

# Rethinking Sentiment Analysis in the News: from Theory to Practice and back

Alexandra Balahur<sup>1,2</sup>, Ralf Steinberger<sup>1</sup>

<sup>1</sup> European Commission – Joint Research Centre

<sup>2</sup> University of Alicante, Department of Software and Computing Systems  
{Alexandra.Balahur-Dobrescu, Ralf.Steinberger}@jrc.ec.europa.eu

**Abstract.** Most sentiment analysis work has been carried out on highly subjective text types where the target is clearly defined and unique across the text (movie or product reviews). However, when applying sentiment analysis to the news domain, it is necessary to clearly define the scope of the task, in a more specific manner than it has been done in the field so far. The main tasks we identified for news opinion mining are: definition of the target; separation of the good and bad news content from the good and bad sentiment expressed on the target; and analysis of clearly marked opinion that is expressed explicitly, not needing interpretation or the use of world knowledge. It is furthermore important to distinguish three different possible views on newspaper articles – author, reader and text. These have to be addressed differently at the time of analysing sentiment, especially the case of author intention and reader interpretation, where specific profiles must be defined if the proper sentiment is to be extracted.

**Keywords:** sentiment analysis, opinion mining, news analysis, reported speech, task definition.

## 1 Introduction

Humans are social beings. They cannot reach the level of what we call “human” unless they develop in organised societies, where they are taught norms, rules and laws governing the existence and co-existence of people. Although most of the times unconsciously, we continuously shape our behaviour and attitudes on the basis of these social conventions, of public and private opinions and events of the world surrounding us. We give and accept advice as part of our every-day lives, as part of a ritual to knowing, better understanding and integrating into our surrounding reality. Our reactions are based on what we expect. Moreover, norms of what is allowed and what is not, what is generally expected to be done or not in a context, trigger our own emotional response and our attitude towards the situations (Ratner, 2000), (Goleman, 1995).

Together with the development of technology and the growing access to information, we have witnessed the birth of a new type of society - that of the interaction and communication (Wiberg, 2004). In this new context, the role of emotion has become crucial. Facts determine emotion in people, who absorb facts and express the effect these facts have on them. Other persons have access to these expressions of sentiments and facts, which they, in their turn, transform under the influence of their own emotional perception. Practically, access to information has also given way to access to emotional response to information, in the light of which information changes. Thus, people react to both facts and attitude on facts. And while society is changing, norms are changing along with it, and world attitude shapes new norms, under which societies further change. We can think of numerous examples of facts, on which the society's and the people's opinion has changed over time and we can also think of facts that remain tabu subjects in many communities around the world. It is thus both interesting, as well as challenging to see what the opinion is on certain subjects, so that trends can be predicted and the right measures taken.

The automatic processing of texts to detect opinion expressed therein, as a unitary body of research, has been denominated *opinion mining* or *sentiment analysis*. Most work on sentiment analysis has been carried out on highly subjective text types such as blogs and product or movie reviews. Authors of such text types mostly express their opinions quite freely. News articles have received much less attention, although news bias across different news sources has been discussed by a few and some initial efforts have concentrated on sentiment analysis in the news area (Fortuna et al., 2009), (Belyaeva and Van der Goot, 2009). News articles and other media reports typically contain much less clearly stated opinions. Although support or criticism are sometimes expressed, the bias or sentiment of the journalist is often expressed indirectly, for instance by highlighting some facts while possibly omitting others or by the choice of words. It may seem that this simply makes opinion mining harder, but by trying to apply known sentiment analysis technology to news texts, we also realised that the task needs to be redefined.

In the remainder of the introduction, we motivate our interest in the task, describe the news data we work with, and discuss some examples highlighting some difficulties. In Section 2, we summarise how others have defined sentiment analysis and how they applied their techniques. Trying to fill the gap between our needs and the current state of affairs, we tempt to redefine the task in Section 3, and we describe our effort to create a new gold standard collection. The paper ends with a conclusion.

