Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), pages 5098–5105
Marseille, 11–16 May 2020

Marking Irony Activators in a Universal Dependencies Treebank: The Case of an Italian Twitter Corpus

Alessandra Teresa Cignarella[1,2], **Manuela Sanguinetti**[1], **Cristina Bosco**[1], **Paolo Rosso**[2]
1. Dipartimento di Informatica, Università degli Studi di Torino, Italy
2. Universitat Politècnica de València, Spain
{cigna|msanguin|bosco}@di.unito.it, prosso@disc.upv.es

Abstract

The recognition of irony is a challenging task in the domain of Sentiment Analysis, and the availability of annotated corpora may be crucial for its automatic processing. In this paper we describe a fine-grained annotation scheme centered on irony, in which we highlight the tokens that are responsible for its activation (*irony activators*), and their morpho-syntactic features. As our case study we therefore introduce a recently released Universal Dependencies treebank for Italian which includes ironic tweets: TWITTIRÒ-UD. For the purposes of this study, we enriched the existing annotation in the treebank with a further level that includes irony activators. A description and discussion of the annotation scheme is provided with a definition of irony activators and the guidelines for their annotation. This qualitative study on the different layers of annotation applied on the same dataset can shed some light on the process of human annotation, and irony annotation in particular, and on the usefulness of this representation for developing computational models of irony to be used for training purposes.

Keywords: Twitter, Italian, Irony, Irony Activators, Syntax, Universal Dependencies

1. Introduction

Irony detection proves to be an important task in the domain of Sentiment Analysis, due to the fact that the presence of ironic devices in a text may reverse the polarity of an opinion, which can be expressed, for example, with positive words but intending a negative meaning, therefore undermining the performance of Sentiment Analysis systems (Bosco et al., 2013; Reyes et al., 2013; Barbieri et al., 2014; Ghosh et al., 2015; Hernández Farías et al., 2015; Joshi et al., 2015).

As a matter of fact, the recognition of irony and the identification of pragmatic and linguistic devices that activate it have always been a controversial topic (Grice, 1975; Grice, 1978; Sperber and Wilson, 1981), and a challenging task for both human annotators and automatic tools.

Additionally, the automatic treatment of such phenomenon is further complicated by the co-occurrence of similar forms of speech such as sarcasm or satire (Hernández Farías and Rosso, 2016; Joshi et al., 2017; Ravi and Ravi, 2017; Zhang et al., 2019) and by the text domain. The application of tools for irony detection has been especially focused on micro-blogging platforms (such as Twitter), since irony is a pragmatic phenomenon widely exploited by users online.

In the last five years in particular, there has been a stronger attention from the NLP community to the topic of irony detection, and this is also attested by the proposal of various shared tasks for different languages and the growing participation of several research groups and industries. As a side effect, novel annotated resources were released where irony is annotated.

The main goal of this paper is to address the issues related to the annotation of a dataset for irony detection by presenting an annotation scheme that allows a fine-grained representation of the phenomenon. A novel annotation scheme is therefore provided, which is centered on the notion of *irony activator* and the morphological and syntactic features that can be observed in its realization in the context of an ironic message.

The Universal Dependencies (UD) format is used to encode the morpho-syntactic features of the irony activators, and a UD-based Italian treebank is exploited as a case study. After a brief overview of the shared tasks and the corpora developed for irony detection, in this paper we describe more in detail TWITTIRÒ-UD along with its annotation layers, focusing in particular on the one that includes irony activators.

2. Background and Motivation

The present work is part of a wider joint project regarding irony detection. Together with other research groups working on English and French (Karoui et al., 2015), in the past years we have developed a fine-grained annotation devoted to the implementation of models for irony and irony-aware systems in a multilingual perspective. The main aim of this project is the investigation of the suitability and usefulness of a fine-grained annotation for the development of computational models of irony and figurative language devices.

As far as irony detection is concerned, several evaluation exercises were organized for different languages in the last few years. The Italian EVALITA campaign included a first pilot task on irony detection in 2014, within the SENTIPOLC shared task (Basile et al., 2014), which was proposed again in 2016 (Barbieri et al., 2016), while in 2018 a fine-grained shared task was proposed on the classification of irony and sarcasm: i.e. IronITA[1] (Cignarella et al., 2018b).

