

Bibliography

HLT-NAACL03

April 23, 2003

References

- [1] Alex Acero. Speech recognition and understanding. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 3–3, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [2] Shazia Akhtar, Ronan G. Reilly, and John Dunnion. Automating xml markup of text documents. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 1–3, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [3] Yaser Al-Onaizan, Radu Florian, Martin Franz, Hany Hassan, Young-Suk Lee, J. Scott McCarley, Kishore Papineni, Salim Roukos, Jeffrey Sorensen, Christoph Tillmann, Todd Ward, and Fei Xia. Tips: A translingual information processing system. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 1–2, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [4] Hiyan Alshawi. Effective utterance classification with unsupervised phonotactic models. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 1–7, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [5] Eiji Aramaki, Sadao Kurohashi, Hideki Kashioka, and Hideki Tanaka. Word selection for ebmt based on monolingual similarity and translation confidence. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 57–64, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [6] Masayuki Asahara and Yuji Matsumoto. Japanese named entity extraction with redundant morphological analysis. In Marti Hearst and Mari

- Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 8–15, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [7] Amittai Axelrod. On building a high performance gazetteer database. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 63–68, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [8] Jason Baldridge and Miles Osborne. Active learning for hpsg parse selection. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 17–24, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [9] Breck Baldwin, Bob Carpenter, and Aaron Ross. Alias-i threat trackers. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 3–4, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [10] Kobus Barnard, Matthew Johnson, and David Forsyth. Word sense disambiguation with pictures. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 1–5, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [11] Regina Barzilay and Lillian Lee. Learning to paraphrase: An unsupervised approach using multiple-sequence alignment. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 16–23, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [12] Oliver Bender, Franz Josef Och, and Hermann Ney. Maximum entropy models for named entity recognition. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 148–151, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [13] Frédéric Bilhaut, Thierry Charnois, Patrice Enjalbert, and Yann Mathet. Geographic reference analysis for geographic document querying. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 55–62, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [14] Jeff A. Bilmes and Katrin Kirchhoff. Factored language models and generalized parallel backoff. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 4–6, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [15] Keith Bonawitz, Anthony Kim, and Seth Tardiff. An architecture for word learning using bidirectional multimodal structural alignment. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 30–37, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [16] Sarah Borys. The importance of prosodic factors in phoneme modeling with applications to speech recognition. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*, pages 7–12, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [17] Ivan Bulyko, Mari Ostendorf, and Andreas Stolcke. Getting more mileage from web text sources for conversational speech language modeling using class-dependent mixtures. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 7–9, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [18] John Burger and John Henderson. Exploiting diversity for answering questions. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 10–12, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [19] Carl Burke, Christy Doran, Abigail Gertner, Andy Gregorowicz, Lisa Harper, Joel Korb, and Dan Loehr. Dialogue complexity with portability? research directions for the information state approach. In Alexander I. Rudnicky and Candace L. Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 13–15, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [20] Chris Callison-Burch and Miles Osborne. Bootstrapping parallel corpora. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 44–49, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [21] Peter Carbonetto and Nando de Freitas. Why can't José read? the problem of learning semantic associations in a robot environment. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003*

Workshop: Learning Word Meaning from Non-Linguistic Data, pages 54–61, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [22] Michael Carl and Sisay Fissaha. Phrase-based evaluation of word-to-word alignments. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 31–35, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [23] Adam Carlson and Steven L. Tanimoto. Learning to identify student preconceptions from text. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 9–16, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [24] Xavier Carreras, Lluís Márquez, and Lluís Padró. Learning a perceptron-based named entity chunker via online recognition feedback. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 156–159, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [25] Xavier Carreras, Lluís Márquez, and Lluís Padró. A simple named entity extractor using adaboost. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 152–155, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [26] Scott Cederberg and Dominic Widdows. Using lsa and noun coordination information to improve the recall and precision of automatic hyponymy extraction. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 111–118, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [27] Joyce Y. Chai, Pengyu Hong, and Michelle X. Zhou. Combining semantic and temporal constraints for multimodal integration in conversation systems. In Alexander I. Rudnicky and Candace L. Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 1–3, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [28] Songsak Channarukul, Susan W. McRoy, and Syed S. Ali. Doghed: A template-based generator for multimodal dialog systems targeting het-