## **2 Motivation**

### **2.1 Motivation for applying opinion mining to news data**

The European Commission's (EC) Joint Research Centre (JRC) has developed a number of news analysis systems. Four of them are publicly accessible (see <http://emm.jrc.it/overview.html>), the first one of them has been online since 2002. They are referred to as the Europe Media Monitor (EMM) family of applications

(Steinberger et al. 2009). These four systems have in common that they process the news data gathered by the EMM news gathering engine. EMM gathers between 80,000 and 100,000 articles per day in about 50 languages, by scraping about 2,200 hand-selected online news sources and a few specialist websites. EC-internal users additionally have access to newswires from about 20 commercial news providers. EMM continuously visits news sites to get the latest news (the most active sites are visited every five minutes) and aggregates and processes them instantly. The live news sites *NewsBrief* and *MedISys* are updated every ten minutes so that users will always see the latest news from the fastest news providers around the world. All news get converted into a standard RSS format. They get classified according to hundreds of classes and named entity software detects and disambiguates mentions of persons, organisations and locations. Related news get grouped into news clusters, and the development of clusters is tracked over time. Unlike these two live systems, the *NewsExplorer* application provides a daily and a more long-term view of the news, and it collects and aggregates information about persons and organisations from the news in 19 languages. The EMM systems are used by European Union (EU) institutions, by national public bodies inside and outside the EU countries, by international organisations and by an average of 30,000 online users per day. From the news in 13 of these languages, an average 3165 reported speech quotations per day are automatically extracted (Pouliquen et al., 2007). The person issuing the quotation is extracted, and so is any entity that is being mentioned inside the quotation. EMM systems provide their users with news articles about subjects of their interest and with some automatically extracted detailed information. However, in addition to reading news about *facts*, some users have an interest in *opinion* expressed in the media, for instance to measure the public opinion about certain policy areas, about persons, organisations, programs or press releases. A comparison across news sources, countries and languages would be of particular interest.

## 2.2 A short history of news opinion mining at the JRC

Due to this user interest, first work on opinion mining from EMM data began in 2008. In parallel, efforts have started on processing blogs. First experiments were carried out on assigning document level sentiment values. The idea was to distinguish good from bad news, to identify trends over time for news on the same subject, and to possibly compare the reporting on the same events across languages. However, the question soon arose: even if we succeed, what do the results mean? We can expect, for instance, that news on natural disasters are usually bad, and a document level analysis will thus give negative values to the entities mentioned in such articles. According to a common definition, sentiment analysis is supposed to detect the opinion of a certain *source* (e.g. the speaker) regarding a certain *target* (e.g. another person, or a certain topic or product). For the task at hand, the source of a news article can simply be assumed to be the journalist or the publishing newspaper. However, in document level sentiment analysis, there is a need to define and identify the target more clearly, in order to avoid that entities get negative opinion results when being mentioned in news about negative events. In other words: document level sentiment analysis does not make the necessary distinction between a number of different

possible targets. Furthermore, it does not distinguish good or bad *news* from good or bad *sentiment* expressed in the article.

In a second effort towards opinion mining, we applied sentiment analysis to reported speech quotations from one person about another, with the objective of identifying support or criticism relationships between people. Tanev (2007) had developed software to detect such relationships in news texts and to construct signed social networks showing these relationships. However, such relationships are not so frequently made explicit in news text. In order to expand this signed social network with more positive or negative relationships, we developed software that classifies quotations of one person about another entity into expressing positive or negative sentiment. The idea was that a negative quotation of one person about another is equivalent to that person expressing criticism towards the other. For that purpose, a Bayesian (sentiment) classifier was trained on quotations involving all sets of persons for which an explicit support or criticism relationship was found. The classifier was then applied to other pairs of persons for which quotations have been found (Tanev et al., forthcoming). While experiments showed that the idea does work in principle, the accuracy of that process is not currently high enough for the application to go online. A likely explanation for this is that quotations mentioning a person are not necessarily mostly about that person, i.e. these persons may not be the target of the sentiment expression. Just like in the previously mentioned work on document level sentiment analysis, problems thus arise because not enough care is taken to clearly identify the target of the sentiment expression.

