.edu/cs/projects/proteus/evalb.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, *3*, 371–405.

Seneff, S. (1995). TINA: A natural language system for spoken language application. *Computational Linguistics*, *18*(1), 62–86.

Seneff, S. (2002). Response planning and generation in the MERCURY flight reservation system. *Computer Speech and Language, Special Issue on Spoken Language Generation*, *16*(3-4), 283–312.

Seneff, S., Lau, R., and Meng, H. (1996). ANGIE: A new framework for speech analysis based on morpho-phonological modelling. In *ICSLP-96*.

Seneff, S. and Polifroni, J. (2000). Dialogue management in the mercury flight reservation system. In *ANLP/NAACL Workshop on Conversational Systems*, Seattle.

Seneff, S. and Zue, V. W. (1988). Transcription and alignment of the TIMIT database. In *Proceedings of the Second Symposium on Advanced Man-Machine Interface through Spoken Language*, Oahu, Hawaii.

Senellart, J., Dienes, P., and Váradi, T. (2001). New generation systran translation system. In *MT Summit 8*.

Settles, B. (2005). ABNER: an open source tool for automatically tagging genes,proteins and other entity names in text. *Bioinformatics*, *21*(14), 3191–3192.

Seuss, D. (1960). *One Fish Two Fish Red Fish Blue Fish*. Random House, New York.

Shannon, C. E. (1938). A symbolic analysis of relay and switching circuits. *Transactions of the American Institute of Electrical Engineers*, *57*, 713–723.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. Continued in the following volume.

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

Sheil, B. A. (1976). Observations on context free parsing. *SMIL: Statistical Methods in Linguistics*, *1*, 71–109.

Shieber, S. M. (1985a). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, *8*, 333–343.

Shieber, S. M. (1985b). Using restriction to extend parsing algorithms for complex-feature-based formalisms. In *ACL-85*, Chicago, pp. 145–152.

Shieber, S. M. (1986). *An Introduction to Unification-Based Approaches to Grammar*. Center for the Study of Language and Information, Stanford University, Stanford, CA.

Shieber, S. M. (1994a). Lessons from a restricted Turing test. *Communications of the ACM*, *37*(6), 70–78.

Shieber, S. M. (1994b). Restricting the weak-generative capacity of synchronous tree-adjoining grammars. *Computational Intelligence*, *10*(4), 371–385.

Shieber, S. M. and Schabes, Y. (1992). Generation and synchronous tree-adjoining grammars. *Computational Intelligence*, *7*(4), 220–228.

Shockey, L. (2003). *Sound Patterns of Spoken English*. Blackwell, Oxford.

Shoup, J. E. (1980). Phonological aspects of speech recognition. In Lea, W. A. (Ed.), *Trends in Speech Recognition*, pp. 125–138. Prentice-Hall.

Shriberg, E. (2002). To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, *31*(1), 153–169.

Shriberg, E. (2005). Spontaneous speech: How people really talk, and why engineers should care. In *INTERSPEECH-05*, Lisbon, Portugal.

Shriberg, E., Ferrer, L., Kajarekar, S., Venkataraman, A., and Stolcke, A. (2005). Modeling prosodic feature sequences for speaker recognition. *Speech Communication*, *46*(3-4), 455–472.

Shriberg, E., Stolcke, A., Hakkani-Tür, D., and Tür, G. (2000). Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication*, *32*(1-2), 127–154.

Shriberg, E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. Ph.D. thesis, University of California, Berkeley, CA. (unpublished).

Shriberg, E., Bates, R., Taylor, P., Stolcke, A., Jurafsky, D., Ries, K., Coccaro, N., Martin, R., Meteer, M., and Van Ess-Dykema, C. (1998). Can prosody aid the automatic classification of dialog acts in conversational speech?. *Language and Speech (Special Issue on Prosody and Conversation)*, *41*(3-4), 439–487.

Shriberg, E., Wade, E., and Price, P. J. (1992). Human-machine problem solving using spoken language systems (SLS): Factors affecting performance and user satisfaction. In *Proceedings DARPA Speech and Natural Language Workshop*, Harriman, NY, pp. 49–54. Morgan Kaufmann.

Siddharthan, A., Nenkova, A., and McKeown, K. R. (2004). Syntactic simplification for improving content selection in multi-document summarization. In *COLING-04*, p. 896.

Sidner, C. L. (1979). Towards a computational theory of definite anaphora comprehension in English discourse. Tech. rep. 537, MIT Artificial Intelligence Laboratory, Cambridge, MA.

Sidner, C. L. (1983). Focusing in the comprehension of definite anaphora. In Brady, M. and Berwick, R. C. (Eds.), *Computational Models of Discourse*, pp. 267–330. MIT Press.

Sills, D. L. and Merton, R. K. (Eds.). (1991). *Social Science Quotations*. MacMillan, New York.

Silverman, K., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P. J., Pierrehumbert, J. B., and Hirschberg, J. (1992). ToBI: a standard for labelling English prosody. In *ICSLP-92*, Vol. 2, pp. 867–870.

Simmons, R. F. (1965). Answering English questions by computer: A survey. *Communications of the ACM*, *8*(1), 53–70.

Simmons, R. F. (1973). Semantic networks: Their computation and use for understanding English sentences. In Schank, R. C. and Colby, K. M. (Eds.), *Computer Models of Thought and Language*, pp. 61–113. W.H. Freeman and Co., San Francisco.

Simmons, R. F. (1978). Rule-based computations on English. In Waterman, D. A. and Hayes-Roth, F. (Eds.), *Pattern-Directed Inference Systems*. Academic Press.

Simmons, R. F. (1983). *Computations from the English*. Prentice Hall.

Singh, S. P., Litman, D. J., Kearns, M., and Walker, M. A. (2002). Optimizing dialogue management with reinforcement learning: Experiments with the njfun system. *J. Artif. Intell. Res. (JAIR)*, *16*, 105–133.

Sleator, D. and Temperley, D. (1993). Parsing English with a link grammar. In *Proceeedings, Third International Workshop on Parsing Technologies*, Tilburg, The Netherlands/Durbuy, Belgium.

Slobin, D. I. (1996). Two ways to travel. In Shibatani, M. and Thompson, S. A. (Eds.), *Grammatical Constructions: Their Form and Meaning*, pp. 195–220. Clarendon Press, Oxford.

Small, S. L., Cottrell, G. W., and Tanenhaus, M. (Eds.). (1988). *Lexical Ambiguity Resolution*. Morgan Kaufman.

Small, S. L. and Rieger, C. (1982). Parsing and comprehending with Word Experts. In Lehnert, W. G. and Ringle, M. H. (Eds.), *Strategies for Natural Language Processing*, pp. 89–147. Lawrence Erlbaum.

Smith, D. A. and Eisner, J. (2007). Bootstrapping feature-rich dependency parsers with entropic priors. In *EMNLP/CoNLL 2007*, Prague, pp. 667–677.

Smith, N. A. and Eisner, J. (2005). Guiding unsupervised grammar induction using contrastive estimation. In *IJCAI Workshop on Grammatical Inference Applications*, Edinburgh, pp. 73–82.

