

Annotating Irony in a Novel Italian Corpus for Sentiment Analysis

Andrea Gianti*, Cristina Bosco*, Viviana Patti*, Andrea Bolioli[◇], Luigi Di Caro*

*Dipartimento di Informatica, Università di Torino
Corso Svizzera, 195, 10149, Torino
andregianti@gmail.com, {bosco,dicaro,patti}@di.unito.it

[◇]CELI srl Torino
Via San Quintino 31, 10121, Torino
abolioli@celi.it

Abstract

In this paper we describe our current work on Senti-TUT, a novel Italian corpus for sentiment analysis. This resource includes annotations concerning both sentiment and morpho-syntax, in order to make available several possibilities of further exploitation related to sentiment analysis. For what concerns the annotation at sentiment level, we focus on irony and we selected therefore texts on politics from a social media, namely Twitter, where irony is usually applied by humans. Our aim is to add a new sentiment dimension, which explicitly accounts for irony, to a sentiment analysis classification framework based on polarity annotation.

The paper describes the data set, the features of the annotation both at sentiment and morpho-syntactic level, the procedures and tools applied in the annotation process. Finally, it shows the preliminary experiments we are carrying on in order to validate the annotation work.

Keywords: Irony, Sentiment analysis, Corpus annotation, Social media, Italian

1. Introduction and Motivation

In this paper we describe an ongoing project for the development of an Italian corpus annotated for sentiment analysis. We concentrate our attention on irony, a hard nut that is still to be cracked in the sentiment analysis context, and on a specific topic for texts where irony is usually applied by humans: politics.

Irony is recognized in literature as a specific phenomenon which can harm sentiment analysis and opinion mining systems (Pang and Lee, 2008; Davidov et al., 2011; Tsur et al., 2010). The rhetorical tradition treated irony as the figure of speech in which the meaning is the opposite of the literal meaning, so that an ironists primary intention is to communicate the opposite of what he/she says. Modern Gricean pragmatic theory has not departed radically from this view (Grice, 1975). Another interesting account of irony, the one proposed within relevance theory (Sperber and Wilson, 1986), suggests that irony is a variety of echoic use of language. This approach accounts for cases of “echoic irony”, where ironical utterances can be viewed as echoic mentions, in which usually the communicator dissociates herself from the opinion echoed.

The literature on irony and its interpretation is very extensive, however most of the proposals aim at explaining the fact that in an ironic sentence the explicit meaning is different or opposite from the real intended meaning. Therefore, in a sentiment analysis setting the presence of ironic devices in a text can work as an unexpected “polarity reverser”, by undermining the accuracy of the systems, especially in application contexts focussing on monitoring political sentiment, where blogs or social media provide the data sources. Recently, such application contexts gained popularity, since message content from social media (microblog-

ging like Twitter¹ especially) turned out to be a powerful real-time indicator of political sentiment. Microblogging messages, like “tweets” or Facebook messages, emerged as a very valuable information data not only in politics, but in a variety of NLP application domains, ranging from the extraction of critical information during times of mass emergency (Verma et al., 2011) to the sentiment analysis for the stock market prediction (Bollen et al., 2010).

However, Twitter communications includes a high percentage of ironic and sarcastic messages (Davidov et al., 2011; Tumasjan et al., 2011), and platforms monitoring the sentiment in Twitter messages experimented the problem to classify as positive many posts which instead express ironic non-positive judgments or opinions. As an example, let us consider the following tweet²:

TWSPINO-1160

*‘Alemanno: “Questa mattina sembra tutto funzionante”.
Gli hanno spiegato come funziona la pala’*

(Alemanno: “This morning everything seems to be working properly.” They’ve showed him how the shovel works)

In absence of irony recognition, such tweet it is classified as positive, while it clearly expresses a criticism w.r.t. the Rome’s mayor ability to deal with the snow emergency in Winter 2011-2012³.

In our tweets, we observed the presence of the well-known lexical devices and features that characterize humorous

¹<http://twitter.com>

²In february 2012, Rome’s mayor, Gianni Alemanno, was widely criticised in Italy for failing to activate an emergency plan after an exceptionally heavy snowfall.