erogeneous devices. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 5–6, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [29] Francine Chen, Ayman Farahat, and Thorsten Brants. Story link detection and new event detection are asymmetric. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 13–15, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [30] Hai Leong Chieu and Hwee Tou Ng. Named entity recognition with a maximum entropy approach. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 160–163, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [31] Jennifer Chu-Carroll, Krzysztof Czuba, John Prager, and Abraham Ittycheriah. In question answering, two heads are better than one. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 24–31, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [32] Grace Chung, Stephanie Seneff, and Chao Wang. Automatic acquisition of names using speak and spell mode in spoken dialogue systems. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 32–39, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [33] Peter Clark, Phil Harrison, and John Thompson. A knowledge-driven approach to text meaning processing. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 1–6, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [34] Stephen Clark, James Curran, and Miles Osborne. Bootstrapping pos-taggers using unlabelled data. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 49–55, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [35] Katri A. Clodfelder. An lsa implementation against parallel texts in french and english. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 111–114, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [36] Sean Colbath and Francis Kubala. Tap-xl: An automated analyst’s assistant. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 7–8, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [37] Cleo Condoravdi, Dick Crouch, Valeria de Paiva, Reinhard Stolle, and Daniel G. Bobrow. Entailment, intensionality and text understanding. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 38–45, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [38] Silviu Cucerzan and David Yarowsky. Minimally supervised induction of grammatical gender. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 40–47, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [39] James Curran and Stephen Clark. Language independent ner using a maximum entropy tagger. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 164–167, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [40] James R. Curran. Blueprint for a high performance nlp infrastructure. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 39–44, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [41] Naomi Daniel, Dragomir Radev, and Timothy Allison. Sub-event based multi-document summarization. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 9–16, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [42] Jody J. Daniels and Helen Wright Hastie. The pragmatics of taking a spoken language system out of the laboratory. In Alexander I. Rudnicky and Candace L. Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 16–18, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [43] Marco De Boni and Suresh Manandhar. An analysis of clarification dialogue for question answering. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 48–55, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [44] Fien De Meulder and Walter Daelemans. Memory-based named entity recognition using unannotated data. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 208–211, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [45] Renato De Mori and Frederic Béchet. Conceptual language models for dialog systems. In Alexander I. Rudnicky and Candace L. Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 4–6, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [46] Herve Dejean, Eric Gaussier, Cyril Goutte, and Kenji Yamada. Reducing parameter space for word alignment. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 23–26, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [47] Yonggang Deng and Sanjeev Khudanpur. Latent semantic information in maximum entropy language models for conversational speech recognition. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 56–63, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [48] Ian Densham and James Reid. System demo: A geo-coding service encompassing a geo-parsing tool and integrated digital gazetteer service. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 79–80, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [49] Dinh Dien and Hoang Kiem. Pos-tagger for english-vietnamese bilingual corpus. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 88–95, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [50] Takao Doi and Eiichiro Sumita. Input sentence splitting and translating. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 104–110, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [51] Takao Doi, Eiichiro Sumita, and Hirofumi Yamamoto. Adaptation using out-of-domain corpus within ebmt. In Marti Hearst and Mari Ostendorf,

editors, *HLT-NAACL 2003: Short Papers*, pages 16–18, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [52] Peter Ford Dominey and Thomas Voegtlin. Learning word meaning and grammatical constructions from narrated video events. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 38–45, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [53] Bonnie Dorr, David Zajic, and Richard Schwartz. Hedge trimmer: A parse-and-trim approach to headline generation. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 1–8, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [54] Shona Douglas. Active learning for classifying phone sequences from unsupervised phonotactic models. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 19–21, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [55] John Dowding and James Hieronymus. A spoken dialogue interface to a geologist’s field assistant. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 9–10, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [56] Daniel M. Dunlavy, John Conroy, and Dianne P. O’Leary. Qcs: A tool for querying, clustering, and summarizing documents. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 11–12, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [57] Jaety Edwards, Ryan White, and David Forsyth. Words and pictures in the news. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 6–13, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [58] Jason Eisner. Simpler and more general minimization for weighted finite-state automata. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 64–71, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [59] Donghui Feng. Cooperative model-based language understanding. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL*

2003 Student Research Workshop., pages 55–60, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [60] David Ferrucci and Adam Lally. Accelerating corporate research in the development, application, and deployment of human language technologies. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 67–74, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [61] Michael Fleischman and Eduard Hovy. A maximum entropy approach to framenet tagging. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 22–24, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [62] Radu Florian, Abe Ittycheriah, Hongyan Jing, and Tong Zhang. Named entity recognition through classifier combination. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 168–171, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [63] Simona Gandrabur and George Foster. Confidence estimation for translation prediction. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 95–102, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [64] Ulrich Germann. Greedy decoding for statistical machine translation in almost linear time. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 72–79, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [65] Roxana Girju, Adriana Badulescu, and Dan Moldovan. Learning semantic constraints for the automatic discovery of part-whole relations. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 80–87, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [66] James Glass and Stephanie Seneff. Flexible and personalizable mixed-initiative dialogue systems. In Alexander I. Rudnicky and Candace L. Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 19–21, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.