In a third set of experiments (Balahur et al. 2009), sentiment analysis was again applied to quotations of one person mentioning another entity. The motivation for this setup was that (a) we found that quotations typically contain more sentiment expressions than other parts of news articles and (b) source and target of the opinion are better defined than in document level opinion mining, although our second set of experiments has shown that entities mentioned in quotations are not necessarily the target of the sentiment. However, the setup proved to be useful in that it gave us a clear insight on the phenomena that we found in news articles and the difficulties of the task, as well as the pitfalls of wrongly defining what we are looking for as far as opinion mining is concerned. In the following sub-section, we show how difficult opinion mining in news analysis can be. In section 3, we will re-define the task to make news opinion mining more feasible.

### 2.3 Examples highlighting the difficulties of news opinion mining

During the second and third sets of experiments, where the task was to decide whether a quotation was positive, negative or neutral, we realised that the inter-annotator agreement even for the short pieces of text was relatively low (below 50%). The following is an example of a quotation where annotators had difficulty to agree:

(1) Politician A said: “We have declared a war on drugs”.

While one annotator may feel that this is a positive statement as it shows action and efficiency to overcome an existing drug-related problem, another annotator may interpret this statement as being negative because (a) ‘war’ is negative and (b) the

situation must be rather bad if an extreme reaction such as ‘declaring war’ is necessary. Depending on the background of the annotator, s/he may even want to argue in a certain context (c) that ‘drugs’ refers to soft drugs and that these are not a problem at all for society, in which case politician A’s attitude would probably be considered as being misled and erroneous. While the source of the quotation (politician A) is clear, there is thus some confusion regarding what the target is. Is it the energetic and problem-solving (or erroneous) attitude of the politician (positive attitude), or is it the alleged fact that there is a drug problem (negative news). A further issue is the question whether the confidence expressed in the politician’s statement should be considered as being positive, i.e. whether a statement such as “we will do something about this problem” should be considered a sentiment statement at all. It is clear that such a statement is intended to cause positive feelings towards the politician. Some existing sentiment or affect vocabulary lists do also include words like ‘war’ and ‘mother’ with the respective values negative or positive.

By adding one level to this example:

(2) Person B mocked politician A’s statement that “We have declared a war on drugs”.

yet another interpretation is possible, namely that ‘Person B’ is the source and ‘politician A’ is the target, questioning also somehow the positive sentiment politician A wanted to create with his/her statement. The journalist writing the article may furthermore express his or her own opinion on this, as in:

(3) Person B unreasonably mocked politician A’s statement that “We have declared a war on drugs”.

In this case, the journalist expresses negative sentiment towards ‘Person B’ for criticising ‘politician A’.

The chain could theoretically be continued: For instance, if the newspaper were a known defender of person A (and the corresponding political party and attitudes), the whole statement (3) could be interpreted as sarcasm, inverting the negative sentiment of ‘Person B’ towards ‘Politician A’, and so on. While this is clearly a constructed example, our low inter-annotator agreement and the clarifying discussions showed that our initial sentiment annotation instructions were under-specified and left too much leeway for interpretation. For these reasons, we re-defined our task and subsequently annotated a set of 1592 quotations, after which the inter-annotator agreement was 0.81%. Section 2 describes related work and especially sheds light on the sentiment definitions used by other researchers and practitioners. Section 3 then presents the annotation guidelines we used, and it gives some details on the gold-standard quotation collection we created according to these guidelines.

### 3 Definitions proposed in previous work

Several definitions of opinion mining have been proposed in the literature, but none of them, as we will see in this section, entirely fits our own purposes. We thus re-defined the task (see below).