Smith, N. A. and Johnson, M. (2007). Weighted and probabilistic context-free grammars are equally expressive. *Computational Linguistics*. To appear.

Smith, R. W. and Gordon, S. A. (1997). Effects of variable initiative on linguistic behavior in human-computer spoken natural language dialogue. *Computational Linguistics*, *23*(1), 141–168.

Smith, V. L. and Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*, 25–38.

Smolensky, P. and Legendre, G. (2006). *The Harmonic Mind*. MIT Press.

Smrž, O. (1998). *Functional Arabic Morphology*. Ph.D. thesis, Charles University in Prague.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *23*, 197–229.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., and Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. In *AMTA-2006*.

Snow, R., Jurafsky, D., and Ng, A. Y. (2005). Learning syntactic patterns for automatic hypernym discovery. In Saul, L. K., Weiss, Y., and Bottou, L. (Eds.), *NIPS 17*, pp. 1297–1304. MIT Press.

Snow, R., Jurafsky, D., and Ng, A. Y. (2006). Semantic taxonomy induction from heterogenous evidence. In *COLING/ACL 2006*.

Snow, R., Prakash, S., Jurafsky, D., and Ng, A. Y. (2007). Learning to merge word senses. In *EMNLP/CoNLL 2007*, pp. 1005–1014.

Soderland, S., Fisher, D., Aseltine, J., and Lehnert, W. G. (1995). CRYSTAL: Inducing a conceptual dictionary. In *IJCAI-95*, Montreal, pp. 1134–1142.

Soon, W. M., Ng, H. T., and Lim, D. C. Y. (2001). A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, *27*(4), 521–544.

Soong, F. K. and Huang, E.-F. (1990). A tree-trellis based fast search for finding the n-best sentence hypotheses in continuous speech recognition. In *Proceedings DARPA Speech and Natural Language Processing Workshop*, Hidden Valley, PA, pp. 705–708. Also in Proceedings of IEEE ICASSP-91, 705-708.

Soricut, R. and Marcu, D. (2006). Discourse generation using utility-trained coherence models. In *COLING/ACL 2006*, pp. 803–810.

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

Sparck Jones, K. (1986). *Synonymy and Semantic Classification*. Edinburgh University Press, Edinburgh. Republication of 1964 PhD Thesis.

Sparck Jones, K. (2007). Automatic summarising: The state of the art. *Information Processing and Management*, *43*(6), 1449–1481.

Sparck Jones, K. and Galliers, J. R. (Eds.). (1996). *Evaluating Natural Language Processing Systems*. Springer.

Sparck Jones, K. and Willett, P. (Eds.). (1997). *Readings in Information Retrieval*. Morgan Kaufmann, San Francisco, CA.

Spiegel, M. F. (2002). Proper name pronunciations for speech technology applications. In *Proceedings of IEEE Workshop on Speech Synthesis*, pp. 175–178.

Spiegel, M. F. (2003). Proper name pronunciations for speech technology applications. *International Journal of Speech Technology*, *6*(4), 419–427.

Sporleder, C. and Lapata, M. (2004). Automatic paragraph identification: A study across languages and domains. In *EMNLP 2004*.

Sporleder, C. and Lapata, M. (2006). Automatic Paragraph Identification: A Study across Languages and Domains. *ACM Transactions on Speech and Language Processing (TSLP)*, *3*(2).

Sporleder, C. and Lascarides, A. (2005). Exploiting Linguistic Cues to Classify Rhetorical Relations. In *Proceedings of the Recent Advances in Natural Language Processing (RANLP-05), Borovets, Bulgaria*.

Sproat, R. (1993). *Morphology and Computation*. MIT Press.

Sproat, R. (1994). English noun-phrase prediction for text-to-speech. *Computer Speech and Language*, *8*, 79–94.

Sproat, R. (1998a). Further issues in text analysis. In Sproat, R. (Ed.), *Multilingual Text-To-Speech Synthesis: The Bell Labs Approach*, pp. 89–114. Kluwer, Dordrecht.

Sproat, R. (Ed.). (1998b). *Multilingual Text-To-Speech Synthesis: The Bell Labs Approach*. Kluwer, Dordrecht.

Sproat, R., Black, A. W., Chen, S. F., Kumar, S., Ostendorf, M., and Richards, C. (2001). Normalization of non-standard words. *Computer Speech & Language*, *15*(3), 287–333.

Sproat, R., Shih, C., Gale, W. A., and Chang, N. (1996). A stochastic finite-state word-segmentation algorithm for Chinese. *Computational Linguistics*, *22*(3), 377–404.

Stabler, E. (1997). Derivational minimalism. In Retoré, C. (Ed.), *Logical Aspects of Computational Linguistics*, pp. 68–95. Springer.

Stalnaker, R. C. (1978). Assertion. In Cole, P. (Ed.), *Pragmatics: Syntax and Semantics Volume 9*, pp. 315–332. Academic Press.

Stanners, R. F., Neiser, J., Hernon, W. P., and Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, *18*, 399–412.

Steedman, M. (1989). Constituency and coordination in a combinatory grammar. In Baltin, M. R. and Kroch, A. S. (Eds.), *Alternative Conceptions of Phrase Structure*, pp. 201–231. University of Chicago, Chicago.

Steedman, M. (1996). *Surface Structure and Interpretation*. MIT Press. Linguistic Inquiry Monograph, 30.

Steedman, M. (2000). *The Syntactic Process*. The MIT Press.

Steedman, M. (2003). Information-structural semantics for English intonation..

Steedman, M. and Baldridge, J. (2007). Combinatory categorial grammar. In Borsley, R. and Borjars, K. (Eds.), *Non-Transformational Syntax*. Blackwell.

Stent, A. (2002). A conversation acts model for generating spoken dialogue contributions. *Computer Speech and Language, Special Issue on Spoken Language Generation*, *16*(3-4).

Stetina, J. and Nagao, M. (1997). Corpus based PP attachment ambiguity resolution with a semantic dictionary. In Zhou, J. and Church, K. W. (Eds.), *Proceedings of the Fifth Workshop on Very Large Corpora*, Beijing, China, pp. 66–80.

Stevens, K. N., Kasowski, S., and Fant, G. M. (1953). An electrical analog of the vocal tract. *Journal of the Acoustical Society of America*, *25*(4), 734–742.

Stevens, K. N. (1998). *Acoustic Phonetics*. MIT Press.

Stevens, K. N. and House, A. S. (1955). Development of a quantitative description of vowel articulation. *Journal of the Acoustical Society of America*, *27*, 484–493.

Stevens, K. N. and House, A. S. (1961). An acoustical theory of vowel production and some of its implications. *Journal of Speech and Hearing Research*, *4*, 303–320.

Stevens, S. S. and Volkmann, J. (1940). The relation of pitch to frequency: A revised scale. *The American Journal of Psychology*, *53*(3), 329–353.

Stevens, S. S., Volkmann, J., and Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, *8*, 185–190.

Stevenson, M. and Wilks, Y. (2001). The interaction of knowledge sources in word sense disambiguation. *Computational Linguistics*, *27*(3), 321–349.