- [67] Joshua Goodman. The state of the art in language modeling. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 4–4, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [68] Peter Gorniak and Deb Roy. Understanding complex visually referring utterances. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 14–21, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [69] Amardeep Grewal, Timothy Allison, Stanko Dimitrov, and Dragomir Radev. Multi-document summarization using off the shelf compression software. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 17–24, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [70] Claire Grover, Ben Hachey, and Chris Korycinski. Summarising legal texts: Sentential tense and argumentative roles. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 33–40, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [71] Iryna Gurevych, Rainer Malaka, Robert Porzel, and Hans-Peter Zorn. Semantic coherence scoring using an ontology. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 88–95, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [72] Iryna Gurevych, Stefan Merten, and Robert Porzel. Automatic creation of interface specifications from ontologies. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 59–66, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [73] Iryna Gurevych, Robert Porzel, Elena Slinko, Norbert Pflieger, Jan Alexandersson, and Stefan Merten. Less is more: Using a single knowledge representation in dialogue systems. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 14–21, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [74] Nizar Habash and Bonnie Dorr. A categorial variation database for english. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003:*

Main Proceedings, pages 96–102, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [75] Kadri Hacioglu and Wayne Ward. Question classification with support vector machines and error correcting codes. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 28–30, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [76] Kadri Hacioglu and Wayne Ward. Target word detection and semantic role chunking using support vector machines. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 25–27, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [77] James Hammerton. Named entity recognition with long short-term memory. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 172–175, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [78] Thomas Hanneforth, Silvan Heintze, and Manfred Stede. Rhetorical parsing with underspecification and forests. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 31–33, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [79] Hilda Hardy, Tomek Strzalkowski, and Min Wu. Dialogue management for an automated multilingual call center. In Alexander I. Rudnicky and Candace L. Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 10–12, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [80] James Henderson. Inducing history representations for broad coverage statistical parsing. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 103–110, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [81] John C. Henderson. Word alignment baselines. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 27–30, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [82] Iris Hendrickx and Antal van den Bosch. Memory-based one-step named-entity recognition: Effects of seed list features, classifier stacking, and

- unannotated data. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 176–179, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [83] Gerd Herzog, Heinz Kirchmann, Stefan Merten, Alassane Ndiaye, and Peter Poller. Multiplatform testbed: An integration platform for multimodal dialog systems. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 75–82, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [84] Dustin Hillard, Mari Ostendorf, and Elizabeth Shriberg. Detection of agreement vs. disagreement in meetings: Training with unlabeled data. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 34–36, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [85] Graeme Hirst. Introduction to non-statistical natural language processing. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 1–1, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [86] Baden Hughes and Steven Bird. Grid-enabling natural language engineering by stealth. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 31–38, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [87] Nancy Ide, Laurent Romary, and Eric de la Clergerie. International standard for a linguistic annotation framework. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 25–30, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [88] Kenji Imamura, Yasuhiro Akiba, and Eiichiro Sumita. Automatic expansion of equivalent sentence set based on syntactic substitution. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 37–39, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [89] Abraham Ittycheriah, Lucian Lita, Nanda Kambhatla, Nicolas Nicolov, Salim Roukos, and Margo Stys. Identifying and tracking entity mentions in a maximum entropy framework. In Marti Hearst and Mari Ostendorf,

editors, *HLT-NAACL 2003: Short Papers*, pages 40–42, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [90] Hongyan Jing, Daniel Lopresti, and Chilin Shih. Summarization of noisy documents: A pilot study. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 25–32, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [91] Howard Johnson and Joel Martin. Unsupervised learning of morphology for english and inuktitut. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 43–45, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [92] Dominic R. Jones and Cynthia A. Thompson. Identifying events using similarity and context. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 135–141, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [93] Hiroyuki Kaji. Word sense acquisition from bilingual comparable corpora. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 111–118, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [94] Dharmendra Kanajiya, Arun Kamar, and Surenda Prasad. Automatic evaluation of students’ answers using syntactically enhanced lsa. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 53–60, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [95] Vangelis Karkaletsis, Constantine D. Spyropoulos, Dimitris Souflis, Claire Grover, Ben Hachey, Maria Teresa Paziienza, Michele Vindigni, Emmanuel Cartier, and Jose Coch. Demonstration of the crossmarc system. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 13–14, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [96] Paul Kingsbury and Karin Kipper. Deriving verb-meaning clusters from syntactic structure. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 70–77, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [97] Katrin Kirchhoff and Mari Ostendorf. Directions for multi-party human-computer interaction research. In Alexander I. Rudnicky and Candace L.