Subjectivity analysis is defined by (Wiebe, 1994) as the “linguistic expression of somebody’s opinions, sentiments, emotions, evaluations, beliefs and speculations”. In her definition, the author was inspired by the work of the linguist Ann Banfield

(Banfield, 1982), who defines as subjective the “sentences that take a character’s point of view (Uspensky, 1973)” and that present private states (Quirk, 1985) (i.e. states that are not open to objective observation or verification) of an experiencer, holding an attitude, optionally towards an object. Subjectivity is opposed to objectivity, which is the expression of facts. As (Kim and Hovy, 2003) notice, opinion is subjective, but may not imply a sentiment. But what about our “war on drugs” example? Can facts express opinions? Is there a difference between interpretations of facts at sentiment level and the direct expression of sentiments? Should we take them into consideration?

(Esuli and Sebastiani, 2006) define opinion mining as a recent discipline at the crossroads of information retrieval and computational linguistics which is concerned not with the topic a document is about, but with the opinion it expresses. This is a very broad definition, that targets opinions expressed at a document level. As we have shown before, news articles contain mentions of different persons and events, the topic in itself might involve a negative tonality and both the author of the text, as well as the facts presented or the interpretation they are given by the reader may lead to a different categorisation of the document. So, this definition is not specific enough for us to understand what we should be looking for when annotating pieces of newspaper articles.

(Dave et al., 2003), define an opinion mining system as one that is able to “process a set of search results for a given item, generating a list of product attributes (quality, features, etc.) and aggregating opinions about each of them (poor, mixed, good).” Opinion mining, in this context, aims therefore at extracting and analysing judgements on various aspects of given products. A similar paradigm is given by (Hu, Liu 2004), which the authors entitle feature-based opinion mining. It is, however, not clear how statements such as “It broke in two days”, “The night photos are blurry”, that are actual fact information (according to the definition of subjectivity, they are verifiable) could be and should be annotated. Do they fall outside the goal of opinion mining? Since in our context, persons, organisations or events have no definable or inferable lists of features, this definition of the task does not help us in the context of opinion mining from newspaper articles.

(Kim and Hovy, 2005) define opinion as a quadruple (Topic, Holder, Claim, Sentiment), in which the Holder believes a Claim about the Topic, and in many cases associates a Sentiment, such as good or bad, with the belief. The authors distinguish among opinions with sentiment and opinions without sentiment and between directly and indirectly expressed opinions with sentiment. In this context, it does not remain clear how an example such as the “Local authorities have provided no help for the victims of the accident.” should be interpreted and why. Some might even argue that a statement they claim to be opinionated but with no sentiment – “Gap is likely to go bankrupt” (which would probably be interesting when assessing favorability in markets), has a sentiment and that sentiment is negative.

In the SemEval 2007 No. 14 Affective Text Task (Mihalcea and Strapparava, 2007), the systems were supposed to classify 1000 newspaper titles according to their valence and emotion contained. A title such as “Scientists proved that men’s perspiration raises women’s hormone levels” or “100 killed in bomb attack” were classified as negative. However, this is factual, verifiable information. Does this mean that when capturing the media sentiment, we should consider these results as being

negative? Do these statements refer to a fact and are we interested in the information conveyed or in the sought effect? If so, which of these aspects would we include in a system doing sentiment analysis from newspaper articles?

In other approaches, capturing favourability versus unfavourability, support versus opposition, criticism versus appreciation, liking versus disliking, even bad versus good news classification were considered to be sentiment analysis.

However, at the moment of annotating sentiment in newspaper articles, we have seen that combining all these aspects together did not help to clear what the task was and how annotation should be done. Even in the case of quotes, which are short pieces of text where the source was known and the possible targets were identified, expressions of opinion that needed some kind of interpretation or knowledge of the situation fell short of agreement, due to personal convictions, background and so on.