Stevenson, S. and Merlo, P. (1999). Automatic verb classification using distributions of grammatical features. In *EACL-99*, Bergen, Norway, pp. 45–52.

Stifelman, L. J., Arons, B., Schmandt, C., and Hulteen, E. A. (1993). VoiceNotes: A speech interface for a hand-held voice notetaker. In *Human Factors in Computing Systems: INTER-CHI '93 Conference Proceedings*, Amsterdam, pp. 179–186. ACM.

Stolcke, A. (1995). An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational Linguistics*, *21*(2), 165–202.

Stolcke, A. (1998). Entropy-based pruning of backoff language models. In *Proc. DARPA Broadcast News Transcription and Understanding Workshop*, Lansdowne, VA, pp. 270–274.

Stolcke, A. (2002). Srilm - an extensible language modeling toolkit. In *ICSLP-02*, Denver, CO.

Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martina, R., Meteer, M., and Van Ess-Dykema, C. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, *26*, 339–371.

Stolcke, A., Shriberg, E., Bates, R., Coccaro, N., Jurafsky, D., Martin, R., Meteer, M., Ries, K., Taylor, P., and Van Ess-Dykema, C. (1998). Dialog act modeling for conversational speech. In Chu-Carroll, J. and Green, N. (Eds.), *Applying Machine Learning to Discourse Processing. Papers from the 1998 AAAI Spring Symposium. Tech. rep. SS-98-01*, Stanford, CA, pp. 98–105. AAAI Press.

Stolcke, A. and Shriberg, E. (1996). Statistical language modeling for speech disfluencies. In *IEEE ICASSP-96*, Atlanta, GA, Vol. 1, pp. 405–408. IEEE.

Stolz, W. S., Tannenbaum, P. H., and Carstensen, F. V. (1965). A stochastic approach to the grammatical coding of English. *Communications of the ACM*, *8*(6), 399–405.

Streeter, L. (1978). Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America*, *63*, 1582–1592.

Strube, M. and Ponzetto, S. P. (2006). WikiRelate! Computing semantic relatedness using Wikipedia. In *AAAI-06*, pp. 1419–1424.

Strube, M. and Hahn, U. (1996). Functional centering. In *ACL-96*, Santa Cruz, CA, pp. 270–277.

Strzalkowski, T. and Harabagiu, S. (Eds.). (2006). *Advances in Open Domain Question Answering*. Springer.

Sundheim, B. (Ed.). (1991). *Proceedings of the Third Message Understanding Conference*.

Sundheim, B. (Ed.). (1992). *Proceedings of the Fourth Message Understanding Conference*.

Sundheim, B. (Ed.). (1993). *Proceedings, Fifth Message Understanding Conference (MUC-5), Baltimore, MD*. Morgan Kaufmann.

Sundheim, B. (1995a). Overview of results of the MUC-6 evaluation. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, Columbia, MD, pp. 13–31.

Sundheim, B. (Ed.). (1995b). *Proceedings of the Sixth Message Understanding Conference*.

Surdeanu, M., Harabagiu, S., Williams, J., and Aarseth, P. (2003). Using predicate-argument structures for information extraction. In *ACL-03*, pp. 8–15.

Sutton, C. and McCallum, A. (2006). An introduction to conditional random fields for relational learning. In Getoor, L. and Taskar, B. (Eds.), *Introduction to Statistical Relational Learning*. MIT Press.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Bradford Books (MIT Press).

Sweet, H. (1877). *A Handbook of Phonetics*. Clarendon Press, Oxford.

Swerts, M., Litman, D. J., and Hirschberg, J. (2000). Corrections in spoken dialogue systems. In *ICSLP-00*, Beijing, China.

Swier, R. and Stevenson, S. (2004). Unsupervised semantic role labelling. In *EMNLP 2004*, pp. 95–102.

Syrdal, A. K. and Conkie, A. (2004). Data-driven perceptually based join costs. In *Proceedings of Fifth ISCA Speech Synthesis Workshop*.

Syrdal, A. K., Wightman, C. W., Conkie, A., Stylianou, Y., Beutnagel, M., Schroeter, J., Strom, V., and Lee, K.-S. (2000). Corpus-based techniques in the AT&T NEXTGEN synthesis system. In *ICSLP-00*, Beijing.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Shopen, T. (Ed.), *Language Typology and Syntactic Description, Volume 3*. Cambridge University Press. Originally appeared as UC Berkeley Cognitive Science Program Report No. 30, 1980.

Talmy, L. (1991). Path to realization: a typology of event conflation. In *BLS-91*, Berkeley, CA, pp. 480–519.

Taskar, B., Klein, D., Collins, M., Koller, D., and Manning, C. D. (2004). Max-margin parsing. In *EMNLP 2004*.

Taylor, P. (2000). Analysis and synthesis of intonation using the Tilt model. *Journal of the Acoustical Society of America*, *107*(3), 1697–1714.

Taylor, P. (2005). Hidden Markov Models for grapheme to phoneme conversion. In *INTERSPEECH-05*, Lisbon, Portugal, pp. 1973–1976.

Taylor, P. (2008). Text-to-speech synthesis. Manuscript.

Taylor, P. and Black, A. W. (1998). Assigning phrase breaks from part of speech sequences. *Computer Speech and Language*, *12*, 99–117.

Taylor, P. and Isard, S. (1991). Automatic diphone segmentation. In *EUROSPEECH-91*, Genova, Italy.

Taylor, P., King, S., Isard, S., and Wright, H. (1998). Intonation and dialog context as constraints for speech recognition. *Language and Speech*, *41*(3-4), 489–508.

Taylor, W. L. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, *30*, 415–433.

Taylor, W. L. (1957). Cloze readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology*, *4*, 19–26.

ter Meulen, A. (1995). *Representing Time in Natural Language*. MIT Press.

Teranishi, R. and Umeda, N. (1968). Use of pronouncing dictionary in speech synthesis experiments. In *6th International Congress on Acoustics*, Tokyo, Japan, pp. B155–158.

Tesar, B. (2006a). Faithful contrastive features in learning. *Cognitive Science*, *30*(5), 863–903.

Tesar, B. (2006b). Learning from paradigmatic information..

Tesar, B. and Prince, A. (2007). Using phonotactics to learn phonological alternations. In *CLS 39*, pp. 200–213.

Tesar, B. and Smolensky, P. (2000). *Learning in Optimality Theory*. MIT Press.

Tesnière, L. (1959). *Éléments de Syntaxe Structurale*. Librairie C. Klincksieck, Paris.

Teufel, S. and van Halteren, H. (2004). Evaluating information content by factoid analysis: human annotation and stability. In *EMNLP 2004*, Barcelona.

Teufel, S. and Moens, M. (2002). Summarizing scientific articles: experiments with relevance and rhetorical status. *Computational Linguistics*, *28*(4), 409–445.

Thede, S. M. and Harper, M. P. (1999). A second-order Hidden Markov Model for part-of-speech tagging. In *ACL-99*, College Park, MA, pp. 175–182.

Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*, 53–71.

Thiessen, E. D. and Saffran, J. R. (2004). Spectral tilt as a cue to word segmentation in infancy and adulthood. *Perception and Psychophysics*, *66*(2), 779–791.