- Sidner, editors, *HLT-NAACL 2003 Workshop: Research Directions in Dialogue Processing*, pages 7–9, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [98] Dan Klein and Christopher D. Manning. A* parsing: Fast exact viterbi parse selection. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 119–126, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [99] Dan Klein, Joseph Smarr, Huy Nguyen, and Christopher D. Manning. Named entity recognition with character-level models. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 180–183, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [100] Kevin Knight and Philipp Koehn. What’s new in statistical machine translation. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 5–5, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [101] Ola Knutsson, Teresa Cerrato Pargman, and Kerstin Severinson Eklundh. Transforming grammar checking technology into a learning environment for second language writing. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 38–45, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [102] Philipp Koehn, Franz J. Och, and Daniel Marcu. Statistical phrase-based translation. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 127–133, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [103] Okan Kolak, William Byrne, and Philip Resnik. A generative probabilistic ocr model for nlp applications. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 134–141, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [104] Grzegorz Kondrak, Daniel Marcu, and Kevin Knight. Cognates can improve statistical translation models. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 46–48, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [105] Hans-Ulrich Krieger. Sdl—a description language for building nlp systems. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 83–90, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [106] Shankar Kumar and William Byrne. A weighted finite state transducer implementation of the alignment template model for statistical machine translation. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 142–149, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [107] Kui Lam Kwok and Qiang Deng. Geoname: a system for back-transliterating pinyin place names. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 26–30, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [108] Thomas Landauer. Pasteur’s quadrant: Computational linguistics, lsa, and education. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 46–52, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [109] Chun-Jen Lee and Jason S. Chang. Acquisition of english-chinese transliterated word pairs from parallel-aligned texts using a statistical machine transliteration model. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 96–103, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [110] Jochen Leidner. Current issues in software engineering for natural language processing. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 45–50, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [111] Jochen L. Leidner, Gail Sinclair, and Bonnie Webber. Grounding spatial named entities for information extraction and question answering. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 31–38, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [112] Huifeng Li, K. Rohini Srihari, Cheng Niu, and Wei Li. Infoextract location normalization: a hybrid approach to geographic references in information extraction. In András Kornai and Beth Sundheim, editors, *HLT-NAACL*

- 2003 Workshop: Analysis of Geographic References*, pages 39–44, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [113] Xin Li, Dan Roth, and Yuancheng Tu. Phrasenet: towards context sensitive lexical semantics. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 87–94, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [114] Chin-Yew Lin and Eduard Hovy. Automatic evaluation of summaries using n-gram co-occurrence statistics. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 150–157, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [115] Chin-Yew Lin and Eduard Hovy. The potential and limitations of automatic sentence extraction for summarization. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 73–80, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [116] Dekang Lin and Colin Cherry. Proalign: Shared task system description. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 11–14, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [117] Dekang Lin and Colin Cherry. Word alignment with cohesion constraint. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 49–51, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [118] Diane Litman, Kate Forbes, and Scott Silliman. Towards emotion prediction in spoken tutoring dialogues. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 52–54, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [119] Yang Liu. Word fragments identification using acoustic-prosodic features in conversational speech. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*, pages 37–42, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [120] Deryle Lonsdale and Diane Strong-Krause. Automated rating of esl essays. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 61–67, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [121] Brian Mak, Manhung Siu, Mimi Ng, Yik-Cheung Tam, Yu-Chung Chan, Kin-Wah Chan, Ka-Yee Leung, Simon Ho, Jimmy Wong, and Jacqueline Lo. Plaser: Pronunciation learning via automatic speech recognition. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 23–29, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [122] Juha Makkonen. Investigations on event evolution on tdt. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*, pages 43–48, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [123] Inderjeet Mani, Barry Schiffman, and Jianping Zhang. Inferring temporal ordering of events in news. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 55–57, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [124] Gideon Mann and David Yarowsky. Unsupervised personal name disambiguation. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 33–40, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [125] Christopher Manning and Dan Klein. Optimization, maxent models, and conditional estimation without magic. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 8–8, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [126] Dimitar Manov, Atanas Kiryakov, Borislav Popov, Kalina Bontcheva, Diana Maynard, and Hamish Cunningham. Experiments with geographic knowledge for information extraction. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 1–9, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [127] Joel Martin, Howard Johnson, Benoit Farley, and Anna Maclachlan. Aligning and using an english-inuktitut parallel corpus. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building*