³English translations of the Italian examples are mainly literal and so may sometimes appear awkward in English.

texts, like linguistic ambiguity, the use of affective terms, and so on, i.e. the tweet TWSPINO-32: *‘Marchionne presenta la nuova Panda. Il timore è che si diffonda tra la popolazione’* (Marchionne has presented the new Panda. It is feared that it may spread throughout the population).⁴ Moreover, we observed many cases of “echoic mentions” (Sperber and Wilson, 1986) among our ironic tweets. For instance in tweet TWNEWS-570 *‘Governo Monti: la rassicurante conferma che in Italia non esistono Tecnocrati, che non siano Gerontocrati. Non è un Paese per giovani’* (Monti’s government: the reassuring confirmation that in Italy do not exist Technocrats which are not Gerontocrats. No country for young men.) the sentence *‘non è un paese per giovani’* (no country for young men) is a case of echoic mention, with a clear reference to the title of the movie *‘Non è un Paese per Vecchi’* (No Country for Old Men⁵). The main aim of this project is to add a new sentiment dimension, which explicitly accounts for irony, to a sentiment analysis classification framework based on polarity annotation. To the best of our knowledge, existing sentiment analysis frameworks consider the following dimensions: subjectivity and objectivity; (positive or negative) polarity; emotional categories; opinions about entities. Accordingly, corpora that are manually annotated for subjectivity, polarity, or emotion, are available in many languages. Nowadays, with few exceptions (Esuli et al., 2008), Italian is among the less-resourced languages with respect to sentiment analysis. For what concerns English, let us mention the MPQA Opinion Corpus⁶, which contains news articles from a wide variety of news sources manually annotated for opinions and other private states (like emotions, sentiments, etc.). A multilingual dataset⁷, automatically annotated for subjectivity, in English, Arabic, French, German, Romanian, and Spanish, is the result of the work described in (Banea et al., 2010), while the multilingual corpus (Spanish, Italian and English) of blog posts in (Boldrini et al., 2010) is annotated according to the *EmotiBlog* annotation schema.

In the last years the authors gained experience both in sentiment analysis applied to social media (CELI and Me-Source, 2009), and in ontology-driven sentiment analysis applied to socially tagged resources (Baldoni et al., 2012), with a focus on the Italian language. Moreover, some among them are actively involved from more than ten years in both the development of linguistic resources morphosyntactically annotated, namely the treebank TUT (Bosco et al., 2000) (see Section 2.2.), and the exploitation of annotated data in several contexts for training and evaluation of NLP tools, see e.g. (Bosco and Mazzei, 2012b) and (Bosco and Mazzei, 2012a). On this line, we are now working to make available a novel Italian corpus for sentiment analysis, that we call Senti-TUT, which includes sentiment annotations concerning irony and consists in a collection of texts from social media. Such kind of resource is currently

⁴Marchionne is CEO of the Italian automotive group Fiat. Panda is the name of a Fiat city car.

⁵For details, see the Wikipedia page: [http://en.wikipedia.org/wiki/No_Country_for_Old_Men_\(film\)](http://en.wikipedia.org/wiki/No_Country_for_Old_Men_(film)).

⁶<http://www.cs.pitt.edu/mpqa/>

⁷<http://www.cse.unt.edu/rada/downloads.html#msa>

missing in particular for Italian. Moreover, we are carrying on some preliminary experiments in classification of our data in order to validate the annotation work.

The paper is organized as follows. In the next section we describe the corpus and the annotation we applied on it. Then, we discuss the preliminary experiments performed for the validation of data. The last section outlines some directions for future work.

2. Data

In this section we describe the data collected for the Senti-TUT project and the annotation we are applying on them. All the data related to the project and the information about download can be found in the Senti-TUT web site: <http://www.di.unito.it/~tutreeb/sentitut.html>.

2.1. The corpus

As confirmed by various references (Davidov et al., 2011) and (Tumasjan et al., 2011) social media, such as Facebook or Twitter, includes a high percentage of ironic and sarcastic messages and can mirror offline political sentiment, as they did for instance in the recent USA and German elections. Our linguistic data are therefore mainly collected by Twitter.

As far as the text style is concerned, in general, Twitter communications are composed by messages called “tweets”, each of which is shorter than 140 characters and can be composed by one or more sentences. In our Italian corpus of messages most of tweets are composed by two short sentences or simple noun phrases, and very rarely by wh-sentences. The typical structure of a tweet is shown in the following post⁸:

TWSPINO-107

‘Napolitano: “Attenti a toccare la Costituzione”.