Similarly, in the framework of periodical evaluation campaigns for NLP tools, a related task was proposed for French: i.e. DEFT[2] at TALN 2017 (Benamara et al., 2017). For what concerns Spanish, the most recent shared task about irony in social media was organized at IberLEF 2019 *Irony Detection in Spanish Variants (IroSvA 2019)*[3], explor-

[1] http://di.unito.it/ironita18
[2] https://deft.limsi.fr/2017/
[3] https://www.autoritas.net/IroSvA2019/

ing the differences among varieties of Spanish from Spain, Cuba and Mexico (Ortega et al., 2019), in which the organizers also proposed a focus on context, stressing the importance of contextual semantics in ironic productions.
In the same year, the task *Irony Detection in Arabic Tweets* (IDAT 2019)[4] was organized within the evaluation campaign FIRE 2019. Its aim was to explore, for the first time, the realization of irony in a non-European language (Ghanem et al., 2019). While concerning English, the first task about this topic is the Task 11 proposed at SemEval-2015, focusing on *Sentiment Analysis of Figurative Language in Twitter*[5] (Ghosh et al., 2015). In 2018 a shared task specifically on irony detection in tweets has been organized within the context of the same evaluation campaign: *SemEval-2018 Task 3: Irony detection in English tweets*[6] (Van Hee et al., 2018). In this latter evaluation exercise the aim was not only to detect irony but also to classify it into three different sub-types: verbal irony by means of a polarity contrast, situational irony and other kind. The setting proposed for the SemEval-2018 is an indication of the growing interest in a deeper analysis of the linguistic phenomena underlying ironic expressions.
As far as resources are concerned, in most of the cases the corpora exploited in the tasks mentioned above are featured by a simple indication of the presence/absence of irony, but the interest for annotation going beyond this point is attested e.g. by (Cignarella et al., 2018a), which describes an Italian corpus of tweets (TWITTIRÒ), annotated with a fine-grained tagset for irony that was presented in Karoui et al. (2015). The same resource has been recently released under the *Universal Dependencies* project after the application of the UD morpho-syntactic format, based on dependency syntax (Nivre et al., 2016), thus creating TWITTIRÒ-UD (Cignarella et al., 2019a).
The starting point for this paper is TWITTIRÒ-UD. Recently, we proposed for this corpus to add a new level of semantic information by explicitly tagging irony activators at token level (Cignarella et al., 2019b) and the present research is a direct follow-up of this last work. Here the annotation of the activators has been applied on the full dataset, consisting of 1,424 tweets, and analyzed according to the perspectives described in the following sections.
We present the annotation process that underlies the enrichment of the TWITTIRÒ-UD corpus that, to the best of our knowledge, is one of the few linguistic resources where sentiment analysis and syntactic annotation are applied within the same framework.
Following the trail of increasing interest for the automatic detection of irony, reflected by the above-mentioned shared tasks in different languages, our aim is to develop a gold-standard resource that could be exploited in the training of automatic systems. Another goal is that of shedding some light on the pragmatic devices that are used to express irony in a finer-grained fashion, in particular creating a bridging interface between the level of semantics and the level of syntax.

[4] https://www.irit.fr/IDAT2019/.
[5] http://alt.qcri.org/semeval2015/task11/
[6] https://competitions.codalab.org/competitions/17468

3. Corpus Description

As mentioned in the previous section, the TWITTIRÒ-UD corpus comprises of many different levels of annotation. In this section we will summarize the size of the dataset and its characteristics. We will describe the two different annotation scopes: the one that includes irony and the one that involves syntactic annotation.

3.1. Irony

The dataset is composed of 1,424 ironic tweets written in Italian and labeled according to different annotation levels:

- the activation type
- the irony category

Two different *types* of irony activation can be distinguished: IMPLICIT and EXPLICIT. The distinction lies on whether the functional elements that activate irony are both lexicalized in the tweet (explicit) or only one of them is (implicit). We also distinguish among eight different *categories*[7], which indicate the modality in which irony is conveyed. Table 1 shows their distribution throughout the corpus.

	IMPLICIT	EXPLICIT	TOTAL
ANALOGY	55	206	261
EUPHEMISM	10	74	84
CONTEXT SHIFT	–	185	185
OXYMORON PARADOX	–	277	277
HYPERBOLE	7	74	81
FALSE ASSERTION	117	–	117
RHETORICAL QUESTION	19	202	221
OTHER	57	141	198
TOTAL			**1,424**

Table 1: Irony types and irony categories in TWITTIRÒ-UD. Cells are with hyphen sign when they represent a category either inherently implicit or explicit.

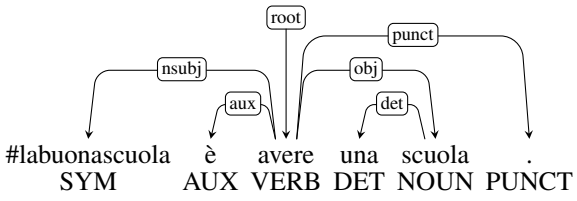
3.2. Syntax

In order to add the syntactic annotation layer to the dataset, we applied the full pipeline of tokenization, lemmatization, PoS-tagging and dependency parsing provided by *UDPipe* (Straka and Straková, 2017), obtaining a representation in CoNLL-U format of the 1,424 tweets[8].

```
# sent_id = _____
# twittiro =  EXPLICIT  OXYMORON / PARADOX
# text = #labuonascuola è avere una scuola.[9]
```

The parser output was also manually revised by two independent annotators. Below, we provide an example of annotated tree:

[7] For a detailed description of the categories and their profound pragmatic meaning, please refer to (Cignarella et al., 2018a).
[8] The TWITTIRÒ-UD treebank is freely available at: https://github.com/UniversalDependencies/UD_Italian-TWITTIRO.
[9] Translation: *#thegoodschool is to have a school.*



4. Annotation of the Irony Activators

Exploiting the availability of the above mentioned scopes of annotation, irony on one side and morpho-syntactic structures and features on the other, we intend to perform a kind of annotation that serves as an interface between these two different levels of analysis; i.e. annotating specific irony activators at a token level and investigating the syntactic relations between them.

