

# Seminar Wintersemester 2011: Sentiment Analysis

Michael Strube  
Heidelberger Institut für Theoretische Studien  
Schloss-Wolfsbrunnenweg 35  
69118 Heidelberg  
<http://michael.kimstrube.de>

27. Januar 2012

## Termine, Themenvorschläge

**20.10.2011**

**fällt aus** – Michael Strube bei AMIA '11 (i2b2 Workshop)

**27.10.2011**

**Einführung, Terminologie:** Pang & Lee (2008a)

**Vorläufer:** Carbonell (1979), Wilks & Bien (1984), Wiebe (1990)

**03.11.2011**

**Subjectivity, private states, point of view:** Wiebe (1990), Wiebe (1994), Wiebe & Bruce (1995)

**zur Vorbereitung:** Wiebe & Bruce (1995) oder Wiebe (1994)

**10.11.2011**

**Referat: Eric Hildebrand – Subjectivity detection – Weakly supervised approaches:** Riloff et al. (2003), Riloff & Wiebe (2003)

**zur Vorbereitung:** Riloff et al. (2003) oder Riloff & Wiebe (2003)

optional: Su & Markert (2009), Lin et al. (2011)

**Subjectivity detection – Recent work**

features for subjectivity detection (Wiebe et al., 2004; Wilson et al., 2005); BLOG track at TREC '06: Ounis et al. (2006); subsentential analysis: Esuli & Sebastiani (2006a), Takamura et al. (2006), Mihalcea et al. (2007)

**17.11.2011**

**Annotating private state, opinions, emotions:** Wiebe et al. (1999), Wilson & Wiebe (2003), Wiebe et al. (2005), Wilson & Wiebe (2005), Somasundaran et al. (2008a),

Yesselina et al. (2010)

**zur Vorbereitung:** Wilson & Wiebe (2005) oder Somasundaran et al. (2008a)

#### **24.11.2011**

**Referat: Jan Pawellek – Emotion in speech, speech analytics:** Batliner et al. (2008), Ranganath et al. (2009)

optional: El Ayadi et al. (2011), Murray & Carenini (2011)

**Referat: Annika Berger – Emotional speech synthesis:** Inanoglu & Young (2009), Yu et al. (2010)

**zur Vorbereitung:** Ranganath et al. (2009) oder (Inanoglu & Young (2009) oder Yu et al. (2010))

#### **01.12.2011**

**Referat: Schigehiko Schamoni – Sentiment polarity classification: Thumbs up – thumbs down:** Pang et al. (2002), Turney (2002)

**Referat: Magali Boizot-Roche – Summarization:** Wang & Liu (2011)

optional: Beineke et al. (2004), Zhou & Hovy (2005), Ku et al. (2006), Carenini et al. (2006), Titov & McDonald (2008a), Titov & McDonald (2008b), Zhang (2008)

**zur Vorbereitung:** (Pang et al. (2002) oder Turney (2002)) oder Wang & Liu (2011)

#### **08.12.2011**

**Referat: Madeleine Remse – Domain adaptation:** Blitzer et al. (2007), Bollegala et al. (2011)

**Referat: Philipp Busch – Language adaptation:** Wei & Pal (2010), Lu et al. (2011)

**zur Vorbereitung:** Wei & Pal (2010) und Blitzer et al. (2007)

optional: Abdul-Mageed et al. (2011), Hassan et al. (2011), He et al. (2011), Zhang & Liu (2011)

#### **15.12.2011**

**Referat: Anja Summa – Lexicon induction:** Hatzivassiloglou & McKeown (1997), Kaji & Kitsuregawa (2007)

**Referat: Jonas Placzek – Sentiment lexicon creation:** Esuli & Sebastiani (2006b), Baccianella et al. (2010)

**zur Vorbereitung:** Esuli & Sebastiani (2006b) und Hatzivassiloglou & McKeown (1997)

optional: Esuli & Sebastiani (2006a), Esuli & Sebastiani (2007), Taboada et al. (2011), Hatzivassiloglou & Wiebe (2000), Turney (2002), Yu & Hatzivassiloglou (2003), Riloff et al. (2003), Turney & Littman (2003), Hu & Liu (2004), Esuli & Sebastiani (2006b)

#### **22.12.2011**

**Referat: Huiqin Körkel-Qu – Degrees of polarity:** Pang & Lee (2005), Goldberg & Zhu (2006)

optional: Wilson et al. (2004), Goldberg & Zhu (2006)