and Using Parallel Texts: Data Driven Machine Translation and Beyond, pages 115–118, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [128] Katsuya Masuda, Takashi Ninomiya, Yusuke Miyao, Tomoko Ohta, and Jun'ichi Tsujii. A robust retrieval engine for proximal and structural search. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 58–60, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [129] Takuya Matsuzaki, Yusuke Miyao, and Jun'ichi Tsujii. An efficient clustering algorithm for class-based language models. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 119–126, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [130] James Mayfield, Paul McNamee, and Christine Piatko. Named entity recognition using hundreds of thousands of features. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 184–187, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [131] Andrew McCallum and Wei Li. Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 188–191, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [132] Kathleen McKeown, Regina Barzilay, John Chen, David Elson, David Evans, Judith Klavans, Ani Nenkova, Barry Schiffman, and Sergey Sigelman. Columbia's newsblaster: New features and future directions. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 15–16, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [133] I. Dan Melamed. Multitext grammars and synchronous parsers. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 158–165, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [134] I. Dan Melamed, Ryan Green, and Joseph P. Turian. Precision and recall of machine translation. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 61–63, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [135] Rada Mihalcea and Ted Pedersen. An evaluation exercise for word alignment. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 1–10, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [136] Ruslan Mitkov and Le An Ha. Computer-aided generation of multiple-choice tests. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 17–22, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [137] Yusuke Miyao and Jun'ichi Tsujii. A model of syntactic disambiguation based on lexicalized grammars. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 1–8, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [138] Behrang Mohit and Sridhar Narayanan. Semantic extraction with wide-coverage lexical resources. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 64–66, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [139] Dan Moldovan, Christine Clark, Sanda Harabagiu, and Steve Maierano. Cogex: A logic prover for question answering. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 166–172, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [140] Thomas Morton and Jeremy LaCivita. Wordfreak: An open tool for linguistic annotation. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 17–18, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [141] Erik T. Mueller. Story understanding through multi-representation model construction. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 46–53, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [142] Robert Munro, Daren Ler, and Jon Patrick. Meta-learning orthographic and contextual models for language independent named entity recognition. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 192–195, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.

- [143] Cosmin Munteanu. Indexing methods for efficient parsing. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*., pages 25–30, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [144] Preslav I. Nakov and Marti A. Hearst. Category-based pseudowords. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 67–69, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [145] Ramesh Nallapati. Semantic language models for topic detection and tracking. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*., pages 1–6, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [146] Mary S. Neff, Roy J. Byrd, and Branimir K. Boguraev. The talent system: Textract architecture and data model. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 1–8, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [147] Ani Nenkova and Kathleen McKeown. References to named entities: a corpus study. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 70–72, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [148] Vincent Ng and Claire Cardie. Weakly supervised natural language learning without redundant views. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 173–180, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [149] Stephen Nightingale and Hideki Tanaka. Comparing the sentence alignment yield from two news corpora using a dictionary-based alignment system. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 119–122, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [150] Sergei Nirenburg, Marjorie McShane, and Stephen Beale. Operative strategies in ontological semantics. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 22–29, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [151] Cheng Niu, Wei Li, Jihong Ding, and Rohini K. Srihari. Bootstrapping for named entity tagging using concept-based seeds. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 73–75, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [152] Eric Nyberg and Robert Frederking. Javelin: A flexible, planner-based architecture for question answering. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 19–20, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [153] Douglas W. Oard. Information retrieval systems as integration platforms for language technologies. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 2–2, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [154] Douglas W. Oard, David Doermann, Bonnie Dorr, Daqing He, Philip Resnik, Amy Weinberg, William Byrne, Sanjeev Khudanpur, David Yarowsky, Anton Leuski, Philipp Koehn, and Kevin Knight. Desparately seeking cebuano. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 76–78, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [155] Tim Oates. Grounding word meanings in sensor data: Dealing with referential uncertainty. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 62–69, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [156] Tom O’Hara and Janyce Wiebe. Preposition semantic classification via treebank and framenet. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 79–86, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [157] Manabu Okumura, Takahiro Fukusima, and Hidetsugu Nanba. Text summarization challenge 2 - text summarization evaluation at ntcir workshop 3. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 49–56, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [158] Andrew Olney, Max Louwerse, Eric Matthews, Johanna Marineau, Heather Hite-Mitchell, and Arthur Graesser. Utterance classification in autotutor. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL*