Provided that irony is activated by the presence of a clash or a contradiction between two elements (both occurring in the message or one within and one without it), we can indicate them as triggers of irony, i.e. **irony activators**, that correspond to some node of the syntactic tree to be marked as T1 and T2 (when it occurs within the message).

After a pilot annotation experiment presented in (Cignarella et al., 2019b), we compiled annotation guidelines that took into account the nature of the eight irony categories.

4.1. Exploiting Syntactic Information

The creation of such a resource, annotated both for sentiment analysis and within the syntactic framework of Universal Dependencies, aims at exploring the possible inter-relation between semantics and syntax in ironical utterances. The main intuition that inspired this kind of work is that we assume that there exist "syntactic patterns" which activate irony, or that exploiting features based on dependency syntax in an automatic system would be useful in the detection of irony. Indeed, a preliminary attempt was made in (Cignarella and Bosco, 2019), where shallow features based on dependency relations have been engineered in order to solve a task concerning automatic irony detection.

In the syntax tree of the tweet "Spero sia colite. Ma ho paura sia amore." (*I hope it's colitis. But I'm afraid it's love.*), we have annotated T1 and T2 according to the guidelines of the category ANALOGY (see Section 4.2.). The syntactic tree looks like this:

```
# twittiro =  EXPLICIT  ANALOGY
# activators =  3  9
```

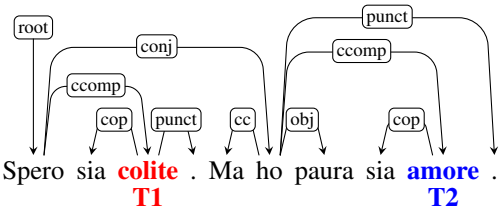


Figure 1: Syntactic tree in UD flat format, with highlighted irony activators T1 and T2.

According to our intuition, if "syntactic patterns" that help to detect irony do exist, they should be particularly evident in the syntactic context of certain lexical elements that create a semantic clash in a text (i.e. T1 and T2). After extracting automatically the "sub-tree" surrounding the irony activators, we would have this tree representation:

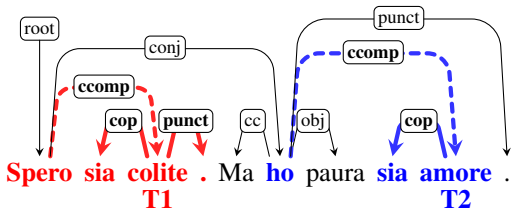


Figure 2: Syntactic tree in UD flat format, with highlighted irony activators T1 and T2 and their respective sub-trees.

The information in which we are interested in, and that motivated us in creating such resource are chiefly two:

1. What are the tokens directly connected through dependency relationships to the irony activators?

2. Is there a syntactic relation that connects T1 and T2?

Concerning the tweet above, the tokens directly connected through dependency relationships to the first question are: $T1 = [spero, sia, colite, .]$ and $T2 = [ho, sia, amore]$.

While, from a dependency relation viewpoint T1 and T2 are connected by means of:
$T1 \rightarrow T2 = [T1 \rightarrow ccomp \rightarrow conj \rightarrow ccomp \rightarrow T2]$.

Such information could be exploited as feature in the implementation of automatic systems for the detection of irony, but it could also be useful to gain new insights on patterns that may underlie the activation of irony.

4.2. Annotation Guidelines

In social media such as Twitter, contrasts in ironic tweets often consist of at least two propositions (but also simple words), which are in contradiction to each other (Karoui et al., 2017). This contradiction can be at a verbal or situational level. As we described in Section 3.1., the two elements in contrasts, i.e. T1 and T2, can be both part of the internal context of an utterance (that is explicitly lexicalized), or one is present and the other one implied.

Starting from this principle – and exploiting the fact that our data are tokenized as per the Universal Dependencies format – we annotate T1 and T2 at a token level such that:

- T1 and T2 can be tokens of any type: no specific constraints are given on the morpho-syntactic category; whenever a compound or multi-word named entity is involved, just its syntactic head is conventionally annotated;

- the identification of the proper T1 and T2 is guided by the irony category.