**zur Vorbereitung:** Pang & Lee (2005)

#### **12.01.2012**

**Gastvortrag: Cäcilia Zirn (Uni Mannheim) – Fine-grained sentiment analysis with structural features** (Zirn et al., 2011)

**zur Vorbereitung:** Somasundaran & Wiebe (2009) oder Zirn et al. (2011)

optional: Pang & Lee (2004), Devitt & Ahmad (2007), Somasundaran et al. (2008a), Somasundaran et al. (2008b), Somasundaran et al. (2009), Kim & Hovy (2006a), Johansson & Moschitti (2010), Johansson & Moschitti (2011), Jiang et al. (2011), Park et al. (2011), Sauper et al. (2011), Täckström & McDonald (2011b), Täckström & McDonald (2011a), Yu et al. (2011), Zirn et al. (2011)

#### **19.01.2012**

**Referat: Thomas Bögel – Sentiment polarity classification: Why do reviewers like or dislike a product?** Kim & Hovy (2006a), Qin et al. (2008)

optional: Branavan et al. (2008), Titov & McDonald (2008a), Titov & McDonald (2008b)

**Referat: Isabelle Augenstein – Ironie, Sarkasmus, etc.:** Carvalho et al. (2011), González-Ibanez et al. (2011)

**zur Vorbereitung:** Carvalho et al. (2011) und(/oder) Qin et al. (2008)

#### **26.01.2012**

**Referat: Nicolas Bellm – Political discourse:** Laver et al. (2003), Mullen & Malouf (2006), Bansal et al. (2008), Martin & Vanberg (2008), Mullen & Malouf (2008)

**Referat: Katharina Sowa – Political orientation of websites:** Grefenstette et al. (2004), Lin & Hauptmann (2006), Lin et al. (2006); **Political debates:** Bansal et al. (2008), Carvalho et al. (2011); **Political blogs:** Abbott et al. (2011), Balasubramanian et al. (2011), Agarwal et al. (2011)

**zur Vorbereitung:** Agarwal et al. (2011) und Mullen & Malouf (2006)

**02.02.2012**

**zur Vorbereitung:** Bitte senden Sie mir offengebliebene Fragen und weitere Themen, die Sie interessieren, bis spätestens Dienstag, 31.1.2012, 13 Uhr zu, damit wir über diese in der letzten Sitzung sprechen können.

**Themenvorschläge:**

**Features: Parts-of-speech (adjectives, nouns), syntax, language models:** beyond unigrams: Dave et al. (2003), Kim & Hovy (2006b); adjectives Hatzivassiloglou & McKeown (1997), Hatzivassiloglou & Wiebe (2000), Mullen & Collier (2004), Whitelaw et al. (2005); syntax: Dave et al. (2003), Gamon (2004), Kudo & Matsumoto (2004), Ng et al. (2006)

**Unsupervised approaches:** Wiebe & Riloff (2005), Pang & Lee (2008b), Maas et al. (2011)

**Interaction topic-sentiment:** Hurst & Nigam (2004), Mullen & Collier (2004), Eguchi & Lavrenko (2006), Eguchi & Shah (2006), Kim & Hovy (2007), Mei et al. (2007)

**Emotion recognition:** Subasic & Huettner (2001), Liu et al. (2003), Alm et al. (2005)

**Deception detection:** Burgoon et al. (2003), Zhou et al. (2003), Graciarena et al. (2006), Bachenko et al. (2008), Hancock et al. (2008)

**Summarization:** Beineke et al. (2004), Zhou & Hovy (2005), Ku et al. (2006), Carenini et al. (2006), Titov & McDonald (2008a), Titov & McDonald (2008b), Zhang (2008), Wang & Liu (2011);

**Perspective detection for summarization:** Seki et al. (2004), Seki et al. (2005), Seki et al. (2006)

**Question answering:** Cardie et al. (2003), Wiebe et al. (2003)

**Bemerkungen**

**Leistungsnachweise:** Lektüre und aktive Teilnahme (1/3), Referat (1/3), Hausarbeit (1/3). Hausarbeit: 8-10 Seiten (Proseminar), 12-15 Seiten (Hauptseminar) inkl. Bibliographie. Die Hausarbeit kann auch per Email an mich geschickt werden, aber *nicht* als Word-Datei sondern nur als PDF-Datei. – Ich empfehle, wissenschaftliche Texte mit Latex und Bibtex zu verfassen.