Thompson, K. (1968). Regular expression search algorithm. *Communications of the ACM*, *11*(6), 419–422.

Tillmann, C. (2003). A projection extension algorithm for statistical machine translation. In *EMNLP 2003*, Sapporo, Japan.

Titone, D. and Connine, C. M. (1997). Syllabification strategies in spoken word processing: Evidence from phonological priming. *Psychological Research*, *60*(4), 251–263.

Titov, I. and Henderson, J. (2006). Loss minimization in parse reranking. In *EMNLP 2006*.

Titov, I. and Henderson, J. (2007). A latent variable model for generative dependency parsing. In *IWPT-07*.

Tjong Kim Sang, E. F. and Veenstra, J. (1999). Representing text chunks. In *Proceedings of EACL 1999*, pp. 173–179.

Tokuda, K., Kobayashi, T., and Imai, S. (1995a). Speech parameter generation from HMM using dynamic features. In *IEEE ICASSP-95*.

Tokuda, K., Masuko, T., and Yamada, T. (1995b). An algorithm for speech parameter generation from continuous mixture HMMs with dynamic features. In *EUROSPEECH-95*, Madrid.

Tokuda, K., Yoshimura, T., Masuko, T., Kobayashi, T., and Kitamura, T. (2000). Speech parameter generation algorithms for HMM-based speech synthesis. In *IEEE ICASSP-00*.

Tokuda, K., Zen, H., and Kitamura, T. (2003). Trajectory modeling based on HMMs with the explicit relationship between static and dynamic features. In *EUROSPEECH-03*.

Tomabechi, H. (1991). Quasi-destructive graph unification. In *ACL-91*, Berkeley, CA, pp. 315–322.

Tomokiyo, L. M. (2001). *Recognizing non-native speech: Characterizing and adapting to non-native usage in speech recognition*. Ph.D. thesis, Carnegie Mellon University.

Tomokiyo, L. M. and Waibel, A. (2001). Adaptation methods for non-native speech. In *Proceedings of Multilinguality in Spoken Language Processing*, Aalborg, Denmark.

Touretzky, D. S., Elvgren III, G., and Wheeler, D. W. (1990). Phonological rule induction: An architectural solution. In *COGSCI-90*, pp. 348–355.

Toutanova, K., Manning, C. D., Flickinger, D., and Oepen, S. (2005). Stochastic HPSG Parse Disambiguation using the Redwoods Corpus. *Research on Language & Computation*, *3*(1), 83–105.

Toutanova, K., Ilhan, H. T., and Manning, C. D. (2002). Extensions to HMM-based statistical word alignment models. In *EMNLP 2002*, pp. 87–94.

Toutanova, K., Klein, D., Manning, C. D., and Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *HLT-NAACL-03*.

Toutanova, K. and Moore, R. C. (2002). Pronunciation modeling for improved spelling correction. In *ACL-02*, Philadelphia, PA, pp. 144–151.

Traum, D. R. (2000). 20 questions for dialogue act taxonomies. *Journal of Semantics*, *17*(1).

Traum, D. R. and Hinkelman, E. A. (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence: Special Issue on Computational Approaches to Non-Literal Language*, *8*(3).

Traum, D. R. and Larsson, S. (2000). Information state and dialogue management in the trindi dialogue move engine toolkit. *Natural Language Engineering*, *6*(323-340), 97–114.

Traum, D. R. and Larsson, S. (2003). The information state approach to dialogue management. In van Kuppevelt, J. and Smith, R. (Eds.), *Current and New Directions in Discourse and Dialogue*. Kluwer.

Treiman, R., Bowey, J., and Bourassa, D. (2002). Segmentation of spoken words into syllables by english-speaking children as compared to adults. *Journal of Experimental Child Psychology*, *83*, 213–238.

Trubetskoi, N. S. (1939). *Grundzüge der Phonologie*, Vol. 7 of *Travaux du cercle linguistique de Prague*. Available in 1969 English translation by Christiane A. M. Baltaxe as *Principles of Phonology*, University of California Press.

Tseng, H. (2003). Semantic classification of Chinese unknown words. In *ACL-03*, pp. 72–79.

Tseng, H., Chang, P., Andrew, G., Jurafsky, D., and Manning, C. D. (2005a). Conditional random field word segmenter. In *Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing*.

Tseng, H., Jurafsky, D., and Manning, C. D. (2005b). Morphological features help pos tagging of unknown words across language varieties. In *Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing*.

Tsujii, J. (1986). Future directions of machine translation. In *COLING-86*, Bonn, pp. 655–668.

Turian, J. P., Shen, L., and Melamed, I. D. (2003). Evaluation of machine translation and its evaluation. In *Proceedings of MT Summit IX*, New Orleans, LA.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *42*, 230–265. Read to the Society in 1936, but published in 1937. Correction in volume 43, 544–546.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*, 433–460.

Turner, J. and Charniak, E. (2005). Supervised and unsupervised learning for sentence compression. In *ACL-05*, pp. 290–297.

Turney, P., Littman, M., Bigham, J., and Shnayder, V. (2003). Combining independent modules to solve multiple-choice synonym and analogy problems. In *Proceedings of RANLP-03*, Borovets, Bulgaria, pp. 482–489.

Tyler, L. K. (1984). The structure of the initial cohort: Evidence from gating. *Perception & Psychophysics*, *36*(5), 417–427.

Umeda, N., Matui, E., Suzuki, T., and Omura, H. (1968). Synthesis of fairy tale using an analog vocal tract. In *6th International Congress on Acoustics*, Tokyo, Japan, pp. B159–162.

Umeda, N. (1976). Linguistic rules for text-to-speech synthesis. *Proceedings of the IEEE*, *64*(4), 443–451.

Usher, A. P. (1954). *A History of Mechanical Inventions*. Harvard University Press.

Uszkoreit, H. (1986). Categorial unification grammars. In *COLING-86*, Bonn, pp. 187–194.

van Benthem, J. and ter Meulen, A. (Eds.). (1997). *Handbook of Logic and Language*. MIT Press.

van Deemter, K. and Kibble, R. (2000). On coreferring: coreference in muc and related annotation schemes. *Computational Linguistics*, *26*(4), 629–637.

van den Bosch, A. (1997). *Learning to Pronounce Written Words: A Study in Inductive Language Learning*. Ph.D. thesis, University of Maastricht, Maastricht, The Netherlands.

van Halteren, H. and Teufel, S. (2003). Examining the consensus between human summaries: initial experiments with factoid analysis. In *HLT-NAACL-03 Workshop on Text Summarization*.

van Halteren, H. (Ed.). (1999). *Syntactic Wordclass Tagging*. Kluwer, Dordrecht.

van Lehn, K. (1978). Determining the scope of English quantifiers. Master's thesis, MIT, Cambridge, MA. MIT Technical Report AI-TR-483.

van Rijsbergen, C. J. (1975). *Information Retrieval*. Butterworths, London.

van Santen, J. P. H. (1994). Assignment of segmental duration in text-to-speech synthesis. *Computer Speech and Language*, *8*(95–128).

van Santen, J. P. H. (1997). Segmental duration and speech timing. In Sagisaka, Y., Campbell, N., and Higuchi, N. (Eds.), *Computing Prosody: Computational Models for Processing Spontaneous Speech*. Springer.

van Santen, J. P. H. (1998). Timing. In Sproat, R. (Ed.), *Multilingual Text-To-Speech Synthesis: The Bell Labs Approach*, pp. 115–140. Kluwer, Dordrecht.

van Santen, J. P. H. and Sproat, R. (1998). Methods and tools. In Sproat, R. (Ed.), *Multilingual Text-To-Speech Synthesis: The Bell Labs Approach*, pp. 7–30. Kluwer, Dordrecht.

van Santen, J. P. H., Sproat, R., Olive, J. P., and Hirschberg, J. (Eds.). (1997). *Progress in Speech Synthesis*. Springer.