The decision of annotating exclusively a single token corresponding to each irony activator found in the tweet lies in a twofold motivation. Firstly, managing only a pair of tokens per tweet is computationally easier and more homogeneous. Secondly, each node of a dependency tree can be seen as a head of a sub-tree, i.e. a word usually semantically richer with respect to its dependents and therefore more interesting for our analysis. Furthermore, in this way we are also able to analyze morphological similarities between irony activators and across irony types and categories, as we will describe in Section 5.

Here we provide the annotation guidelines specific for each category of irony, along with an annotated tweet as example.[10]

Analogy - Find the elements that are put in comparison, (either literally or metaphorically).

(1) cmq [Mario]$_{T1}$ Monti è come [Carlo]$_{T2}$ d'Inghilterra. poverino, non sarà mai re
[Mario]$_{T1}$ Monti is like [Prince]$_{T2}$ Charles. poor guy, he will never be king

Euphemism - Select T1 and T2 so that one of the two elements is supposedly used to soften or just substitute the other one in the pair.

(2) Salento, extracomunitario [muore]$_{T1}$ mentre raccoglie pomodori. Era l'unico modo per prendersi una [pausa]$_{T2}$.
Salento, immigrant [dies]$_{T1}$ while harvesting tomatoes. It was his only way to take a [break]$_{T2}$

Context shift - Select the words that most strikingly show the change of topic/frame in the tweet, but also a change of register, or the style of the conversation.

(3) Bossi: «Il governo Monti fa [schifo]$_{T1}$». Un giudizio [tecnico]$_{T2}$.
Bossi: «The Monti administration [sucks]$_{T1}$». A [technical]$_{T2}$ judgement.

Oxymoron/Paradox - Select the activators so that the type of relation triggered will be a contrast or a contradiction.

(4) la cosa bella del governo Monti è che ha [acceso]$_{T1}$ le speranze di tutti..... e le [spegnerà]$_{T2}$ pure...
the good thing about the Monti government is that it has [kindled]$_{T1}$ everyone's hopes..... and it will [stifle]$_{T2}$ them as well...

Hyperbole - Select the activator that expresses the exaggeration; in case of explicit hyperbole, both the hyperbolic expression and the emphasized element can be annotated.

(5) [Mario]$_{T1}$ Monti il [messia]$_{T2}$, accolto in Senato da un lungo applauso
[Mario]$_{T1}$ Monti the [messiah]$_{T2}$, welcomed in Senate with a warm applause

False assertion - This category is always implicit. Mark only one activator, selecting the syntactic head of the assertion or reported speech that it is believed to be false.

(6) #labuonascuola è scuola per imparare a [fare]$_{T1}$ dolci
#thegoodschool is the school to learn how to [bake]$_{T1}$

Rhetorical question - Select T1 and T2 such that the first element is the syntactic head of the rhetorical question present in the tweet, and the second (if present) an element in contrast with it.

(7) Mario Monti non [vuoi]$_{T1}$ una nipotina? [#ADOTTAMI]$_{T2}$
Mario Monti don't you [want]$_{T1}$ a little niece? [#ADOPTME]$_{T2}$

Other - No specific guidelines are provided for this category, due to its heterogeneous nature that may comprise different linguistic and pragmatic phenomena.

(8) Ma 3 [Mario]$_{T1}$ Monti fanno un Giulio [Tremonti]$_{T2}$?
Do 3 [Mario]$_{T1}$ Monti make a Giulio [Tremonti]$_{T2}$?

In Example 8, for example, irony is activated by the presence of a wordplay. In fact, "3 Monti", that means "three times the president Mario Monti" (English: *three mounts*) is a homophone of the surname of Giulio Tremonti, another Italian politician.

4.3. Procedure

In the previous step of this work (Cignarella et al., 2019b), two skilled annotators first carried out a pilot annotation experiment on small sample of 50 tweets of the category OXYMORON/PARADOX; after a discussion of the annotation results, they drafted a first version of the guidelines and then proceeded with the annotation of all the tweets labeled as OXYMORON/PARADOX, that correspond to about 20% of the corpus (see Table 1). An expanded version of the guidelines was then produced (see Section 4.2.) and the annotation process was carried out for all the remaining categories.
Here we comment on the observed agreement on the 1,424 instances between the two skilled annotators, also introducing some examples.