Regelmäßige Teilnahme (d.i. nicht mehr als einmal unentschuldigtes Fehlen) ist Voraussetzung für den Scheinerwerb. Zu jeder Sitzung müssen jeweils zwei Fragen (!) zu einem Papier abgegeben werden, das in der aktuellen Sitzung vorgestellt wird. Abgabe entweder per Email bis spätestens 13 Uhr am Tag der Sitzung oder schriftlich direkt vor der Sitzung. Dies geht in die Bewertung für aktive Teilnahme am Seminar ein.

**Literatur:** Viele Papiere können direkt aus der *ACL Anthology* kopiert werden (<http://acl.ldc.upenn.edu/>), insbesondere alle Papiere der (E/NA)ACL-, *Coling*- und *EMNLP*-Konferenzen, alle Workshops, die im Rahmen dieser Konferenzen veranstaltet wurden und die Zeitschrift *Computational Linguistics*. Papiere, die von der *AAAI* publiziert wurden (AAAI-Konferenz, AAI-Workshops, AAI-Symposia, etc.) sind in

der *AAAI Digital Library* verfügbar (<http://www.aaai.org/Library>). – Die meisten weiteren Zeitschriften sind elektronisch verfügbar über die UB (<http://rzblx1.uni-regensburg.de/ezeit/search.phtml?bibid=UBHE>) – oder stehen dort im Regal.

**Sprechstunde:** Auf Vereinbarung (Email, Telefon) bei mir im Büro, ggf. auch im Anschluß an das Seminar.

**Hausarbeiten:**

Maximal 8-10 Seiten (Proseminar), 12-15 Seiten (Hauptseminar) inkl. Abbildungen, inkl. Literaturverzeichnis.

**Inhalt:** Fokus auf das vorgestellte Papier; NICHT *Related Work*-Kapitel referieren, wenn die entsprechenden Papiere nicht gelesen wurden; Evaluierung berichten; WICHTIG: mit eigener Meinung oder Bewertung abschließen.

**Stil:** Wissenschaftlichkeit drückt sich nicht durch lange, komplizierte Sätze und exzessiven Gebrauch von Fremdwörtern aus – deshalb bitte kurze Sätze, einfache Sprache; Hausarbeiten vor der Abgabe Korrektur lesen oder Korrektur lesen lassen (s. auch *Dos and donts: Hinweise zur Abfassung wissenschaftlicher Arbeiten* von Prof. Frank – [http://www.cl.uni-heidelberg.de/~frank/materials/dos\\_and\\_donts.pdf](http://www.cl.uni-heidelberg.de/~frank/materials/dos_and_donts.pdf)). Ich schätze Wikipedia als Gegenstand meiner Forschung sehr, nicht aber als Quelle für wissenschaftliche Arbeiten. Hausarbeiten, die Wikipedia (oder auch andere allgemeine Enzyklopädien) als Beleg zitieren, werde ich zurückweisen. Bitte lesen und zitieren Sie Fachliteratur!

Seminararbeit (d.i. eine praktische Arbeit) ist auch möglich. Sollte durch 5-6 Seiten Bericht begleitet werden.

**Abgabetermin:** bis spätestens 1. März 2012; per Email als PDF-Datei (kein Microsoft Word!) oder ausgedruckt per Post – Matrikelnummer und Studiengang nicht vergessen!