2003 Workshop: Building Educational Applications Using Natural Language Processing, pages 1–8, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [159] Bo Pang, Kevin Knight, and Daniel Marcu. Syntax-based alignment of multiple translations: Extracting paraphrases and generating new sentences. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 181–188, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [160] Brian E. Pangburn, S. Sitharama Iyengar, Robert C. Mathews, and Jonathan P. Ayo. Ebla: A perceptually grounded model of language acquisition. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 46–53, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [161] Patrick Pantel and Dekang Lin. Automatically discovering word senses. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 21–22, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [162] Fuchun Peng, Dale Schuurmans, and Shaojun Wang. Language and task independent text categorization with simple language models. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 189–196, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [163] Leonid Peshkin, Avi Pfeffer, and Virginia Savova. Bayesian nets for syntactic categorization of novel words. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 79–81, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [164] Jochen Peters. Lm studies on filled pauses in spontaneous medical dictation. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 82–84, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [165] Katharina Probst. Using 'smart' bilingual projection to feature-tag a monolingual dictionary. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 103–110, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [166] Amruta Purandare. Discriminating among word senses using mcquitty's similarity analysis. In Costas Boulis, Eric Breck, and Victor Lavrenko,

- editors, *HLT-NAACL 2003 Student Research Workshop*, pages 19–24, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [167] James Pustejovsky and Inderjeet Mani. Annotation of temporal and event expressions. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 6–6, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [168] Victor Raskin, Sergei Nirenburg, Christian F. Hempelmann, Inna Nirenburg, and Katrina E. Triezenberg. The genesis of a script for bankruptcy in ontological semantics. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 30–37, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [169] Erik Rauch, Michael Bukatin, and Kenneth Baker. A confidence-based framework for disambiguating geographic terms. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 50–54, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [170] Deepak Ravichandran, Abraham Ittycheriah, and Salim Roukos. Automatic derivation of surface text patterns for a maximum entropy based question answering system. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 85–87, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [171] Ehud Reiter and Somayajulu Sripada. Learning the meaning and usage of time phrases from a parallel text-data corpus. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 78–85, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [172] Doug Reynolds and Marc Zissman. Automatic speaker and language recognition. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 9–9, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [173] Stefan Riezler, Tracy H. King, Richard Crouch, and Annie Zaenen. Statistical sentence condensation using ambiguity packing and stochastic disambiguation methods for lexical-functional grammar. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 197–204, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.

- [174] Ellen Riloff, Janyce Wiebe, and Theresa Wilson. Learning subjective nouns using extraction pattern bootstrapping. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 25–32, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [175] Brian Roark and Michiel Bacchiani. Supervised and unsupervised pcfg adaptation to novel domains. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 205–212, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [176] Carolyn P. Rosé, Diane Litman, Dumisizwe Bhembe, Kate Forbes, Scott Silliman, Ramesh Srivastava, and Kurt VanLehn. A comparison of tutor and student behavior in speech versus text based tutoring. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 30–37, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [177] Carolyn P. Rose, Antonio Roque, Dumisizwe Bhembe, and Kurt VanLehn. A hybrid approach to content analysis for automatic essay grading. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 88–90, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [178] Carolyn P. Rosé, Antonio Roque, Dumisizwe Bhembe, and Kurt VanLehn. A hybrid text classification approach for analysis of student essays. In Jill Burstein and Claudia Leacock, editors, *HLT-NAACL 2003 Workshop: Building Educational Applications Using Natural Language Processing*, pages 68–75, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [179] Mihai Rotaru and Diane J. Litman. Exceptionality and natural language learning. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 63–70, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [180] Deb Roy, Kai-Yuh Hsiao, and Nikolaos Mavridis. Conversational robots: Building blocks for grounding word meaning. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 70–77, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [181] Ulrich Schäfer. What: An xslt-based infrastructure for the integration of natural language processing components. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 9–16, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [182] Gerold Schneider. A low-complexity, broad-coverage probabilistic dependency parser for english. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*., pages 31–36, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [183] Lenhart Schubert and Matthew Tong. Extracting and evaluating general world knowledge from the brown corpus. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 7–13, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [184] Satoshi Sekine and Chikashi Nobata. A survey for multi-document summarization. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 65–72, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [185] Sid-Ahmed Selouani, Hesham Tolba, and Douglas O’Shaughnessy. Auditory-based acoustic distinctive features and spectral cues for robust automatic speech recognition in low-snr car environments. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 91–93, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [186] Riccardo Serafin, Barbara Di Eugenio, and Michael Glass. Latent semantic analysis for dialogue act classification. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 94–96, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [187] Fei Sha and Fernando Pereira. Shallow parsing with conditional random fields. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 213–220, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [188] Libin Shen and Aravind K. Joshi. An svm-based voting algorithm with application to parse reranking. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*,

pages 9–16, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.