We further on then present an annotation effort for newspaper quotes, that shed light on the issue and helped define guidelines for labelling that led, from a level of agreement of under 50%, to a level of agreement of 81%. We give some details on the gold-standard quotation collection we created according to these guidelines. Finally, we redefine the task of sentiment analysis in the news, capturing the different aspects of sentiment in text that we identified and pinpointing what exactly we expect a sentiment analysis system to discover from news under the different aspects.

## 4 Redefining the task

Although some definitions of the task were proposed, none of them, as we have seen, could give an indication of the specific aspects that a system implementing opinion mining in a news context should contemplate. To clarify the task, we selected a collection of 1592 quotes (reported speech) from newspaper articles in English, whose source and target were known (their extraction patterns are designed with that scope) which we set out to annotate. The first experiments had an agreement of under 50%. Specifying that just the sentiment on the target should be annotated and separated from the good and bad news that was described led to an increase in the agreement up to 60%. We realised that by delimiting a few aspects, the task became much clearer. Following are the annotation guidelines we used:

### 4.1 Annotation guidelines

We are trying to decide whether the entity (person or organisation) in the text snippet is being talked about in a positive (POS) or in a negative light (NEG), or if the statement is rather objective/neutral (OBJ). We thus distinguish three cases of sentiment: two cases of subjectivity, in which case we can directly indicate the polarity (POS, NEG), and the case of non-subjectivity, objectivity or neutrality (OBJ). OBJ is the default, so no need to label neutral/objective examples.

Here are some more clarifications that may help:

1. If you can, try not to make use of your world knowledge, such as the political views of the entity. If you cannot decide without knowing the political views of the entity, just leave it neutral/objective (OBJ).
2. It may help to imagine that you are the one being talked about: would you like or dislike the statement without using your world knowledge?
3. We are not interested in knowing whether the whole piece of text is positive or negative, but exclusively the sentiment expressed towards the entity.
4. Another hint to identify subjectivity – those news items whose content cannot be verified and whose content is expressly changed to induce a negative/positive opinion should be annotated as positive or negative. E.g. “Negotiations with Turkey have been delayed” – factual (objective from a sentiment point of view) vs. “EU stalls negotiations with Turkey” – (subjective, negative sentiment).
5. Note that, in the text snippet “X supported Y for criticising Z”, there is negative sentiment towards Z from X and Y, but positive sentiment between X and Y.
6. Please try to separate good news vs. bad news from the sentiment expressed. We should NOT annotate the good versus bad content of the news. E.g. if the news talks about 7000 people being killed by a bomb, the news is factual/objective (OBJ), even if there is negative news content.
7. Expressions of attitude: “EU is willing to make efforts to prevent this from becoming a crisis” (This shows positive attitude, i.e. positive sentiment, POS); On the other hand, the sentence “EU sent help to the earthquake-affected Aquila citizens” is objective from a sentiment point of view (OBJ).
8. Should there be both positive and negative statements in the snippet, please consider the statement to be objective (OBJ). (strictly speaking, it would be subjective, but balanced; but we are not trying to distinguish this case).
9. It is certain that there will be many cases of doubt. In case of doubt, just leave the example un-annotated (neutral/objective, OBJ).

## 4.2 Annotation

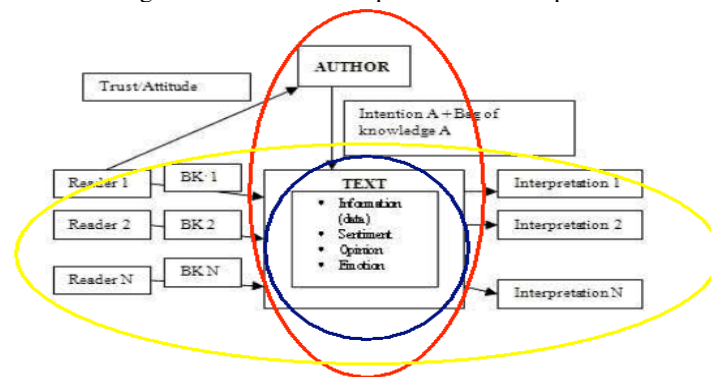
The original data set we decided to annotate contained 1592 quotes extracted from news in April 2008. The average final agreement was 81%, between 3 pairs of two annotators each.