## Literatur

- Abbott, Rob, Marilyn Walker, Pranav Anand, Jean E. Fox Tree, Robeson Bowmani & Joseph King (2011). How can you say such things?!?: Recognizing disagreement in informal political argument. In *ws-lsm-11*, pp. 2–11.
- Abdul-Mageed, Muhammad, Mona Diab & Mohammed Korayem (2011). Subjectivity and sentiment analysis of modern standard Arabic. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 587–591.
- Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow & Rebecca Passonneau (2011). Sentiment analysis of Twitter data. In *Proceedings of the ACL-11 Workshop on Language in Social Media*, Portland, Oreg., 23 June 2011, pp. 30–38.
- Alm, Cecilia Ovesdotter, Dan Roth & Richard Sproat (2005). Emotions from text: Machine learning for text-based emotion prediction. In *Proceedings of the Human Language Technology Conference and the 2005 Conference on Empirical Methods in Natural Language Processing*, Vancouver, B.C., Canada, 6–8 October 2005, pp. 579–586.
- Baccianella, Stefano, Andrea Esuli & Fabrizio Sebastiani (2010). SentWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the 7th International Conference on Language Resources and Evaluation*, La Valetta, Malta, 17–23 May 2010, pp. 2200–2204.
- Bachenko, Joan, Eileen Fitzpatrick & Michael Schonwetter (2008). Verification and implementation of language-based deception indicators in civil and criminal narratives. In *Proceedings of the 22nd International Conference on Computational Linguistics*, Manchester, U.K., 18–22 August 2008, pp. 41–48.
- Balasubramanian, Ramnath, William W. Cohen, Doug Pierce & David P. Redlawsk (2011). What pushes their buttons? predicting polarity from the content of political blog posts. In *Proceedings of the ACL-11 Workshop on Language in Social Media*, Portland, Oreg., 23 June 2011, pp. 12–19.
- Bansal, Mohit, Claire Cardie & Lillian Lee (2008). The power of negative thinking: Exploiting label disagreement in the min-cut classification framework. In *Proceedings of the Poster Session at the 22nd International Conference on Computational Linguistics*, Manchester, U.K., 18–22 August 2008, pp. 15–18.
- Batliner, Anton, Stefan Steidl, Christian Hacker & Elmar Nöth (2008). Private emotions vs. social interactions – a data driven approach towards analysing emotion in speech. *User Modelling and User-Adapted Interaction*, 18:175–206.
- Beineke, Philip, Trevor Hastie, Christopher Manning & Shivakumar Vaithyanathan (2004). Exploring sentiment summarization. In *Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*, Palo Alto, Cal., 22–24 March 2004, pp. 12–15.
- Blitzer, John, Mark Dredze & Fernando Pereira (2007). Biographies, Bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, Prague, Czech Republic, 23–30 June 2007, pp. 440–447.
- Bollegala, Danushka, David Weir & John Carroll (2011). Using multiple sources to construct a sentiment sensitive thesaurus for cross-domain sentiment classification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 132–141.
- Branavan, S.R.K., Harr Chen, Jacob Eisenstein & Regina Brazilay (2008). Learning document-level semantic properties from free-text annotations. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Columbus, Ohio, 15–20 June 2008, pp. 263–271.
- Burgoon, Judee K., J.P. Blair, Tiantian Qin & Jr. Nunamaker, Jay F. (2003). Detecting deception through linguistic analysis. In *Proceedings of Intelligence and Security Informatics (ISI)*, p. 958.
- Carbonell, Jaime G. (1979). *Subjective Understanding: Computer Models of Belief Systems*,

- (Ph.D. thesis). Yale University.
- Cardie, Claire, Janyce Wiebe, Theresa Wilson & Diane Litman (2003). Combining low-level and summary representations of opinions for multi-perspective question answering. In *Proceedings of the AAAI Spring Symposium on New Directions in Question Answering*, pp. 20–27.
- Carenini, Giuseppe, Raymond Ng & Adam Pauls (2006). Multi-document summarization of evaluative text. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, 3–7 April 2006, pp. 305–312.
- Carvalho, Paula, Luís Sarmento, Jorge Teixeira & Mário J. Silva (2011). Liars and saviors in a sentiment annotated corpus of comments to political debates. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 564–568.
- Dave, Kushal, Steve Lawrence & David M. Pennock (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th World Wide Web Conference*, Budapest, Hungary, 20–24 May 2003, pp. 519–528.
- Devitt, Ann & Khurshid Ahmad (2007). Sentiment analysis in financial news: A cohesion-based approach. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, Prague, Czech Republic, 23–30 June 2007, pp. 984–991.
- Eguchi, Koji & Victor Lavrenko (2006). Sentiment retrieval using generative models. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, Sydney, Australia, 22–23 July 2006, pp. 345–354.
- Eguchi, Koji & Chirag Shah (2006). Opinion retrieval experiments using generative models: Experiments for the TREC 2006 blog track. In *Proceedings of the Fifteenth Text REtrieval Conference*, Gaithersburg, Md., 14–17 November 2006.
- El Ayadi, Moataz, Mohamed S. Kamel & Fakhri Karray (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3):572–587.
- Esuli, Andrea & Fabrizio Sebastiani (2006a). Determining term subjectivity and term orientation for opinion mining. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, 3–7 April 2006, pp. 193–200.
- Esuli, Andrea & Fabrizio Sebastiani (2006b). SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa, Italy, 22–28 May 2006.
- Esuli, Andrea & Fabrizio Sebastiani (2007). PageRanking WordNet synsets: An application to opinion mining. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, Prague, Czech Republic, 23–30 June 2007, pp. 424–431.
- Gamon, Michael (2004). Sentiment classification on customer feedback data: Noisy data, large feature vectors, and the role of linguistic analysis. In *Proceedings of the 20th International Conference on Computational Linguistics*, Geneva, Switzerland, 23–27 August 2004, pp. 611–617.
- Goldberg, Andrew B. & Jerry Zhu (2006). Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization. In *TextGraphs: HLT/NAACL Workshop on Graph-based Algorithms for Natural Language Processing*, pp. 45–53.
- González-Ibanez, Roberto, Smaranda Muresan & Nina Wacholder (2011). Identifying sarcasm in Twitter: A closer look. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 581–586.
- Graciarena, Martin, Elizabeth Shriberg, Andreas Stolcke, Frank Enos, Julia Hirschberg & Sachin Kajarekar (2006). Combining prosodic, lexical and cepstral systems for deceptive speech detection. In *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France, 15–19 June 2006, pp. 1033–1036.
- Grefenstette, Gregory, Yan Qu, James G. Shanahan & David A. Evans (2004). Coupling niche browsers and affect analysis for an opinion mining application. In *Proceedings of Recherche d'Information Assistée par Ordinateur (RIA/O)*, Avignon, France, 26–28 April 2004.
- Hancock, Jeffrey T., Lauren E. Curry, Saurabh Goorha & Michael Woodworth (2008). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, 45:1–23.