Bisogna aspettare il medico legale.’

(Napolitano: “Be careful you don’t touch the Constitution”.

We have to wait for the forensic surgeon to arrive first.)

With respect to the composition and size of the data set, it is organized in two subcorpora, namely TWNEWS and TWSPINO. The former is currently composed of around three thousands of tweets, published in the weeks after the new Italian prime minister Mario Monti announced his Cabinet (from October 2011 the 16th to February 2012 the third). The latter is instead composed of more than one thousand tweets extracted from the Twitter section of Spinoza, published from July 2009 to February 2012. Spinoza⁹, is a very popular collective Italian blog which includes a high percentage of posts with sharp satire on politics, which is published on Twitter since 2009. This subcorpus has been therefore added in order to enlarge our data set with texts where various forms of irony are involved. The collection of all the data has been done by exploiting a collaborative annotation tool, which is part of the Blogmeter social media monitoring platform (CELI and Me-Source,

⁸Giorgio Napolitano is the current President of the Italian Republic.

⁹<http://www.spinoza.it/>

```

1 La (IL ART DEF F SING) [7;VERB-SUBJ]
2 spazzatura (SPAZZATURA NOUN COMMON F SING) [1;DET+DEF-ARG]
3 di (DI PREP MONO) [2;PREP-RMOD]
4 Napoli (NAPOLI NOUN PROPER F SING ££CITY) [3;PREP-ARG]
5 si (SI PRON REFL-IMPERS ALLVAL ALLVAL 3 LSUBJ+LOBJ+LIOBJ CLITIC) [7;VERB-OBJ]
6 sta (STAREVERB AUX IND PRES INTRANS 3 SING) [7;AUX]
7 decomponendo (DECOMPORREVERB MAIN GERUND PRES TRANS) [0;TOP-VERB]
8 . (#. PUNCT) [7;END]

1 Concorrerà (CONCORREREVERB MAIN IND FUT INTRANS 3 SING) [0;TOP-VERB]
1.10 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [1;VERB-SUBJ]
2 al (A PREP MONO) [1;VERB-INDCOMPL]
2.1 al (IL ART DEF M SING) [2;PREP-ARG]
3 Nobel (NOBEL NOUN PROPER) [2.1;DET+DEF-ARG]
4 per (PER PREP MONO) [3;PREP-RMOD]
5 la (IL ART DEF F SING) [4;PREP-ARG]
6 chimica (CHIMICA NOUN COMMON F SING) [5;DET+DEF-ARG]
7 . (#. PUNCT) [1;END]

```

Figure 1: The tweet 216 from the Spinoza corpus (TWSPINO-216) as annotated in TUT format.

2009). These data are only a portion of the whole material collected by this tool for the above mentioned periods (which are about 11,000 tweets).

2.2. The annotation

The project for the development of the Senti-TUT involves the annotation of the linguistic data with respect to two distinguished levels. While the first one includes morphological and syntactic tags as usual e.g. in treebanks, the second refers instead to concepts typical of sentiment analysis.

2.2.1. Morphological and syntactic annotation

For what concerns the morphological and syntactic annotation, this is done according to the format developed and applied in the Turin University Treebank (henceforth TUT) project (Bosco et al., 2000). This treebank is a freely available resource developed by the Natural Language Processing group of the University of Turin (for more details and examples see <http://www.di.unito.it/~tutreeb>) including 102,150 annotated tokens (around 3,500 sentences), which has been successfully exploited as testbed in various evaluation campaigns for Italian parsing (<http://www.evalita.it/>, (Bosco and Mazzei, 2012b) and (Bosco and Mazzei, 2012a)). We selected this format for two main reasons: the reliability of TUT format for the involved language and the availability of a variety of tools implemented within TUT project, first of all the Turin University Linguistic Environment (TULE, <http://www.tule.di.unito.it/>, (Lesmo, 2007) and (Lesmo, 2009)), whose pipeline includes tokenization, morphological and syntactic analysis.