Van Son, R. J. J. H., Koopmans-van Beinum, F. J., and Pols, L. C. W. (1998). Efficiency as an organizing principle of natural speech. In *ICSLP-98*, Sydney.

Van Son, R. J. J. H. and Pols, L. C. W. (2003). How efficient is speech?. *Proceedings of the Institute of Phonetic Sciences*, *25*, 171–184.

Van Valin, Jr., R. D. and La Polla, R. (1997). Syntax: Structure, meaning, and function..

Vanderwende, L. (1994). Algorithm for the automatic interpretation of noun sequences. In *COLING-94*, Kyoto, pp. 782–788.

Vanderwende, L., Suzuki, H., Brockett, C., and Nenkova, A. (2007a). Beyond sumbasic: Task-focused summarization with sentence simplification and lexical expansion. *Information Processing and Management*, *43*(6), 1606–1618.

Vanderwende, L., Suzuki, H., Brockett, C., and Nenkova, A. (2007b). Beyond sumbasic: Task-focused summarization with sentence simplification and lexical expansion. *Information Processing and Management, Special issue on summarization*, *43*(6).

VanLehn, K., Jordan, P. W., Rosé, C., Bhembe, D., Böttner, M., Gaydos, A., Makatchev, M., Pappuswamy, U., Ringenberg, M., Roque, A., Siler, S., Srivastava, R., and Wilson, R. (2002). The architecture of Why2-Atlas: A coach for qualitative physics essay writing. In *Proc. Intelligent Tutoring Systems*.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag.

Vasilescu, F., Langlais, P., and Lapalme, G. (2004). Evaluating variants of the lesk approach for disambiguating words. In *LREC-04*, Lisbon, Portugal, pp. 633–636. ELRA.

Vauquois, B. (1968). A survey of formal grammars and algorithms for recognition and transformation in machine translation. In *IFIP Congress 1968*, Edinburgh, pp. 254–260.

Veale, T. and Keane, M. T. (1992). Conceptual scaffolding: A spatially founded meaning representation for metaphor comprehension. *Computational Intelligence*, *8*(3), 494–519.

Veblen, T. (1899). *Theory of the Leisure Class*. Macmillan Company, New York.

Veldhuis, R. (2000). Consistent pitch marking. In *ICSLP-00*, Beijing, China.

Velichko, V. M. and Zagoruyko, N. G. (1970). Automatic recognition of 200 words. *International Journal of Man-Machine Studies*, *2*, 223–234.

Venditti, J. J. (2005). The j_tobi model of japanese intonation. In Jun, S.-A. (Ed.), *Prosodic Typology and Transcription: A Unified Approach*. Oxford University Press.

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY.

Venugopal, A., Vogel, S., and Waibel, A. (2003). Effective phrase translation extraction from alignment models. In *ACL-03*, pp. 319–326.

Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Katz, G., and Pustejovsky, J. (2007). Semeval-2007 task 15: Tempeval temporal relation identification. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, Czech Republic, pp. 75–80.

Veronis, J. and Ide, N. M. (1990). Word sense disambiguation with very large neural networks extracted from machine readable dictionaries. In *COLING-90*, Helsinki, Finland, pp. 389–394.

Verspoor, C. M., Joslyn, C., and Papcun, G. J. (2003). The gene ontology as a source of lexical semantic knowledge for a biological natural language processing application. In *Proceedings of the SIGIR'03 Workshop on Text Analysis and Search for Bioinformatics*, Toronto, CA.

Vieira, R. and Poesio, M. (2000). An empirically based system for processing definite descriptions. *Computational Linguistics*, *26*(4), 539–593.

Vilain, M., Burger, J. D., Aberdeen, J., Connolly, D., and Hirschman, L. (1995). A model-theoretic coreference scoring scheme. In *MUC6 '95: Proceedings of the 6th Conference on Message Understanding*.

Vintsyuk, T. K. (1968). Speech discrimination by dynamic programming. *Cybernetics*, *4*(1), 52–57. Russian Kibernetika 4(1):81-88 (1968).

Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, *IT-13*(2), 260–269.

Voiers, W., Sharpley, A., and Hehmsoth, C. (1975). Research on diagnostic evaluation of speech intelligibility. Research Report AFCRL-72-0694.

von Neumann, J. (1963). *Collected Works: Volume V*. Macmillan Company, New York.

Voorhees, E. M. (1998). Using WordNet for text retrieval. In Fellbaum, C. (Ed.), *WordNet: An Electronic Lexical Database*, pp. 285–303. MIT Press.

Voorhees, E. M. and Harman, D. K. (2005). *TREC: Experiment and Evaluation in Information Retrieval*. MIT Press.

Voorhees, E. M. and Tice, D. M. (1999). The TREC-8 question answering track evaluation. Proceedings of the TREC-8 Workshop.

Voutilainen, A. (1995). Morphological disambiguation. In Karlsson, F., Voutilainen, A., Heikkilä, J., and Anttila, A. (Eds.), *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*, pp. 165–284. Mouton de Gruyter, Berlin.

Voutilainen, A. (1999). Handcrafted rules. In van Halteren, H. (Ed.), *Syntactic Wordclass Tagging*, pp. 217–246. Kluwer, Dordrecht.

Wade, E., Shriberg, E., and Price, P. J. (1992). User behaviors affecting speech recognition. In *ICSLP-92*, pp. 995–998.

Wagner, R. A. and Fischer, M. J. (1974). The string-to-string correction problem. *Journal of the Association for Computing Machinery*, *21*, 168–173.

Waibel, A. (1988). *Prosody and Speech Recognition*. Morgan Kaufmann.

Wald, B. and Shopen, T. (1981). A researcher's guide to the sociolinguistic variable (ING). In Shopen, T. and Williams, J. M. (Eds.), *Style and Variables in English*, pp. 219–249. Winthrop Publishers.

Walker, M. A., Fromer, J. C., and Narayanan, S. S. (1998). Learning optimal dialogue strategies: a case study of a spoken dialogue agent for email. In *COLING/ACL-98*, Montreal, Canada, pp. 1345–1351.

Walker, M. A., Iida, M., and Cote, S. (1994). Japanese discourse and the process of centering. *Computational Linguistics*, *20*(2).

Walker, M. A., Joshi, A. K., and Prince, E. (Eds.). (1998). *Centering in Discourse*. Oxford University Press.