4.4. Agreement

As mentioned in the previous sections, two annotators (Ann1 and Ann2) had to mark two irony activators at a token level (or just one in case of implicit irony and leave an underscore as placeholder). With the kind of data we are dealing with, it was not feasible to use standard measures as Cohen's κ or Krippendorff's α to calculate inter-annotator agreement (IAA). For this reason we calculated

[10]The guidelines discussed in this paper for the annotation of T1 and T2 at a token level are inspired to a great extent by this document:
`https://github.com/IronyAndTweets/Scheme`.

the observed agreement (expressed in percentage) between Ann1 and Ann2 on the 1,424 tweets distinguishing three scenarios:

- Agreement:
  the two annotators agree on both tokens.
  $T1_{Ann1} = T1_{Ann2} \wedge T2_{Ann1} = T2_{Ann2}$
- Partial Agreement:
  the two annotators agree on only one token.
  $T1_{Ann1} = (T1_{Ann2} \vee T2_{Ann2})$
  $\vee$
  $T2_{Ann1} = (T1_{Ann2} \vee T2_{Ann2})$
- Disagreement:
  the two annotators disagree on both tokens.
  $T1_{Ann1} \neq (T1_{Ann2} \vee T2_{Ann2})$
  $\wedge$
  $T2_{Ann1} \neq (T1_{Ann2} \vee T2_{Ann2})$

In Figure 3 it can be seen how a complete agreement was reached on 485 tweets (34.1%), while a partial agreement was reached on 613 tweets (43%). The remaining tweets (i.e. 326), which correspond to a 22.9% of the corpus have a clear cut disagreement.

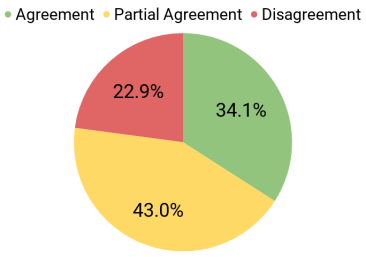


Figure 3: Observed IAA on the annotation of T1 and T2.

We also propose a more detailed representation of the agreement, that takes into account the distinction between irony types (implicit *vs* explicit). As it can be observed in Figure 4, it seems that the cases with implicit type of contradiction are easier to detect, and in general they show less disagreement. This low percentage of disagreement on the implicit class (6,77%) is also naturally attributable to the nature of the annotation guidelines. In fact, in case of implicit type of irony, only one token is marked, while the second is left blank.
Another point worth mentioning is also that, in many of the cases of partial agreement the two annotators tagged as T1 the same token, while they marked as T2 tokens that belong to the same syntactic sub-tree (but not exactly the same token). As in the example below:

(9) RT @user: sono riuscita a cambiare il profilo, mo nevica.
RT @user: I managed to change my profile, now it'll snow.

Ann1: [riuscita]$_{T1}$, [nevica]$_{T2}$
Ann2: [cambiare]$_{T1}$, [nevica]$_{T2}$

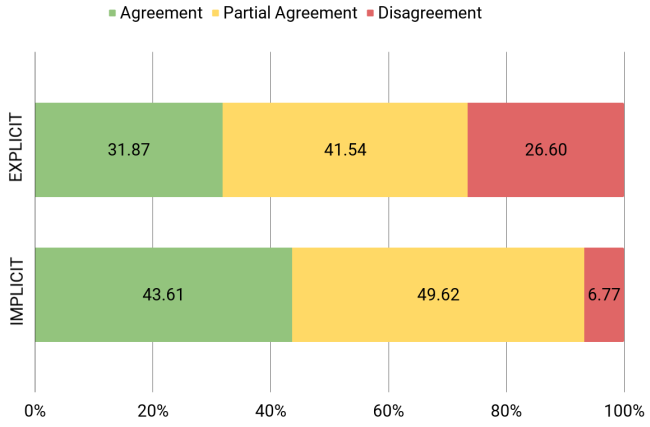


Figure 4: Observed IAA with respect to irony type.

In Example 9, [riuscita] and [cambiare] belong to the same syntactic sub-tree. Nonetheless, Ann1 opted for marking [riuscita] as T1, while Ann2 opted for the latter solution. This kind of disagreement issues are not considered as strikingly hard to solve, and they were easily resolved with the intervention of a third skilled annotator, as we describe in Section 4.5.
Similarly, Figure 5 shows the observed agreement between the two annotators, according to the irony categories. As expected, the categories ANALOGY and OXYMORON/PARADOX, which are those in which a stronger correlation between syntax and semantics is present, should have been the ones with the highest agreement. In fact, in analogies, the comparison between two entities is often made explicit by a conjunction such as "come" (*English: like*), and it is easier to be detected, as in the example below:

(10) Il [governo]$_{T1}$ Monti è **come** il [medico]$_{T2}$ al capezzale del malato.
*Monti's [government]$_{T1}$ is **like** the [doctor]$_{T2}$ at the sick person's bedside.*

Similarly, irony by means of oxymoron and paradoxes is activated through the semantic contrast between two elements. These intuitions are confirmed by the data: in fact, analogies reach 45.98% of complete agreement and oxymorons 40.07%, both categories being among the top three agreement scores.
Surprisingly, a higher than expected agreement is also obtained for the category of FALSE ASSERTION (49.57%). Firstly, this category is inherently implicit, due to the restrictions posed by the annotation scheme (Karoui et al., 2015); meaning that only one token was annotated, thus reducing the probabilities of disagreement. Secondly, as described in section 4.2., the annotators had to "select the syntactic head of the assertion or reported speech [they] believed it is false". Due to this factor, and also because tweets in our collection are not longer than 140 characters[11], most of the time only one false assertion is present