In figure 1 and 2, a post extracted from our tweet corpus is represented according to TUT format: TWSPINO-216 ‘*La spazzatura di Napoli si sta decomponendo. Concorrerà al Nobel per la chimica.*’ (The garbage of Naples is becoming rotten. It will apply for the chemistry Nobel prize.). In particular, we can observe that TUT format is featured by a very detailed morphological tag set, which is useful for the description of a language with a rich inflection, and by a large inventory of grammatical relations exploited in the labeling of the edges of the dependency trees. For each

word, the lemma, the morphological category and related features are annotated together with the index of the father in the dependency tree and the relation linking the word with the father itself. Moreover, in order to offer an explicit representation of all the elements involved in the predicate argument structure, e.g. the subject which is often dropped in Italian, TUT format includes also null elements, see e.g. the annotation of the node 1.10 (t) which is the subject of the second sentence of the tweet represented in the figures. The morpho-syntactic annotation of the Senti-TUT corpus is automatically performed by TULE and then semi-automatically corrected by exploiting the tools developed within the TUT project. Nevertheless, the application of these tools, TULE especially, to the Senti-TUT corpus shows that, in order to achieve reliable annotations, the integration in the parsing process of various patterns typical of the social media language is needed. These patterns vary from the use of several citations from the Web to the words and phrases not formal or literary. Twitter, and social media in general, represent in fact a text genre different from those previously analyzed by exploiting TULE, e.g. newspaper or legal, which has never been analyzed in our knowledge for Italian. It is known in literature that in order to obtain a reliable morphological and syntactic analysis of a specific text genre, the parsing systems should be carefully tuned on the basis of it (Gildea, 2001). This is clearly showed by the current performance scores of TULE parser, which are far from those obtained on the text genres included in TUT, in particular with respect to the syntactic analysis. Nevertheless, the Evalita experiences showed evidences that TULE and other parsing systems for Italian can achieve, if trained and tuned, performances close to the state of the art for English for various text genres.

2.2.2. Annotation for sentiment analysis

As far as the annotation at the level useful for sentiment analysis is concerned, the data are currently annotated at tweet level, since one sentiment tag is applied to each tweet (considering that a tweet can be composed by more than one sentence). Nevertheless, even if, for the present time,

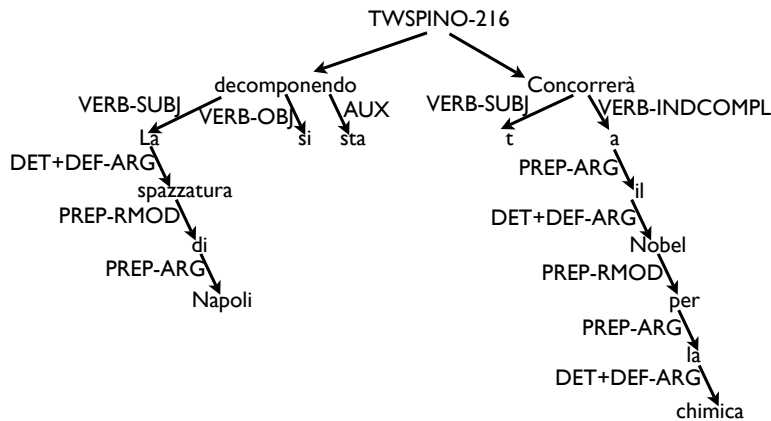


Figure 2: The TUT dependency tree for the tweet 216 from the Spinoza corpus (TWSPINO-216).

the focus of the Senti-TUT is mainly the annotation at tweet level, the resource we are currently developing has to be seen in the wider framework of a project for sentiment analysis and opinion mining. And within this context it should be considered also the availability of the morpho-syntactic annotation on the same data, which allows in the future for the application of other more fine-grained annotations and analysis related to sentiment analysis. For instance, the availability of Part of Speech tags and lemmas for words allows for investigations that relate morphological and sentiment features, e.g. adjective which are carried on sentimental meaning. As in (Tsur et al., 2010) syntactic features can be useful in the identification of irony, e.g. the use of punctuation.

In the table below the sentiment tags used for the annotation of Senti-TUT are described.

Sentiment tag	Meaning
POS	positive
NEG	negative
HUM	ironic
NONE	objective (none of the above)
MIXED	POS and NEG both

Table 1: The sentiment tags applied in Senti-TUT.

The following are examples of the annotation of tweets with the above mentioned sentiment tags.

TWSPINO-30 (tagged as HUM)
'C'e' cosi' tanta crisi che Babbo Natale invece delle letterine riceve curriculum.'
 (The economic crisis is so hard that Santa Claus receives curricula vitae instead of letters.)

TWNEWS-123 (tagged as NONE)
'Mario Monti premier? Tutte le indiscrezioni.'
 (Mario Monti premier? All the gossips.)