Walker, M. A., Kamm, C. A., and Litman, D. J. (2001). Towards developing general models of usability with PAR-ADISE. *Natural Language Engineering: Special Issue on Best Practice in Spoken Dialogue Systems*, *6*(3).

Walker, M. A., Litman, D. J., Kamm, C. A., and Abella, A. (1997). PARADISE: A framework for evaluating spoken dialogue agents. In *ACL/EACL-97*, Madrid, Spain, pp. 271–280.

Walker, M. A., Maier, E., Allen, J., Carletta, J., Condon, S., Flammia, G., Hirschberg, J., Isard, S., Ishizaki, M., Levin, L., Luperfoy, S., Traum, D. R., and Whittaker, S. (1996). Penn multiparty standard coding scheme: Draft annotation manual. `www.cis.upenn.edu/˜ircs/discourse-tagging/newcoding.html`.

Walker, M. A., Passonneau, R., Rudnicky, A. I., Aberdeen, J., Boland, J., Bratt, E., Garofolo, J., Hirschman, L., Le, A., Lee, S., Narayanan, S. S., Papineni, K., Pellom, B., Polifroni, J., Potamianos, A., Prabhu, P., Rudnicky, A. I., Sanders, G., Seneff, S., Stallard, D., and Whittaker, S. (2002). Cross-site evaluation in DARPA Communicator: The June 2000 data collection. submitted.

Walker, M. A. and Rambow, O. (2002). Spoken language generation. *Computer Speech and Language, Special Issue on Spoken Language Generation*, *16*(3-4), 273–281.

Walker, M. A. and Whittaker, S. (1990). Mixed initiative in dialogue: An investigation into discourse segmentation. In *ACL-90*, Pittsburgh, PA, pp. 70–78.

Walker, M. A. *et al*. (2001). Cross-site evaluation in darpa communicator: The june 2000 data collection. Submitted ms.

Wang, M. Q. and Hirschberg, J. (1992). Automatic classification of intonational phrasing boundaries. *Computer Speech and Language*, *6*(2), 175–196.

Wang, Y. Y. and Waibel, A. (1997). Decoding algorithm in statistical machine translation. In *ACL/EACL-97*, pp. 366–372.

Wang, Z., Schultz, T., and Waibel, A. (2003). Comparison of acoustic model adaptation techniques on non-native speech. In *IEEE ICASSP*, Vol. 1, pp. 540–543.

Wanner, E. and Maratsos, M. (1978). An ATN approach to comprehension. In Halle, M., Bresnan, J., and Miller, G. A. (Eds.), *Linguistic Theory and Psychological Reality*, pp. 119–161. MIT Press.

Ward, N. (1994). *A Connectionist Language Generator*. Ablex.

Ward, N. and Tsukahara, W. (2000). Prosodic features which cue back-channel feedback in English and Japanese. *Journal of Pragmatics*, *32*, 1177–1207.

Ward, W. (1989). Modelling non-verbal sounds for speech recognition. In *HLT '89: Proceedings of the Workshop on Speech and Natural Language*, Cape Cod, MA, pp. 47–50.

Ward, W. and Issar, S. (1994). Recent improvements in the cmu spoken language understanding system. In *ARPA Human Language Technologies Workshop*, Plainsboro, N.J.

Warnke, V., Kompe, R., Niemann, H., and Nöth, E. (1997). Integrated dialog act segmentation and classification using prosodic features and language models. In *EUROSPEECH-97*, Vol. 1, pp. 207–210.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 392–393.

Warwick, S. and Russell, G. (1990). Bilingual concordancing and bilingual lexicography. In *EURALEX 4th International Congress*.

Wattarujeekrit, T., Shah, P. K., and Collier, N. (2004). PAS-Bio: predicate-argument structures for event extraction in molecular biology. *BMC Bioinformatics*, *5*(155).

Waugh, L. R. (1976). The semantics and paradigmatics of word order. *Language*, *52*(1), 82–107.

Weaver, W. (1949/1955). Translation. In Locke, W. N. and Boothe, A. D. (Eds.), *Machine Translation of Languages*, pp. 15–23. MIT Press. Reprinted from a memorandum written by Weaver in 1949.

Webber, B. L. (2004). D-LTAG: extending lexicalized TAG to discourse. *Cognitive Science*, *28*(5), 751–79.

Webber, B. L., Knott, A., Stone, M., and Joshi, A. K. (1999). Discourse relations: A structural and presuppositional account using lexicalised TAG. In *ACL-99*, College Park, MD, pp. 41–48.

Webber, B. L. (1978). *A Formal Approach to Discourse Anaphora*. Ph.D. thesis, Harvard University.

Webber, B. L. (1983). So what can we talk about now?. In Brady, M. and Berwick, R. C. (Eds.), *Computational Models of Discourse*, pp. 331–371. The MIT Press. Reprinted in Grosz et al. (1986).

Webber, B. L. (1991). Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, *6*(2), 107–135.

Weber, D. J., Black, H. A., and McConnel, S. R. (1988). AMPLE: A tool for exploring morphology. Tech. rep. Occasional Publications in Academic Computing No. 12, Summer Institute of Linguistics, Dallas.

Weber, D. J. and Mann, W. C. (1981). Prospects for computer-assisted dialect adaptation. *American Journal of Computational Linguistics*, *7*, 165–177. Abridged from Summer Institute of Linguistics *Notes on Linguistics* Special Publication 1, 1979.

Weeds, J. (2003). *Measures and Applications of Lexical Distributional Similarity*. Ph.D. thesis, University of Sussex.

Weeds, J. and Weir, D. J. (2005). Co-occurrence retrieval: a general framework for lexical distributional similarity. *Computational Linguistics*, *31*(4), 439–476.

Wegmann, S., McAllaster, D., Orloff, J., and Peskin, B. (1996). Speaker normalisation on conversational telephone speech. In *IEEE ICASSP-96*, Atlanta, GA.

Weinschenk, S. and Barker, D. T. (2000). *Designing effective speech interfaces*. Wiley.

Weintraub, M., Taussig, K., Hunicke-Smith, K., and Snodgras, A. (1996). Effect of speaking style on LVCSR performance. In *ICSLP-96*, Philadelphia, PA, pp. 16–19.

Weischedel, R. (1995). BBN: Description of the PLUM system as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, San Francisco, pp. 55–70. Morgan Kaufmann.

Weischedel, R., Meteer, M., Schwartz, R., Ramshaw, L. A., and Palmucci, J. (1993). Coping with ambiguity and unknown words through probabilistic models. *Computational Linguistics*, *19*(2), 359–382.

Weizenbaum, J. (1966). ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgement to Calculation*. W.H. Freeman and Company, San Francisco.

Welling, L., Ney, H., and Kanthak, S. (2002). Speaker adaptive modeling by vocal tract normalisation. *IEEE Transactions on Speech and Audio Processing*, *10*, 415–426.

Wells, J. C. (1982). *Accents of English*. Cambridge University Press.