[11]The tweets in our collection were retrieved before Twitter ex-

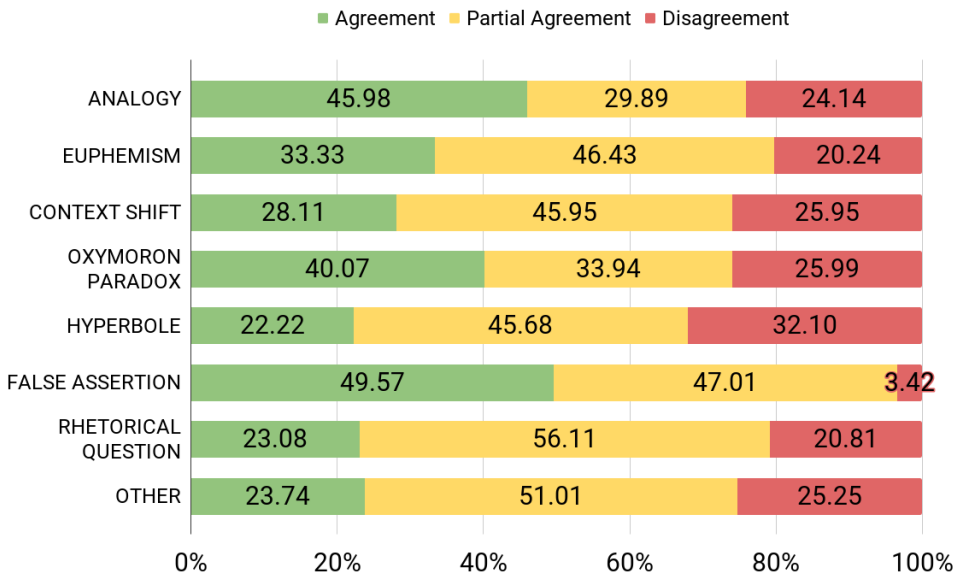


Figure 5: Observed IAA with respect to irony category.

in the text. Therefore, the syntactic head of the only false assertion present (or false reported speech) has been easily marked, thus lowering the possibility of disagreement.

On the other hand, the highest disagreement score is reached with the category of HYPERBOLE. A reason to explain this fact can be found by looking at the example below:

(11) POI DICONO CHE I MATTI SONO AL MANICOMIO..... NO AL GOVERNO MONTI... E' PERICOLOSISSIMO .....!
AND THEY SAY THAT CRAZY PEOPLE ARE IN THE MADHOUSE ..... NO [THEY ARE] IN MONTI'S GOVERNMENT... IT'S DANGEROUS .....!

In this tweet, an exaggerated expression is used (crazy people are in the government), but irony here is not activated by a single word or a word-pair, and therefore, it is not easy to mark it at a token level. In fact, in Example 11, what is also crucial for the activation of irony is the extensive use of punctuation and the exploitation of capital letters to express exaggeration.

4.5. Gold Standard

As described in the previous Section, Ann1 and Ann2 immediately reached a complete agreement on 485 tweets. The partial agreement was reached on 613 tweets and the remaining 326 tweet had a clear cut disagreement or were lacking annotation. In a second phase, a third skilled annotator (Ann3) solved all the disagreement cases between Ann1 and Ann2 choosing between the options that had been already highlighted, or leaning towards a third solution upon which all three annotators agreed unanimously in a joint session. After this procedure a gold standard of 1,424 tweets has been created and has been publicly released.[12]

5. Preliminary Analysis

In order to go in the direction of a deeper and finer-grained analysis on the relationship between semantics and syntax in the case of ironic expressions, we performed a preliminary analysis on the new gold standard. In particular, we wanted to investigate whether there exist some co-occurrencies, related to PoS tags and dependency relations, that are more frequent than other.

In Table 2, we show for each category, for both its implicit or explicit type, which are the preferred PoS tags that correspond to the irony activators. It is worth noticing that in the case of implicit realization of ironic utterances, the preferred trigger (T1) is a verb (in the case of EUPHEMISM, FALSE ASSERTION, RHETORICAL QUESTION and OTHER). The category RHETORICAL QUESTION seems to prefer irony activators that are verbs both in implicit and explicit realizations.

Additionally, we performed a count of the dependency relations linking T1 and T2 (if existing), discriminating for each different category. For each irony category we report the two most exploited dependency relations. The results are shown in Table 3. As it can be observed all eight categories have a preference for the `parataxis` dependency relation[13]. The ANALOGY category exploits mostly the `conj` dependency relation[14]. Interestingly T1 and T2 in the EUPHEMISM category have a preference for

panded the character limit to 280.

[12]The dataset is available at: `https://github.com/AleT-Cig/TWITTIRO-UD/`.