TWNEWS-24 (tagged as POS)
'Marc Lazar: "Napolitano? L'Europa lo ammira. Mario

Monti? Puo' salvare l'Italia'"
 (Marc Lazar: "Napolitano? Europe admires him. Mario Monti? He can save Italy")

TWNEWS-124 (tagged as NEG)
'Monti e' un uomo dei poteri che stanno affondando il nostro paese.'
 (Monti is a man of the powers that are sinking our country.)

TWNEWS (tagged as MIXED)
'Brindo alle dimissioni di Berlusconi ma sul governo Monti non mi faccio illusioni'
 (I drink a toast to the Berlusconi's resignation, but I have no illusion about the Monti's government)

We also used the tag UN in order to mark tweets which are not classifiable, e.g. tweets containing incomplete or meaningless sentences, which are therefore discarded. The distribution of the tags can be seen by observing a preliminary data set composed by around 1,500 tweets: around a third is classified as NONE, 400 as NEG, 300 as HUM, 250 as POS, and the remaining as MIXED or UN.

While the morpho-syntactic annotation is automatically performed by TULE, the annotation of the sentiment tags at the tweet level is currently manually performed by exploiting a collaborative annotation tool, which is part of the Blogmeter social media monitoring platform (CELI and Me-Source, 2009). Among the utilities made available by Blogmeter we applied, in particular, those related to filtering out the non relevant data, e.g. the *re-tweets* (i.e. the forwarded tweets).

Five human skilled annotators have been involved until now in this annotation task producing for each tweet not less than two independent annotations. This manual annotation helped by Blogmeter has been followed by an inter-annotator agreement check, as usual in the development of linguistic resources. In order to solve the disagreement, which can be referred to about 25% of the data, the independent annotation of a third human has been applied to the ambiguous tweets (i.e. those where each of the two annotators selected a tag different from the other annotator). The cases where the disagreement persists (i.e. tweets where

each of the three annotators selected a tag different from the others), which are around 3%, have been then discarded since considered as too ambiguous to be classified.

3. Preliminary experiments

We are carrying on some preliminary experiments in classification of our data in order to validate the annotation work. These experiments are based on a portion of the Senti-TUT corpus and more precisely on about 1,550 annotated tweets from TWNEWS with a balanced tagging of the four above indicated sentiment labels.

Starting from the promising results for other languages (Strapparava et al., 2011; Davidov et al., 2011), we are setting up a framework where irony recognition in our tweets can be formulated as a classification task and machine learning algorithms can be applied.

Making use of a simple evaluation scheme for classification-based tasks called Confusion Matrix (Stehman, 1997), it is possible to look at the existing overlapping among the classes, i.e., how much one class is misclassified as another one. This mechanism usually gives some hint on the lexical overlapping between the texts of two different classes. In our case, we noticed a significant lap between humorous texts and negative ones, while the same does not happen when comparing humorous with positive texts. This somehow confirms what already discovered by (Mihalcea and Pulman, 2007). Another interesting point of analysis concerns the discriminatory power of the words within the classification procedure. This can be easily done by calculating the Information Gain (or Kullback-Leibler divergence (Kullback and Leibler, 1951)) of the terms with respect to the class labels. In case of comparisons between texts sharing both temporal and domain characteristics, it helps to discover current targets of humor. For instance, using our recent tweets talking about Italian politics, terms like ‘Monti’ and ‘Passera’ resulted to be highly relevant during classification (the first one refers to the current Italian prime minister Mario Monti, whereas the second is the Italian minister of economy and development Corrado Passera). Notice that both ‘monti’ and ‘passera’ are words of the Italian vocabulary (e.g. the word ‘monti’ means ‘mountains’, while ‘passera’ means ‘hen sparrow’ but it is also used in adult slang as masculinist metaphor), and many jokes in our tweets exploit such forms of ambiguity.

As a second result, this tool allows to individuate those recurrent patterns that are strictly related to the information sources. In our scenario, the token “http” usually indicates the presence of news instead of humorous texts. This is due to the shortness nature of Twitter that obliges the users to be concise. Indeed, most of non-humorous and informative tweets contain few words followed by one hyperlink (e.g. TWNEWS-186: ‘Chi è Mario Monti? <http://t.co/BZewchzZ>’ (Who is Mario Monti? <http://t.co/BZewchzZ>)).