Weng, F., Varges, S., Raghunathan, B., Ratiu, F., Pon-Barry, H., Lathrop, B., Zhang, Q., Scheideck, T., Bratt, H., Xu, K., Purver, M., Mishra, R., Raya, M., Peters, S., Meng, Y., Cavedon, L., and Shriberg, E. (2006). Chat: A conversational helper for automotive tasks. In *ICSLP-06*, pp. 1061–1064.

Wierzbicka, A. (1992). *Semantics, Culture, and Cognition: University Human Concepts in Culture-Specific Configurations*. Oxford University Press.

Wierzbicka, A. (1996). *Semantics: Primes and Universals*. Oxford University Press.

Wilensky, R. (1983). *Planning and Understanding: A Computational Approach to Human Reasoning*. Addison-Wesley.

Wilensky, R. and Arens, Y. (1980). PHRAN: A knowledge-based natural language understander. In *ACL-80*, Philadelphia, PA, pp. 117–121.

Wilks, Y. (1975a). An intelligent analyzer and understander of English. *Communications of the ACM*, *18*(5), 264–274.

Wilks, Y. (1975b). Preference semantics. In Keenan, E. L. (Ed.), *The Formal Semantics of Natural Language*, pp. 329–350. Cambridge Univ. Press.

Wilks, Y. (1975c). A preferential, pattern-seeking, semantics for natural language inference. *Artificial Intelligence*, *6*(1), 53–74.

Wilks, Y. (1978). Making preferences more active. *Artificial Intelligence*, *11*(3), 197–223.

Wilks, Y., Slator, B. M., and Guthrie, L. M. (1996). *Electric Words: Dictionaries, Computers, and Meanings*. MIT Press.

Williams, J. D. and Young, S. J. (2000). Partially observable markov decision processes for spoken dialog systems. *Computer Speech and Language*, *21*(1), 393–422.

Williams, J. D. and Young, S. J. (2005). Scaling up pomdps for dialog management: The "summary pomdp" method. In *IEEE ASRU-05*.

Winograd, T. (1972a). Understanding natural language. *Cognitive Psychology*, *3*(1), 1–191. Reprinted as a book by Academic Press, 1972.

Winograd, T. (1972b). *Understanding Natural Language*. Academic Press.

Wise, B., Cole, R. A., van Vuuren, S., Schwartz, S., Snyder, L., Ngampatipatpong, N., Tuantranont, J., and Pellom, B. (2007). Learning to Read with a Virtual Tutor: Foundations to Literacy. In Kinzer, C. and Verhoeven, L. (Eds.), *Interactive Literacy Education: Facilitating Literacy Environments Through Technology*. Lawrence Erlbaum.

Witten, I. H. and Bell, T. C. (1991). The zero-frequency problem: Estimating the probabilities of novel events in adaptive text compression. *IEEE Transactions on Information Theory*, *37*(4), 1085–1094.

Witten, I. H. and Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann. 2nd ed.

Wittgenstein, L. (1953). *Philosophical Investigations. (Translated by Anscombe, G.E.M.)*. Blackwell, Oxford.

Wolf, F. and Gibson, E. (2005). Representing discourse coherence: A corpus-based analysis. *Computational Linguistics*, *31*(2), 249–287.

Wolfram, W. A. (1969). *A Sociolinguistic Description of Detroit Negro Speech*. Center for Applied Linguistics, Washington, D.C.

Woodland, P. C., Leggetter, C. J., Odell, J. J., Valtchev, V., and Young, S. J. (1995). The 1994 htk large vocabulary speech recognition system. In *IEEE ICASSP*.

Woodland, P. C. and Povey, D. (2002). Large scale discriminative training of hidden Markov models for speech recognition. *Computer Speech and Language,*, *16*, 25–47.

Woodland, P. C. (2001). Speaker adaptation for continuous density HMMs: A review. In Juncqua, J.-C. and Wellekens, C. (Eds.), *Proceedings of the ITRW 'Adaptation Methods For Speech Recognition'*, Sophia-Antipolis, France.

Woods, W. A. (1967). *Semantics for a Question-Answering System*. Ph.D. thesis, Harvard University.

Woods, W. A. (1973). Progress in natural language understanding. In *Proceedings of AFIPS National Conference*, pp. 441–450.

Woods, W. A. (1975). What's in a link: Foundations for semantic networks. In Bobrow, D. G. and Collins, A. M. (Eds.), *Representation and Understanding: Studies in Cognitive Science*, pp. 35–82. Academic Press.

Woods, W. A. (1977). Lunar rocks in natural English: Explorations in natural language question answering. In Zampolli, A. (Ed.), *Linguistic Structures Processing*, pp. 521–569. North Holland, Amsterdam.

Woods, W. A. (1978). Semantics and quantification in natural language question answering. In Yovits, M. (Ed.), *Advances in Computers*, Vol. 17, pp. 2–87. Academic Press.

Woods, W. A., Kaplan, R. M., and Nash-Webber, B. L. (1972a). The Lunar Sciences Natural Language Information System: Final report. Tech. rep. 2378, Bolt, Beranek, and Newman, Inc., Cambridge, MA.

Woods, W. A., Kaplan, R. M., and Nash-Webber, B. L. (1972b). The lunar sciences natural language information system: Final report.. BBN Report 2378.

Woszczyna, M. and Waibel, A. (1994). Inferring linguistic structure in spoken language. In *ICSLP-94*, Yokohama, Japan, pp. 847–850.

Wouters, J. and Macon, M. (1998). Perceptual evaluation of distance measures for concatenative speech synthesis. In *ICSLP-98*, Sydney, pp. 2747–2750.

Wu, D. (1992). *Automatic Inference: A Probabilistic Basis for Natural Language Interpretation*. Ph.D. thesis, University of California, Berkeley, Berkeley, CA. UCB/CSD 92-692.

Wu, D. (1996). A polynomial-time algorithm for statistical machine translation. In *ACL-96*, Santa Cruz, CA, pp. 152–158.

Wu, D. (2000). Bracketing and aligning words and constituents in parallel text using stochastic inversion transduction grammars. In Veronis, J. (Ed.), *Parallel Text Processing: Alignment and Use of Translation Corpora*. Kluwer, Dordrecht.

Wu, D. and Fung, P. (2005). Inversion transduction grammar constraints for mining parallel sentences from quasi-comparable corpora. In *IJCNLP-2005*, Jeju, Korea.

Wu, Z. and Palmer, M. (1994). Verb semantics and lexical selection. In *ACL-94*, Las Cruces, NM, pp. 133–138.

Wundt, W. (1900). *Völkerpsychologie: eine Untersuchung der Entwicklungsgesetze von Sprache, Mythus, und Sitte*. W. Engelmann, Leipzig. Band II: Die Sprache, Zweiter Teil.

Xia, F. and Palmer, M. (2001). Converting dependency structures to phrase structures. In *HLT-01*, San Diego, pp. 1–5.

Xu, P., Chelba, C., and Jelinek, F. (2002). A study on richer syntactic dependencies for structured language modeling. In *ACL-02*, pp. 191–198.

Xu, W. and Rudnicky, A. I. (2000). Task-based dialog management using an agenda. In *ANLP/NAACL Workshop on Conversational Systems*, Somerset, New Jersey, pp. 42–47.

Xue, N. and Palmer, M. (2004). Calibrating features for semantic role labeling. In *EMNLP 2004*.