- [189] Stuart M. Shieber and Xiaopeng Tao. Comma restoration using constituency information. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 221–227, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [190] Mitsuo Shimohata, Eiichiro Sumita, and Yuji Matsumoto. Retrieving meaning-equivalent sentences for example-based rough translation. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 50–56, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [191] Hiroyuki Shinnou and Minoru Sasaki. Unsupervised learning of word sense disambiguation rules by estimating an optimum iteration number in the em algorithm. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 41–48, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [192] Michel Simard. Translation spotting for translation memories. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 65–72, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [193] Michel Simard and Philippe Langlais. Statistical translation alignment with compositionality constraints. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 19–22, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [194] Andrew E. Smith. Automatic extraction of semantic networks from text using leximancer. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 23–24, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [195] David A. Smith and Gideon S. Mann. Bootstrapping toponym classifiers. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 45–49, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [196] Radu Soricut and Daniel Marcu. Sentence level discourse parsing using syntactic and lexical information. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 228–235, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [197] Humphrey Southall. Defining and identifying the roles of geographic references within text. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 69–78, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [198] Rohini K. Srihari, Wei Li, Cheng Niu, and Thomas Cornell. Infoextract: A customizable intermediate level information extraction engine. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 51–58, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [199] Manfred Stede. Surfaces and depths in text understanding: The case of newspaper commentary. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 62–69, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [200] Mark Steedman, Rebecca Hwa, Stephen Clark, Miles Osborne, Anoop Sarkar, Julia Hockenmaier, Paul Ruhlen, Steven Baker, and Jeremiah Crim. Example selection for bootstrapping statistical parsers. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 236–243, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [201] Suzanne Stevenson and Eric Joanis. Semi-supervised verb class discovery using noisy features. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 71–78, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [202] Nicola Stokes. Spoken and written news story segmentation using lexical chains. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*, pages 49–54, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [203] Kiyoshi Sudo, Satoshi Sekine, and Ralph Grishman. pre-codie-crosslingual on-demand information extraction. In Bob Frederking and

- Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 25–26, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [204] Valentin Tablan, Kalina Bontcheva, Diana Maynard, and Hamish Cunningham. Ollie: On-line learning for information extraction. In Hamish Cunningham and Jon Patrick, editors, *HLT-NAACL 2003 Workshop: Software Engineering and Architecture of Language Technology Systems (SEALTS)*, pages 17–24, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [205] Rong Tang, Kwong Bor Ng, Tomek Strzalkowski, and Paul B. Kantor. Automatically predicting information quality in news documents. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 97–99, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [206] Elena Terenzi and Barbara Di Eugenio. Building lexical semantic representations for natural language instructions. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 100–102, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [207] Egidio L. Terra and Charles L. A. Clarke. Frequency estimates for statistical word similarity measures. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 244–251, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [208] Virongrong Tesprasit, Paisarn Charoenpornasawat, and Virach Sortlertlamvanich. A context-sensitive homograph disambiguation in thai text-to-speech synthesis. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 103–105, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [209] Paul Thompson. Dynamic integration of distributed semantic services: Infrastructure for process queries and question answering. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 27–28, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [210] Bridget Thomson McInnes and Ted Pedersen. The duluth word alignment system. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 40–43, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [211] Christoph Tillmann and Fei Xia. A phrase-based unigram model for statistical machine translation. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 106–108, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [212] Erik F. Tjong Kim Sang and Fien De Meulder. Introduction to the conll-2003 shared task: Language-independent named entity recognition. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 142–147, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [213] Kristina Toutanova, Dan Klein, Christopher D. Manning, and Yoram Singer. Feature-rich part-of-speech tagging with a cyclic dependency network. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 252–259, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [214] Yoshimasa Tsuruoka and Jun'ichi Tsujii. Training a naive bayes classifier via the em algorithm with a class distribution constraint. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 127–134, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [215] Dan Tufiş, Ana-Maria Barbu, and Radu Ion. Treq-al: A word alignment system with limited language resources. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 36–39, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [216] Olga Uryupina. Semi-supervised learning of geographical gazetteer from the internet. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*, pages 18–25, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [217] Benjamin Van Durme, Yifen Huang, Anna Kupść, and Eric Nyberg. Towards light semantic processing for question answering. In Graeme Hirst and Sergei Nirenburg, editors, *HLT-NAACL 2003 Workshop: Text Meaning*, pages 54–61, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [218] Hans van Halteren and Simone Teufel. Examining the consensus between human summaries: initial experiments with factoid analysis. In Dragomir

- Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 57–64, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [219] Ulla Vanhatalo and Hilary Chan. Population testing: Extracting semantic information on near-synonymy from native speakers. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 86–91, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [220] Ellen M. Voorhees. Evaluating answers to definition questions. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 109–111, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [221] Ellen M. Voorhees. Evaluating the evaluation: A case study using the trec 2002 question answering track. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 260–267, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [222] Nina Wacholder and Peng Song. Toward a task-based gold standard for evaluation of np chunks and technical terms. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 268–275, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [223] Sven Wachsmuth, Suzanne Stevenson, and Sven Dickinson. Towards a framework for learning structured shape models from text-annotated images. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 22–29, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [224] Alex Waibel, Ahmed Badran, Alan W Black, Robert Frederking, Donna Gates, Alon Lavie, Lori Levin, Kevin Lenzo, Laura Mayfield Tomokiyo, Juergen Reichert, Tanja Schultz, Dorcas Wallace, Monika Woszczyna, and Jing Zhang. Speechalator: Two-way speech-to-speech translation in your hand. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 29–30, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [225] Richard Waldinger, Peter Jarvis, and Jennifer Dungan. Pointing to places in a deductive geospatial theory. In András Kornai and Beth Sundheim, editors, *HLT-NAACL 2003 Workshop: Analysis of Geographic References*,

- pages 10–17, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [226] Mark Wasson. Nlp r&d and commercial deployment. In James Allan, Jason Eisner, and Wayne Ward, editors, *HLT-NAACL 2003 Tutorials*, pages 7–7, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [227] Casey Whitelaw and Jon Patrick. Named entity recognition using a character-based probabilistic approach. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 196–199, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [228] Brian Whitman, Deb Roy, and Barry Vercoe. Learning word meanings and descriptive parameter spaces from music. In Regina Barzilay, Ehud Reiter, and Jeffrey Mark Siskind, editors, *HLT-NAACL 2003 Workshop: Learning Word Meaning from Non-Linguistic Data*, pages 92–99, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [229] Dominic Widdows. Unsupervised methods for developing taxonomies by combining syntactic and statistical information. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 276–283, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [230] Dominic Widdows and Scott Cederberg. Monolingual and bilingual concept visualization from corpora. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 31–32, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [231] Sandra Williams. Language choice models for microplanning and readability. In Costas Boulis, Eric Breck, and Victor Lavrenko, editors, *HLT-NAACL 2003 Student Research Workshop*., pages 13–18, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [232] Theresa Wilson, David R. Pierce, and Janyce Wiebe. Identifying opinionated sentences. In Bob Frederking and Bob Younger, editors, *HLT-NAACL 2003: Demo Session*, pages 33–34, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [233] Dekai Wu, Grace Ngai, and Marine Carpuat. A stacked, voted, stacked model for named entity recognition. In Walter Daelemans and Miles

- Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 200–203, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [234] Kaoru Yamamoto, Taku Kudo, Yuta Tsuboi, and Yuji Matsumoto. Learning sequence-to-sequence correspondences from parallel corpora via sequential pattern mining. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 73–80, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [235] Hua Yu and Tanja Schultz. Implicit trajectory modeling through gaussian transition models for speech recognition. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Short Papers*, pages 112–114, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.
- [236] Haiqin Zhang, Zheng Chen, Wei-ying Ma, and Qingsheng Cai. A study for document summarization based on personal annotation. In Dragomir Radev and Simone Teufel, editors, *HLT-NAACL 2003 Workshop: Text Summarization (DUC03)*, pages 41–48, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics.
- [237] Tong Zhang, Fred Damerou, and David Johnson. Updating an nlp system to fit new domains: an empirical study on the sentence segmentation problem. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 56–62, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [238] Tong Zhang and David Johnson. A robust risk minimization based named entity recognition system. In Walter Daelemans and Miles Osborne, editors, *Seventh Conference on Natural Language Learning (CoNLL-03)*, pages 204–207, Edmonton, Alberta, Canada, May 31 - June 1 2003. Association for Computational Linguistics. In association with HLT-NAACL 2003.
- [239] Bing Zhao and Stephan Vogel. Word alignment based on bilingual bracketing. In Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 15–18, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.
- [240] Bing Zhao, Klaus Zechner, Stephen Vogel, and Alex Waibel. Efficient optimization for bilingual sentence alignment based on linear regression. In

Rada Mihalcea and Ted Pedersen, editors, *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, pages 81–87, Edmonton, Alberta, Canada, May 31 2003. Association for Computational Linguistics.

- [241] Liang Zhou and Eduard Hovy. A web-trained extraction summarization system. In Marti Hearst and Mari Ostendorf, editors, *HLT-NAACL 2003: Main Proceedings*, pages 284–290, Edmonton, Alberta, Canada, May 27 - June 1 2003. Association for Computational Linguistics.