- Hassan, Ahmed, Amjad Abu-Jbara, Rahul Jha & Dragomir Radev (2011). Identifying the semantic orientation of foreign words. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 592–597.
- Hatzivassiloglou, Vasileios & Kathleen R. McKeown (1997). Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and of the 8th Conference of the European Chapter of the Association for Computational Linguistics*, Madrid, Spain, 7–12 July 1997, pp. 174–181.
- Hatzivassiloglou, Vasileios & Janyce M. Wiebe (2000). Effects of adjective orientation and gradability on sentence subjectivity. In *Proceedings of the 18th International Conference on Computational Linguistics*, Saarbrücken, Germany, 31 July – 4 August 2000, pp. 299–305.
- He, Yulan, Chengua Lin & Harith Alani (2011). Automatically extracting polarity-bearing topics for cross-domain sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 123–131.
- Hu, Minqing & Bing Liu (2004). Mining opinion features in customer reviews. In *Proceedings of the 19th National Conference on Artificial Intelligence*, San Jose, Cal., 25–29 July 2004, pp. 755–760.
- Hurst, Matthew & Kamal Nigam (2004). Retrieving topical sentiments from online document collections. In *Document Recognition and Retrieval XI*, pp. 27–34.
- Inanoglu, Zeynep & Steve Young (2009). Data-driven emotion conversion in spoken English. *Speech Communication*, 51:268–283.
- Jiang, Long, Mo Yu, Ming Zhou, Xiaohua Liu & Tiejun Zhao (2011). Target-dependent Twitter sentiment classification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 151–160.
- Johansson, Richard & Alessandro Moschitti (2010). Reranking models in fine-grained opinion analysis. In *Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, China, 23–27 August 2010, pp. 519–527.
- Johansson, Richard & Alessandro Moschitti (2011). Extracting opinion expressions and their polarities – Exploration of pipelines and joint models. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 101–106.
- Kaji, Nobuhiro & Masaru Kitsuregawa (2007). Building lexicon for sentiment analysis from massive collection of HTML documents. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Language Learning*, Prague, Czech Republic, 28–30 June 2007, pp. 1075–1083.
- Kim, Soo-Min & Eduard Hovy (2006a). Automatic identification of pro and con reasons in online reviews. In *Proceedings of the Poster Session at the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, Sydney, Australia, 17–21 July 2006, pp. 483–490.
- Kim, Soo-Min & Eduard Hovy (2006b). Identifying and analyzing judgment opinions. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, New York, N.Y., 4–9 June 2006, pp. 200–207.
- Kim, Soo-Min & Eduard Hovy (2007). Crystal: Analyzing predictive opinions on the web. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Language Learning*, Prague, Czech Republic, 28–30 June 2007, pp. 1056–1064.
- Ku, Lun-Wei, Yu-Ting Liang & Hsin-Hsi Chen (2006). Tagging heterogeneous evaluation corpora for opinionated tasks. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa, Italy, 22–28 May 2006.
- Kudo, Taku & Yuji Matsumoto (2004). A boosting algorithm for classification of semi-structured text. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 25–26 July 2004, pp. 230–237.
- Laver, Michael, Kenneth Benoit & John Garry (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2):311–331.
- Lin, Chenghua, Yulan He & Richard Everson (2011). Sentence subjectivity detection with weakly-supervised learning. In *Proceedings of the 5th International Joint Conference on*