Still, Information Gain can be used to mine those linguistic expressions, rather than single words, that can be useful to identify the humorous nature of the text. For example, meaningful terms that turn out to be important in this sense are “speriamo” (i.e., “we wish”) and “bene”(“good”),

which refer to the Italian expression “speriamo bene” (“fingers crossed”). Other highly-scored terms include “fiducia” (“trust”), “finalmente” (finally), and so forth. One next step in this direction would be to evaluate such discriminatory power with respect to each one of the classes.

In future works, we aim at using linguistic resources to preprocess the input texts in order to remove noise and uninformative terms. Then, the use of data morpho-syntactically annotated could be crucial in the identification of whole syntactic structures (e.g., “bank director”) as well as linguistic expressions. Finally, the time and the mood of verbs can be another way of studying linguistic differences between humorous and objective texts.

All the above points only represent some issues that came out from our first experiments, thus they are to be considered as preliminary results.

4. Conclusion and future work

In this paper we described our current work on Senti-TUT, a novel Italian corpus for sentiment analysis which includes sentiment annotations concerning irony and consists in a collection of texts from Twitter.

For what concerns issues arising in the manual annotation of the sentiment of our tweets, useful guidelines were found in (Wiebe et al., 2005), where a general annotation scheme to distinguish subjective information from material presented as fact is defined. Tweets in our corpus often express opinion about news entities while reporting on recent events (Godbole et al., 2007), or report opinions of news entities (e.g. politicians) about the breaking news. Following (Wiebe et al., 2005) in both cases we considered the tweets as subjective (with a positive or negative polarity).

Concerning the specific issue of determining if a tweet is ironic, this is not an easy task, mainly due to the fact that irony is very subjective and personal appreciation can lead to different perceptions. We mainly recognized the following features in our tweets: frequent use of adult slang and dirty words, use of echoic irony, language jokes, which often exploit ambiguities involving the politicians’ proper nouns, as confirmed by first experiments. Moreover, we observed many cases of quotation or explicit reference to popular, Italian or international, television series, see e.g. the following tweet referring to the American reality television series Jersey Shore: TWNEWS-844 ‘@mtvitaly ma è vero che Mario Monti parteciperà a Jersey Shore? <http://t.co/d0H1Kmp6>’ (@mtvitaly Is it true that Mario Monti will be a cast member of Jersey Shore? <http://t.co/d0H1Kmp6>). Therefore, a problem that needs to be taken into account is that sometimes in our context the recognition of irony can be hard, because strongly depends not only to the annotator knowledge about the Italian political situation but also to his/her degree of “addiction” to tv shows.

Since the perception of irony can vary from a subject to another, different annotators could consider a given post ironic or sarcastic “to some degree”. In order to face this issue, it would be useful to assign scores to ironic annotations, as suggested in (Davidov et al., 2011). Moreover, we are also considering to extend the annotation framework by adding a more fine-grained annotation where the entire

text is divided in pieces (or fragments) representing both the facts under discussion and the expressions about the judgement. In such richer setting, it will be possible to evaluate the system at different levels of granularity and to use the information to measure different degree of irony. Moreover, during the annotation work, we have observed many different typologies of ironic statements, as for instance *sarcastic tweets*, conveying bitter or cutting expressions or remarks, *hilarious or facetious tweets*, aimed at producing a comic effect, *language jokes*, and so on. In order to tackle this issue, as future work we aim at studying a more sophisticated classification of ironic tweets, where different ways of expressing irony can be distinguished (and possibly organized in a taxonomy) and tweets can be annotated accordingly. In this framework it will be also interesting to test the results of enabling multi-value-annotations.

5. Acknowledgements

The work reported in this paper was partly carried out in the framework of the project PARLI ("Portal for the Access to the Linguistic Resources for Italian") benefitting from a research funding PRIN-2008 from the Ministry of Public Instruction and University.

The authors thank all the persons who supported the work. We are grateful to our annotators, in particular to Gianna Cernuschi and Andrea Marchetti and to CELI Torino for providing the facilities offered by the Blogmeter social media monitoring platform.