Table 1. Results of the data annotation

	Number of quotes	Number of agreed quotes	Number of agreed negative quotes	Number of agreed positive quotes	Number of agreed objective quotes
	1592	1292	234	193	865
Agreement (%)		81%	0.78	0.78	0.83

The result of the annotation guidelines and labelling process was a corpus in which we agreed what sentiment was and what it was not. The number of agreed sentiment-containing quotes was one third of the total number of agreed quotes, showing that only clear, expressly stated opinion was marked, i.e. opinions that required no subjective interpretation from the annotator's part. The result of our labelling showed that in the case of newspapers, it is mandatory to distinguish between three different "components": the *author*, the *reader* and the *text* itself (Figure 1).

Figure 1: The three components of text opinion



While *the author* might convey certain opinions, by omitting or stressing upon some aspect of the text and by thus inserting their own opinion towards the facts, the spotting of such phenomena is outside the aim of sentiment analysis as we have defined it. Instead, such phenomena should be analysed as part of work on perspective determination or news bias research. From the *reader's point of view*, the interpretations of the text can be multiple and they depend on the personal

background knowledge, culture, social class, religion etc. as far as what is normal (expected) and what is not are concerned. Lastly, the opinion stated *strictly in the text* is the one that one should concentrate on at this level, being expressed directly or indirectly, by the target, towards the source, with all the information needed to draw this conclusion on polarity present in the text.

From the author and the reader's perspective and not from the text's pure informational point of view, opinion is conveyed through facts that are interpretable by the emotion they convey. However, emotions are not universal in their meaning. They are determined socially, culturally and historically. There are general emotions, but most of the times they relate to the norms, their significance and the cultural environment. Emotions imply an evaluation, which is both cognitive and affective, of a behaviour, with respect to a norm and the mutual expectation it raises. Some norms are common sense and are accepted and understood by all. Normative expectations link the behaviour (reaction) to a meaning and on this ground, by the understanding it is given. From the reader's point of view, sentiment analysis would be defined as *the assessment of a "target", based on its characteristics and factual information related to it, according to whether or not the results of the assessments are "according to" or "against" the "norm"*.

From the author's point of view, news bias or perspective determination should be concerned with discovering the ways in which expression of facts, word choice, omissions, debate limitations, story framing, selection and use of sources of quotes and the quote boundaries, for example, conveys a certain sentiment or not. The sentiment content of the text, finally, is what is expressly stated, and not what is left to be understood between the lines. Although pragmatics, through the speech act or other theories would argue there is no text that has no intended meaning, the sentiment or factual information conveyed is different from reader to reader and can thus not be done at a general level, as sentiment analysis intends to. For example, the text "The results of the match between Juventus Torino and Real Madrid last night are 3-0." would maybe be interpreted as something positive, a motive for pride in an Italian newspaper, it would be a negative, sad thing if reported by a Spanish source, it would be bad or good depending on whether or not an interested reader were pro or against the two teams and it would constitute just factual news from the strict point of view of the text. Given these three views one must be aware of at the time of constructing a sentiment analysis system for news, we can see that the task becomes much clearer and the agreement at the time of annotating texts, implementing and evaluating systems is higher.

Should one want to discover the possible interpretations of texts, sources' and readers' profiles must be defined and taken into consideration, for a whole understanding of the possible sentiment effects text has or is intended to have, and not just a general, often misunderstood one.