- Natural Language Processing*, Chiang Mai, Thailand, 8–13 November 2011, pp. 1153–1161.
- Lin, Wei-Hao & Alexander Hauptmann (2006). Are these documents written from different perspectives? A test of different perspectives based on statistical distribution divergence. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, Sydney, Australia, 17–21 July 2006, pp. 1057–1064.
- Lin, Wei-Hao, Theresa Wilson, Janyce M. Wiebe & Alexander Hauptmann (2006). Which side are you on? Identifying perspectives at the document and sentence levels. In *Proceedings of the 10th Conference on Computational Natural Language Learning*, New York, N.Y., USA, 8–9 June 2006, pp. 109–116.
- Liu, Hugo, Henry Lieberman & Ted Selker (2003). A model of textual affect sensing using real-world knowledge. In *Proceedings of Intelligent User Interfaces (IUI)*, pp. 125–132.
- Lu, Bin, Chenhao Tan, Claire Cardie & Benjamin K. Tsou (2011). Joint bilingual sentiment classification with unlabeled parallel corpora. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 320–330.
- Maas, Andrew L., Raymond L. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng & Christopher Potts (2011). Learning word vectors for sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 142–150.
- Martin, Lanny W. & Georg Vanberg (2008). A robust transformation procedure for interpreting political text. *Political Analysis*, 16(1):93–100.
- Mei, Qiaozhu, Xu Ling, Matthew Wondra, Hang Su & ChengXiang Zhai (2007). Topic sentiment mixture: Modeling facets and opinions in weblogs. In *Proceedings of the 16th World Wide Web Conference*, Banff, Canada, 8–12 May, 2007, pp. 171–180.
- Mihalcea, Rada, Carmen Banea & Janyce M. Wiebe (2007). Learning multilingual subjective language via cross-lingual projections. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, Prague, Czech Republic, 23–30 June 2007, pp. 976–983.
- Mullen, Tony & Nigel Collier (2004). Sentiment analysis using Support Vector Machines with diverse information sources. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 25–26 July 2004, pp. 412–418.
- Mullen, Tony & Robert Malouf (2006). A preliminary investigation into sentiment analysis of informal political discourse. In *AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW)*, pp. 159–162.
- Mullen, Tony & Robert Malouf (2008). Taking sides: User classification for informal online political discourse. *Internet Research*, 18:177–190.
- Murray, Gabriel & Giuseppe Carenini (2011). Subjectivity detection in spoken and written conversations. *Natural Language Engineering*, 17(3):397–418.
- Ng, Vincent, Sajib Dasgupta & S.M. Niaz Arifin (2006). Examining the role of linguistic knowledge sources in the automatic identification and classification of reviews. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, Sydney, Australia, 17–21 July 2006, pp. 611–618.
- Ounis, Iadh, Maarten de Rijke, Craig Macdonald, Gilad Mishne & Ian Soboroff (2006). Overview of the TREC-2006 blog track. In *Proceedings of the Fifteenth Text REtrieval Conference*, Gaithersburg, Md., 14–17 November 2006.
- Pang, Bo & Lillian Lee (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*, Barcelona, Spain, 21–26 July 2004, pp. 272–279.
- Pang, Bo & Lillian Lee (2005). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, Ann Arbor, Mich., 25–30 June 2005, pp. 115–124.
- Pang, Bo & Lillian Lee (2008a). Opinion mining and sentiment analysis. *Foundations and*