[13]Parataxis: `https://universaldependencies.org/u/dep/parataxis.html`.

[14]Conjunction: `https://universaldependencies.org/u/dep/conj.html`.

type	category	T1	T2
EXPLICIT	ANALOGY	NOUN	NOUN
IMPLICIT		NOUN	—
EXPLICIT	EUPHEMISM	NOUN	NOUN
IMPLICIT		VERB	—
EXPLICIT	CONTEXT SHIFT	NOUN	NOUN
EXPLICIT	OXYMORON / PARADOX	NOUN	NOUN
EXPLICIT	HYPERBOLE	NOUN	VERB
IMPLICIT		NOUN	—
IMPLICIT	FALSE ASSERTION	VERB	—
IMPLICIT	RHETORICAL QUESTION	VERB	—
EXPLICIT		VERB	VERB
IMPLICIT	OTHER	VERB	–
EXPLICIT		NOUN	NOUN

Table 2: Preferential PoS Tags of T1 and T2 in each irony type and category.

the `advcl` dependency relation[15]. The link between T1 and T2 in CONTEXT SHIFT is to be found through the `parataxis:appos` dependency relation[16], a sub-group of the `parataxis` deprel.

category	dependency relation
ANALOGY	`parataxis, conj`
EUPHEMISM	`parataxis, advcl`
CONTEXT SHIFT	`parataxis, parataxis:appos`
OXYM./PARADOX	`parataxis, conj`
HYPERBOLE	`parataxis, conj`
FALSE ASSERTION	—
RHETORICAL QUEST.	`parataxis, conj`
OTHER	`parataxis, conj`

Table 3: Preferential dependency relations occurring between T1 and T2 in each ironic category.

We can not display data related to the category FALSE ASSERTION, because due to its inherently implicit nature, only one trigger was annotated, therefore we could not explicitly extract any link between T1 and T2. Specific correlations between the categories of irony and the syntactic connection between irony activators T1 and T2 do not emerge from this preliminary analysis. As Table 3 shows the most present dependency relation across all eight irony categories is indeed `parataxis`. For this reason it is worth investigating more and in a finer-grained manner on the nature of the relationship that underlies between triggers on a morphological level, but also taking into account a semantic approach.

6. Conclusion

We have described a fine-grained annotation scheme concerning the pragmatic phenomenon of irony, in which we marked irony activators, and their morpho-syntactic features. We provide a description and discussion of the annotation scheme with the definition of irony activators and the guidelines for their annotation. This research, based on a qualitative study on the different layers of annotation applied on the treebank can shed some light on the process of human annotation, and irony annotation in particular, and on the usefulness of this representation for developing computational models of irony to be used for training purposes.

[15] Adverbial clause modifier: `https://universaldependencies.org/u/dep/advcl.html`.

[16] Paratactic apposition: `https://universaldependencies.org/it/dep/parataxis-appos.html`.

Acknowledgments

The work of C. Bosco and M. Sanguinetti was partially funded by Progetto di Ateneo/CSP 2016 (*Immigrants, Hate and Prejudice in Social Media*, S1618L2BOSC01). The work of P. Rosso was partially funded by the Spanish MICINN under the research project MISMIS-FAKEnHATE on MISinformation and MIScommunication in social media: FAKE news and HATE speech (PGC2018-096212-B-C31).

References

Barbieri, F., Saggion, H., and Ronzano, F. (2014). Modelling sarcasm in Twitter, a novel approach. In *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 50–58. ACL.

Barbieri, F., Basile, V., Croce, D., Nissim, M., Novielli, N., and Patti, V. (2016). Overview of the EVALITA 2016 SENTIment POLarity Classification Task. In *Proceedings of Third Italian Conference on Computational Linguistics (CLiC-it 2016) & Fifth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (EVALITA 2016)*, volume 1749. CEUR-WS.org.

Basile, V., Bolioli, A., Nissim, M., Patti, V., and Rosso, P. (2014). Overview of the EVALITA 2014 SENTIment POLarity Classification task. In *Proceedings of the 4th Evaluation Campaign of Natural Language Processing and Speech tools for Italian (EVALITA 2014)*. ELRA.

Benamara, F., Grouin, C., Karoui, J., Moriceau, V., and Robba, I. (2017). Analyse d'opinion et langage figuratif dans des tweets: présentation et résultats du Défi Fouille de Textes DEFT2017. In *Actes de l'atelier DEFT2017 associé à la conférence TALN*. Association pour le Traitement Automatique des Langues (ATALA).

Bosco, C., Patti, V., and Bolioli, A. (2013). Developing Corpora for Sentiment Analysis: The Case of Irony and Senti-TUT. *IEEE Intelligent Systems*, 28(2):55–63.

Cignarella, A. T. and Bosco, C. (2019). ATC at IroSvA 2019: Shallow Syntactic Dependency-based Features for Irony Detection in Spanish Variants. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2019), co-located with 34th Conference of the Spanish Society for Natural Language Processing (SEPLN 2019)*. CEUR-WS.org.