Xue, N. and Shen, L. (2003). Chinese word segmentation as lmr tagging. In *Proceedings of the 2nd SIGHAN Workshop on Chinese Language Processing*, Sapporo, Japan.

Yamada, H. and Matsumoto, Y. (2003). Statistical dependency analysis with support vector machines. In Noord, G. V. (Ed.), *IWPT-03*, pp. 195–206.

Yamada, K. and Knight, K. (2001). A syntax-based statistical translation model. In *ACL-01*, Toulouse, France.

Yamada, K. and Knight, K. (2002). A decoder for syntax-based statistical MT. In *ACL-02*, Philadelphia, PA.

Yankelovich, N., Levow, G.-A., and Marx, M. (1995). Designing SpeechActs: Issues in speech user interfaces. In *Human Factors in Computing Systems: CHI '95 Conference Proceedings*, Denver, CO, pp. 369–376. ACM.

Yarowsky, D. and Wicentowski, R. (2000). Minimally supervised morphological analysis by multimodal alignment. In *ACL-00*, Hong Kong, pp. 207–216.

Yarowsky, D. (1994). Decision lists for lexical ambiguity resolution: Application to accent restoration in Spanish and French. In *ACL-94*, Las Cruces, NM, pp. 88–95. ACL.

Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods. In *ACL-95*, Cambridge, MA, pp. 189–196.

Yarowsky, D. (1997). Homograph disambiguation in text-to-speech synthesis. In van Santen, J. P. H., Sproat, R., Olive, J. P., and Hirschberg, J. (Eds.), *Progress in Speech Synthesis*, pp. 157–172. Springer.

Yeh, A., Morgan, A., Colosimo, M. E., and Hirschman, L. (2005). BioCreatve task 1A: gene mention finding evaluation. *BMC Bioinformatics*, *6*(1).

Yngve, V. H. (1955). Syntax and the problem of multiple meaning. In Locke, W. N. and Booth, A. D. (Eds.), *Machine Translation of Languages*, pp. 208–226. MIT Press.

Yngve, V. H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, *104*, 444–466.

Yngve, V. H. (1970). On getting a word in edgewise. In *CLS-70*, pp. 567–577. University of Chicago.

Young, M. and Rounds, W. C. (1993). A logical semantics for nonmonotonic sorts. In *ACL-93*, Columbus, OH, pp. 209–215.

Young, S. J. (1984). Generating multiple solutions from connected word dp recognition algorithms. *Proceedings of the Institute of Acoustics*, *6*(4), 351–354.

Young, S. J., Odell, J. J., and Woodland, P. C. (1994). Tree-based state tying for high accuracy acoustic modelling. In *Proceedings ARPA Workshop on Human Language Technology*, pp. 307–312.

Young, S. J., Russell, N. H., and Thornton, J. H. S. (1989). Token passing: A simple conceptual model for connected speech recognition systems. Tech. rep. CUED/F-INFENG/TR.38, Cambridge University Engineering Department, Cambridge, England.

Young, S. J. and Woodland, P. C. (1994). State clustering in HMM-based continuous speech recognition. *Computer Speech and Language*, *8*(4), 369–394.

Young, S. J. (2002). The statistical approach to the design of spoken dialogue systems. Tech. rep. CUED/F-INFENG/TR.433, Cambridge University Engineering Department, Cambridge, England.

Young, S. J., Evermann, G., Gales, M., Hain, T., Kershaw, D., Moore, G., Odell, J. J., Ollason, D., Povey, D., Valtchev, V., and Woodland, P. C. (2005). *The HTK Book*. Cambridge University Engineering Department.

Younger, D. H. (1967). Recognition and parsing of context-free languages in time n^3. *Information and Control*, *10*, 189–208.

Yuan, J., Brenier, J. M., and Jurafsky, D. (2005). Pitch accent prediction: Effects of genre and speaker. In *EUROSPEECH-05*.

Yuret, D. (1998). *Discovery of Linguistic Relations Using Lexical Attraction*. Ph.D. thesis, MIT.

Yuret, D. (2004). Some experiments with a Naive Bayes WSD system. In *Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*.

Zajic, D., Dorr, B., Lin, J., and Schwartz, R. (2007). Multi-candidate reduction: Sentence compression as a tool for document summarization tasks. *Information Processing and Management*, *43*(6), 1549–1570.

Zappa, F. and Zappa, M. U. (1982). Valley girl. From Frank Zappa album *Ship Arriving Too Late To Save A Drowning Witch*.

Zavrel, J. and Daelemans, W. (1997). Memory-based learning: Using similarity for smoothing. In *ACL/EACL-97*, Madrid, Spain, pp. 436–443.

Zens, R. and Ney, H. (2007). Efficient phrase-table representation for machine translation with applications to online MT and speech translation. In *NAACL-HLT 07*, Rochester, NY, pp. 492–499.

Zens, R., Och, F. J., and Ney, H. (2002). Phrase-based statistical machine translation. In *KI 2002*, pp. 18–32.

Zernik, U. (1987). *Strategies in Language Acquisition: Learning Phrases from Examples in Context*. Ph.D. thesis, University of California, Los Angeles, Computer Science Department, Los Angeles, CA.

Zernik, U. (1991). Train1 vs. train2: Tagging word senses in corpus. In *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pp. 91–112. Lawrence Erlbaum.

Zhang, H. and Gildea, D. (2005). Stochastic lexicalized inversion transduction grammar for alignment. In *ACL-05*, Ann Arbor, MI.

Zheng, Y., Sproat, R., Gu, L., Shafran, I., Zhou, H., Su, Y., Jurafsky, D., Starr, R., and Yoon, S.-Y. (2005). Accent detection and speech recognition for Shanghai-accented Mandarin. In *InterSpeech 2005*, Lisbon, Portugal.

Zhou, G. and Lua, K. (1998). Word association and MI-trigger-based language modelling. In *COLING/ACL-98*, Montreal, Canada, pp. 1465–1471.

Zhou, G., Su, J., Zhang, J., and Zhang, M. (2005). Exploring various knowledge in relation extraction. In *ACL-05*, Ann Arbor, MI, pp. 427–434.

Zhu, X. and Rosenfeld, R. (2001). Improving trigram language modeling with the world wide web. In *IEEE ICASSP-01*, Salt Lake City, UT, Vol. I, pp. 533–536.

Zue, V. W., Glass, J., Goodine, D., Leung, H., Phillips, M., Polifroni, J., and Seneff, S. (1989). Preliminary evaluation of the VOYAGER spoken language system. In *Proceedings DARPA Speech and Natural Language Workshop*, Cape Cod, MA, pp. 160–167. Morgan Kaufmann.

Zweig, G. (1998). *Speech Recognition with Dynamic Bayesian Networks*. Ph.D. thesis, University of California, Berkeley.

Zwicky, A. (1972). On Casual Speech. In *CLS-72*, pp. 607–615. University of Chicago.

Zwicky, A. and Sadock, J. (1975). Ambiguity tests and how to fail them. In Kimball, J. (Ed.), *Syntax and Semantics 4*, pp. 1–36. Academic Press.