- Trends in Information Retrieval*, 2(1-2):1–135.
- Pang, Bo & Lillian Lee (2008b). Using very simple statistics for review search: An exploration. In *Proceedings of the 22nd International Conference on Computational Linguistics*, Manchester, U.K., 18–22 August 2008, pp. 75–78.
- Pang, Bo, Lillian Lee & Shivakumar Vaithyanathan (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, Philadelphia, Penn., 6–7 July 2002, pp. 79–86.
- Park, Souneil, Kyungsoo Lee & Junewha Song (2011). Contrasting opposing views of news articles on contentious issues. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 340–349.
- Qin, Bin, Yanyan Zhao, Leilei Gao & Ting Liu (2008). Recommended or not? give advice on online products. In *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, Jinan, Shandong, China, 18–20 October 2008, pp. 208–212.
- Ranganath, Rajesh, Dan Jurafsky & Dan McFarland (2009). It’s not you, it’s me: Detecting flirting and its misperception in speed-dates. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, Singapore, 6–7 August 2009, pp. 334–342.
- Riloff, Ellen & Janyce Wiebe (2003). Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, Sapporo, Japan, 11–12 July 2003, pp. 105–112.
- Riloff, Ellen, Janyce Wiebe & Theresa Wilson (2003). Learning subjective nouns using extraction pattern bootstrapping. In *Proceedings of the 7th Conference on Computational Natural Language Learning*, Edmonton, Alberta, Canada, 31 May – 1 June 2003, pp. 25–32.
- Sauper, Christina, Aria Haghighi & Regina Barzilay (2011). Content models with attitude. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 350–358.
- Seki, Yohei, Koji Eguchi & Noriko Kando (2004). Analysis of multi-document viewpoint summarization using multi-dimensional genres. In *Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*, Palo Alto, Cal., 22–24 March 2004, pp. 142–145.
- Seki, Yohei, Koji Eguchi, Noriko Kando & Masaki Aono (2005). Multi-document summarization with subjectivity analysis at DUC 2005. In *Proceedings of the 2005 Document Understanding Conference held at the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, Vancouver, B.C., Canada, 9–10 October 2005.
- Seki, Yohei, Koji Eguchi, Noriko Kando & Masaki Aono (2006). Opinion-focused summarization and its analysis at DUC 2006. In *Proceedings of the 2006 Document Understanding Conference held at the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, New York, N.Y., 8–9 June 2006.
- Somasundaran, Swapna, Galileo Namata, Janyce Wiebe & Lise Getoor (2009). Supervised and unsupervised methods in employing discourse relations for improving opinion polarity classification. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, Singapore, 6–7 August 2009.
- Somasundaran, Swapna, Josef Ruppenhofer & Janyce M. Wiebe (2008a). Discourse level opinion relations: An annotation study. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, Columbus, Ohio, 19–20 June 2008, pp. 129–137.
- Somasundaran, Swapna & Janyce Wiebe (2009). Recognizing stances in online debates. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing*, Singapore, 2–7 August 2009.
- Somasundaran, Swapna, Janyce M. Wiebe & Josef Ruppenhofer (2008b). Discourse level opinion interpretation. In *Proceedings of the 22nd International Conference on Computational Linguistics*, Manchester, U.K., 18–22 August 2008, pp. 801–808.
- Su, Fangzhong & Katja Markert (2009). Subjectivity recognition on word senses via semi-supervised mincuts. In *Proceedings of Human Language Technologies 2009: The Conference*

- of the North American Chapter of the Association for Computational Linguistics, Boulder, Col., 31 May – 5 June 2009, pp. 1–9.
- Subasic, Pero & Alison Huettner (2001). Affect analysis of text using fuzzy semantic typing. *IEEE Transactions on Fuzzy Systems*, 9(4):483–496.
- Taboada, Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll & Manfred Stede (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(4):267–307.
- Täckström, Oscar & Ryan McDonald (2011a). Discovering fine-grained sentiment with latent variable structured prediction models. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 569–574.
- Täckström, Oscar & Ryan McDonald (2011b). Semi-supervised latent variable models for sentence-level sentiment analysis. In *Proceedings of the 33rd European Conference on Information Retrieval*, Dublin, Ireland, 18–21 April 2011, p. ??
- Takamura, Hiroya, Takashi Inui & Manabu Okumura (2006). Latent variable models for semantic orientations of phrases. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, 3–7 April 2006, pp. 201–208.
- Titov, Ivan & Ryan McDonald (2008a). A joint model of text and aspect ratings for sentiment summarization. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Columbus, Ohio, 15–20 June 2008, pp. 308–316.
- Titov, Ivan & Ryan McDonald (2008b). Modeling online reviews with multi-grain topic models. In *Proceedings of the 17th World Wide Web Conference*, Beijing, China, 21–25 April, 2008, pp. 111–120.
- Turney, Peter (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, Penn., 7–12 July 2002, pp. 417–424.
- Turney, Peter D. & Michael L. Littman (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4):315–346.
- Wang, Dong & Yang Liu (2011). A pilot study of opinion summarization in conversations. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 331–339.
- Wei, Bin & Christopher Pal (2010). Cross lingual adaption: An experiment on sentiment classification. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, 11–16 July 2010, pp. 258–262.
- Whitelaw, Casey, Navendu Garg & Shlomo Argamon (2005). Using appraisal groups for sentiment analysis. In *Proceedings of the ACM SIGIR Conference on Information and Knowledge Management (CIKM)*, pp. 625–631.
- Wiebe, Janyce (1990). *Recognizing subjective sentences: A computational investigation of narrative text*, (Ph.D. thesis). SUNY Buffalo, Department of Computer Science. Also published as Technical report 90-03.
- Wiebe, Janyce, Eric Breck, Christopher Buckley, Claire Cardie, Paul Davis, Bruce Fraser, Diane Litman, David Pierce, Ellen Riloff, Theresa Wilson, David Day & Mark Maybury (2003). Recognizing and organizing opinions expressed in the world press. In *Proceedings of the AAI Spring Symposium on New Directions in Question Answering*, pp. 12–19.
- Wiebe, Janyce M. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20(2):233–287.
- Wiebe, Janyce M. & Rebecca Bruce (1995). Probabilistic classifiers for tracking point of view. In *Proceedings of the AAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, pp. 181–187.
- Wiebe, Janyce M., Rebecca F. Bruce & Thomas P. O’Hara (1999). Development and use of a gold-standard data set for subjectivity classification. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, College Park, Md., 20–26 June 1999, pp. 246–253.
- Wiebe, Janyce M. & Ellen Riloff (2005). Creating subjective and objective sentence classifiers