## 5 Conclusions

In this paper, we summarised our insights regarding sentiment classification for news. Realising that the inter-annotator agreement for sentiment annotation was very

low, we concluded that we need to clarify what sentiment analysis means for the news. We have seen that a) first of all there is a need to clearly define, before the annotation is done, what the source and the target of the sentiment are, subsequently b) separate the good and bad *news content* from the good and bad *sentiment* expressed on the target and, finally, c) annotate only clearly marked opinion that is expressed explicitly, not needing interpretation or the use of world knowledge. We have furthermore seen that there are three different possible views on newspaper articles – author, reader and text – and they have to be addressed differently at the time of analysing sentiment. This is especially the case with author intention and reader interpretation, where specific profiles must be defined if the proper sentiment is to be extracted.

At this moment, having the tasks clearly defined, we have started experimenting with adequate methods to perform sentiment analysis considering these insights.

## References

- Balahur Alexandra, Ralf Steinberger, Erik van der Goot, Bruno Pouliquen, Mijail Kabadjov (2009). Opinion Mining on Newspaper Quotations. Proceedings of the workshop 'Intelligent Analysis and Processing of Web News Content' (IAPWNC), held at the 2009 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology. Milano, Italy, 15.09.2009.
- Banfield, A. (1982). *Unspeakable sentences: Narration and Representation in the Language of Fiction*. Routledge and Kegan Paul.
- Belyaeva Evgenia, Erik van Der Goot. News bias of online headlines across languages. The study of conflict between Russia and Georgia. August 2008. *Rhetorics of the Media*. Conference Proceedings (2009) Lodz University Publishing House
- Esuli, A. and Sebastiani, F. (2006). Sentiwordnet: A publicly available resource for opinion mining. In Proceedings of the 6th International Conference on Language Resources and Evaluation
- Fortuna Blaz, Carolina Galleguillos and Nello Cristianini. Detecting the bias in media with statistical learning methods *Text Mining: Theory and Applications* (2009) Taylor and Francis Publisher.
- Goleman, D. (1995). *Emotional Intelligence*. Bantam Books.
- K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in Proceedings of WWW, pp. 519–528, 2003.
- Kim, S.-M. and Hovy, E. (2004). Determining the Sentiment of Opinions. In Proceedings of COLING 2004.
- Pouliquen Bruno, Ralf Steinberger & Clive Best (2007). Automatic Detection of Quotations in Multilingual News. In: Proceedings of the International Conference Recent Advances in Natural Language Processing (RANLP'2007), pp. 487-492. Borovets, Bulgaria, 27-29.09.2007.
- Ratner, C. (2000). A cultural-psychological analysis of emotions. *Culture and Psychology*, (6).
- Steinberger Ralf, Bruno Pouliquen & Erik van der Goot (2009). An Introduction to the Europe Media Monitor Family of Applications. In: Fredric Gey, Noriko Kando & Jussi Karlgren (eds.): *Information Access in a Multilingual World - Proceedings of the SIGIR 2009 Workshop (SIGIR-CLIR'2009)*, pp. 1-8. Boston, USA. 23 July 2009.
- Strapparava, C. and Mihalcea, R. (2007). Semeval 2007 task 14: Affective text. In Proceedings of ACL 2007.
- Tanev Hristo (2007). Unsupervised Learning of Social Networks from a Multiple-Source News Corpus. Proceedings of the Workshop Multi-source Multilingual Information Extraction

- and Summarization (MMIES'2007) held at RANLP'2007, pp. 33-40. Borovets, Bulgaria, 26 September 2007.
- Tanev Hristo, Bruno Pouliquen, Vanni Zavarella & Ralf Steinberger (forthcoming). Automatic Expansion of a Social Network Using Sentiment Analysis. In: *Annals of Information Systems, Special Issue on Data Mining for Social Network Data*.
- Uspensky, B. (1973). *A Poetics of Composition*. University of California Press, Berkeley, California.
- Wiberg, M. (2004). *The Interaction Society: Theories Practice and Supportive Technologies*. Idea Group Inc.
- Wiebe, J. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20.