6. References

- Matteo Baldoni, Cristina Baroglio, Viviana Patti, and Paolo Rena. 2012. From Tags to Emotions: Ontology-driven Sentiment Analysis in the Social Semantic Web. *Intelligenza Artificiale: the International Journal of the AI*IA*. In press.
- Carmen Banea, Rada Mihalcea, and Janyce Wiebe. 2010. Multilingual subjectivity: are more languages better? In *Proceedings of the 23rd International Conference on Computational Linguistics, COLING '10*, pages 28–36, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Ester Boldrini, Alexandra Balahur, Patricio Martínez-Barco, and Andrés Montoyo. 2010. Emotiblog: a finer-grained and more precise learning of subjectivity expression models. In *Proceedings of the Fourth Linguistic Annotation Workshop, LAW IV '10*, pages 1–10, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Johan Bollen, Huina Mao, and Xiao-Jun Zeng. 2010. Twitter mood predicts the stock market. *CoRR*, abs/1010.3003.
- Cristina Bosco and Alessandro Mazzei. 2012a. The evalita 2011 parsing task: the constituency track. In *Evalita'11 Working Notes*, Roma.
- Cristina Bosco and Alessandro Mazzei. 2012b. The evalita 2011 parsing task: the dependency track. In *Evalita'11 Working Notes*, Roma.
- Cristina Bosco, Vincenzo Lombardo, Leonardo Lesmo, and Daniela Vassallo. 2000. Building a treebank for Italian: a data-driven annotation schema. In *Proceedings of LREC'00*.
- CELI and Me-Source. 2009. Blogmeter. <http://www.blogmeter.eu/>.
- Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2011. Semi-supervised recognition of sarcastic sentences in twitter and amazon. In *Proceedings of the fourteenth conference on computational natural language learning*, pages 107–116.
- Andrea Esuli, Fabrizio Sebastiani, and Ilaria Urciuoli. 2008. Annotating expressions of opinion and emotion in the italian content annotation bank. In *Proceedings of the 6th Conference on Language Resources and Evaluation (LREC 2008)*.
- Daniel Gildea. 2001. Corpus variation and parser performance. In *Proceedings of the Empirical Methods in NLP Conference*, Pittsburg.
- Namrata Godbole, Manjunath Srinivasaiah, and Steven Skiena. 2007. Large-scale sentiment analysis for news and blogs. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.
- H. Paul Grice. 1975. Logic and conversation. In Peter Cole and Jerry Morgan, editors, *Syntax and Semantics 3: Speech Acts*, pages 64–75. Academic Press, New York.
- Solomon Kullback and Richard A. Leibler. 1951. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86.
- Leonardo Lesmo. 2007. The rule-based parser of the NLP group of the University of Torino. *Intelligenza artificiale*, 2:46–47.
- Leonardo Lesmo. 2009. The Turin University Parser at Evalita 2009. In *Proceedings of Evalita'09*.
- Rada Mihalcea and Stephen Pulman. 2007. Characterizing humour: An exploration of features in humorous texts. In *Proceedings of the 8th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing '07*, pages 337–347, Berlin, Heidelberg. Springer-Verlag.
- Bo Pang and Lillian Lee. 2008. *Opinion Mining and Sentiment Analysis (Foundations and Trends(R) in Information Retrieval)*. Now Publishers Inc.
- Dan Sperber and Deirdre Wilson. 1986. *Relevance: communication and cognition*. Harvard University Press, Cambridge, MA, USA.
- Stephen V. Stehman. 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1):77 – 89.
- Carlo Strapparava, Oliviero Stock, and Rada Mihalcea. 2011. Computational humour. In Roddy Cowie, Catherine Pelachaud, and Paolo Petta, editors, *Emotion-Oriented Systems*, Cognitive Technologies, pages 609–634. Springer Berlin Heidelberg.
- Oren Tsur, Dmitry Davidov, and Ari Rappoport. 2010. Icwsm - a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In William W. Cohen and Samuel Gosling, editors, *In Proceedings of the Fourth International Conference on Weblogs and Social Media (ICWSM-2010)*. The AAAI Press.

- Andranik Tumasjan, Timm O. Sprenger, Philipp G. Sandner, and Isabell M. Welpe. 2011. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pages 178–185. AAAI.
- Sudha Verma, Sarah Vieweg, William J. Corvey, Leysia Palen, James H. Martin, Martha Palmer, Aaron Schram, and Kenneth M. Anderson. 2011. Natural language processing to the rescue? extracting situational awareness tweets during mass emergency. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pages 385–392. AAAI.
- Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2-3):165–210.