- from unannotated texts. In *Proceedings of the Conference on Computational Linguistics and Intelligent Text Processing (CICLing)*, pp. 486–497.
- Wiebe, Janyce M., Theresa Wilson, Rebecca Bruce, Matthew Bell & Melanie Martin (2004). Learning subjective language. *Computational Linguistics*, 30(3):277–308.
- Wiebe, Janyce M., Theresa Wilson & Claire Cardie (2005). Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2/3):164–210.
- Wilks, Yorick & Janusz Bien (1984). Beliefs, points of view and multiple environments. In *Proceedings of the International NATO Symposium on Artificial and Human Intelligence*, pp. 141–171.
- Wilson, Theresa & Janyce Wiebe (2003). Annotating opinions in the world press. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, Sapporo, Japan, 5-6 July 2003, pp. 13–22.
- Wilson, Theresa & Janyce Wiebe (2005). Annotating attributions and private states. In *Proceedings of the Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*, Ann Arbor, Mich., 29 June 2005, pp. 53–60.
- Wilson, Theresa, Janyce Wiebe & Rebecca Hwa (2004). Just how mad are you? finding strong and weak opinion clauses. In *Proceedings of the 19th National Conference on Artificial Intelligence*, San Jose, Cal., 25–29 July 2004, pp. 761–769.
- Wilson, Theresa, Janyce M. Wiebe & Paul Hoffmann (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Human Language Technology Conference and the 2005 Conference on Empirical Methods in Natural Language Processing*, Vancouver, B.C., Canada, 6–8 October 2005, pp. 347–354.
- Yesselina, Ainur, Yejin Choi & Claire Cardie (2010). Automatically generating annotator rationales to improve sentiment classification. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, 11–16 July 2010, pp. 336–341.
- Yu, Hong & Vasileios Hatzivassiloglou (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, Sapporo, Japan, 11–12 July 2003, pp. 129–136.
- Yu, Jianxing, Zheng-Jun Zha, Meng Wang & Tat-Seng Chua (2011). Aspect ranking: Identifying important product aspects from online consumer reviews. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oreg., USA, 19–24 June 2011, pp. 1496–1505.
- Yu, Kai, François Mairesse & Steve Young (2010). Word-level emphasis modelling in HMM-based speech synthesis. In *Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing*.
- Zhang, Lei & Bing Liu (2011). Identifying noun product features that imply opinions. In *Proceedings of the ACL 2011 Conference Short Papers*, Portland, Oreg., USA, 19–24 June 2011, pp. 575–580.
- Zhang, Zhu (2008). Weighing stars: Aggregating online product reviews for intelligent E-commerce applications. *IEEE Intelligent Systems*, 23(5):42–49.
- Zhou, Liang & Eduard Hovy (2005). Digesting virtual "Geekculture: The summarization of technical internet relay chats. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, Ann Arbor, Mich., 25–30 June 2005, pp. 298–305.
- Zhou, Lina, Judee K. Burgeon & Douglas P. Twitchell (2003). A longitudinal analysis of language behavior of deception in e-mail. In *Proceedings of the First NSF/NIJ Symposium on Intelligence and Security Informatics*, Tucson, Ariz., 2003, p. ??
- Zirn, Căcilia, Matthias Niepert, Heiner Stuckenschmidt & Michael Strube (2011). Fine-grained sentiment analysis with structural features. In *Proceedings of the 5th International Joint Conference on Natural Language Processing*, Chiang Mai, Thailand, 8–13 November 2011, pp. 336–344.