Cignarella, A. T., Bosco, C., Patti, V., and Lai, M. (2018a). Application and Analysis of a Multi-layered Scheme for Irony on the Italian Twitter Corpus TWITTIRÒ. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, pages 4204–4211. ELRA.

Cignarella, A. T., Frenda, S., Basile, V., Bosco, C., Patti, V., Rosso, P., et al. (2018b). Overview of the EVALITA 2018 task on Irony Detection in Italian Tweets (IronITA). In *Sixth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (EVALITA 2018)*, volume 2263, pages 1–6. CEUR-WS.org.

Cignarella, A. T., Bosco, C., and Rosso, P. (2019a). Presenting TWITTIRÒ-UD: An Italian Twitter Treebank in Universal Dependencies. In *Proceedings of SyntaxFest 2019*.

Cignarella, A. T., Sanguinetti, M., Bosco, C., and Rosso, P. (2019b). Is This an Effective Way to Annotate Irony Activators? In *Proceedings of Sixth Italian Conference on Computational Linguistics (CLiC-it 2019)*. CEUR-WS.org.

Ghanem, B., Karoui, J., Benamara, F., Moriceau, V., and Rosso, P. (2019). IDAT@FIRE2019: Overview of the Track on Irony Detection in Arabic Tweets. In *Notebook Papers of FIRE 2019*. CEUR-WS.org.

Ghosh, A., Li, G., Veale, T., Rosso, P., Shutova, E., Barnden, J., and Reyes, A. (2015). Semeval-2015 Task 11: Sentiment Analysis of Figurative Language in Twitter. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval-2015)*, pages 470–478. ACL.

Grice, P. H. (1975). Logic and Conversation. *Syntax and Semantics 3: Speech Arts*, pages 41–58.

Grice, P. H. (1978). Further Notes on Logic and Conversation. *Pragmatics*, 1:13–128.

Hernández Farías, D. I., Benedí, J.-M., and Rosso, P. (2015). Applying basic features from sentiment analysis for automatic irony detection. In *Pattern Recognition and Image Analysis*, volume 9117, pages 337–344. Springer.

Hernández Farías, D. I. and Rosso, P. (2016). Irony, Sarcasm, and Sentiment Analysis. In *Sentiment Analysis in Social Networks*, pages 113–128. Elsevier Science and Technology.

Joshi, A., Sharma, V., and Bhattacharyya, P. (2015). Harnessing context incongruity for sarcasm detection. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 757–762. ACL, July.

Joshi, A., Bhattacharyya, P., and Carman, M. J. (2017). Automatic sarcasm detection: A survey. *ACM Computing Surveys (CSUR)*, 50(5):73.

Karoui, J., Benamara, F., Moriceau, V., Aussenac-Gilles, N., and Belguith, L. H. (2015). Towards a contextual pragmatic model to detect irony in tweets. In *53rd Annual Meeting of the Association for Computational Linguistics (ACL 2015)*.

Karoui, J., Benamara, F., Moriceau, V., Patti, V., and Bosco, C. (2017). Exploring the Impact of Pragmatic Phenomena on Irony Detection in Tweets: A Multilingual Corpus Study. In *Proceedings of the European Chapter of the Association for Computational Linguistics*.

Nivre, J., De Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., McDonald, R. T., Petrov, S., Pyysalo, S., Silveira, N., et al. (2016). Universal Dependencies v1: A Multilingual Treebank Collection. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016)*.

Ortega, R., Rangel, F., Hernández, I., Rosso, P., Montes, M., and Medina, J. E. (2019). Overview of the Task on Irony Detection in Spanish Variants. In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2019), co-located with 34th Conference of the Spanish Society for Natural Language Processing (SEPLN 2019)*. CEUR-WS.org.

Ravi, K. and Ravi, V. (2017). A novel automatic satire and irony detection using ensembled feature selection and data mining. *Knowledge-Based Systems*, 120:15–33.

Reyes, A., Rosso, P., and Veale, T. (2013). A Multidimensional Approach for Detecting Irony in Twitter. *Language Resources and Evaluation*, 47(1):239–268.

Sperber, D. and Wilson, D. (1981). Irony and the Use-Mention Distinction. *Philosophy*, 3:143–184.

Straka, M. and Straková, J. (2017). Tokenizing, POS tagging, lemmatizing and parsing UD 2.0 with UDPipe. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 88–99.

Van Hee, C., Lefever, E., and Hoste, V. (2018). SemEval-2018 Task 3: Irony detection in English Tweets. In *In Proceedings of the International Workshop on Semantic Evaluation (SemEval-2018)*. ACL.

Zhang, S., Zhang, X., Chan, J., and Rosso, P. (2019). Irony detection via sentiment-based transfer learning. *Information Processing & Management*, 56(5):1633–1644.