Exploring the Realization of Irony in Twitter Data

Cynthia Van Hee, Els Lefever, Véronique Hoste

CLiN26
Amsterdam, december 18th, 2015.
Introduction

• Handling figurative language represents one of the most challenging tasks in natural language processing.
• However, figurative language is a ubiquitous part of spoken and written discourse (Veale & Hao, 2007) and is widespread in Web content (Reyes et al., 2013; Maynard & Greenwood, 2014).
Research Motivation

• Irony has important implications for tasks such as sentiment analysis (Reyes et al., 2013) where the aim is to automatically extract positive and negative opinions from online text.

  (1) It was so nice of my dad to come to my graduation party. #not

  (2) Going to the dentist for a root canal this afternoon. Yay, I cant wait.

• If we want to push the state-of-the-art in sentiment analysis or improve tasks such as cyberbullying detection, it is important to build computational models that are capable of recognizing irony.

---

1 Example taken from the SemEval-2015 T11 training corpus
2 Riloff et al., 2013
Related Research

• Traditional distinction between *situation*al and *verbal* irony
• Definition of verbal irony (Grice, 1975; Wilson & Sperber, 1992; Giora, 1995; Kumon-Nakamura et al., 1995; Attardo, 2000)
  → polarity-reversing mechanism
• Distinction between *irony* and *sarcasm*
  – ridicule (Kreuz & Glucksberg, 1989)
  – hostility (Clift, 1999)
  – intentionality (Haiman, 1998)
  – target presence (Kreuz & Roberts, 1993)
  ↓
  No formal agreement
Related Research

• This research focuses on verbal irony, which is defined as
  
  “an evaluative expression whose polarity is changed between the literal and the intended evaluation, resulting in an incongruence between the literal evaluation and its context.”
  
• Guidelines for annotating verbal irony in social media text (Van Hee et al., 2015)

3 Based on Burgers (2010)
Data Collection

• Tweets containing irony-related hashtags #irony, #sarcasm, #not

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tweets</td>
<td>3,000</td>
<td>3,179</td>
</tr>
</tbody>
</table>

Twitter streaming API

Twiqs (http://twiqs.nl)
Master’s thesis corpus
Data Annotation

• Guidelines for annotating irony (Van Hee et al., 2015)
• 3 main steps:
  1. Based on the definition, indicate for each text whether it is ironic, possibly ironic or not ironic.
  2. If the text is ironic:
     - Indicate whether an irony-related hashtag (#not, #sarcasm, #irony) is needed to perceive the irony.
     - Indicate the harshness of the irony on a two-point scale (0-1).
  3. Annotate all evaluative expressions contained by the text.
Data Annotation

1. Ironic / possibly ironic / not ironic
   - **Ironic:** instances containing a polarity change between the literal and the intended evaluation.
   - **Possibly ironic:** there is no polarity clash between the literal and the intended evaluation, but the tweet is ironic.
   - **Not ironic:** instances that don’t contain indications of ironic content or where there is insufficient context to perceive the irony.
1. Ironic / possibly ironic / not ironic
   - ironic:
     - Opposition: the text expresses an evaluation whose literal polarity is the opposite of the intended polarity.
       * e.g., Exams start tomorrow. Yay, can’t wait! #sarcasm
     - Hyperbole: the text expresses an evaluation whose literal polarity is stronger than the intended polarity.
       * e.g., Ohh he can climb a rope? Just like every other commando? Super talented. #sarcasm #TakeMeOut
     - Understatement: the text expresses an evaluation whose literal polarity is less strong than the intended polarity
       * e.g., A++ for Writing? Well that’s not extremely bad.
1. Ironic / possibly ironic / not ironic
   - **Ironic**: instances containing a polarity change between the literal and the intended evaluation.
   - **Possibly Ironic**: there is no polarity change between the literal and the intended evaluation, however, the text is ironic.

   *e.g.*, *The Chicago firehouse restaurant goes up in flames. #irony*
1. Ironic / possibly ironic / not ironic
   - **Ironic:** instances containing a polarity change between the literal and the intended evaluation.
   - **Possibly ironic:** there is no polarity change between the literal and the intended evaluation, but the text is ironic.
   - **Not ironic:** instances that don’t contain indications of ironic content or where there is insufficient context to perceive the irony.

   *e.g.,* #Iony is too complex a word for him to understand @BigJdog916
   *e.g.,* Chasing my dog through the house #not fun
2. If the text is ironic:
   ▪ Hashtag indication
     *e.g.*, *What a golden morning. #not*
     
     Dentists oh the joy #not
   ▪ Harshness
     *e.g.*, *A burned tongue is a lovely way to start the day #sarcasm #ouch*
     
     Shoutout to my mom for being hella supportive of me #sarcasm

3. Annotation of evaluative expressions
Data Annotation

- Brat online annotation tool (Stenetorp et al., 2012)

1. **Iro_oppos [1_high_confidence][High]**
   - Evaluation [Positive]
   - Shoutout to my mom for being hella supportive of me #sarcasm

1. **Iro_oppos [High]**
   - Mod [Intensifier]
   - **Evaluation [Positive]**
   - Mod [Intensifier]
   - Target [Negative]
   - Ah yes. Just what I wanted today - a pounding migraine. #Sarcasm

1. **Situational_Irony [High]**
   - The dude who told me money isn't everything is arguing with his son over money. In public. #Irony

1. **Non_Iro [High]**
   - Now i officially look single. Ha the #irony
Inter-annotator Study

- A subset of the corpus was annotated to assess whether the annotation scheme can be reliably applied to real-world data
- 3 annotators with a background in Linguistics

<table>
<thead>
<tr>
<th>Annotation</th>
<th>Irony (binary)</th>
<th>Hashtag indication needed</th>
<th>Polarity clash</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kappa</strong></td>
<td>0.77</td>
<td>0.84</td>
<td>0.60</td>
</tr>
<tr>
<td><strong>Polarity clash</strong></td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Annotation</th>
<th>Irony (binary)</th>
<th>Hashtag indication needed</th>
<th>Polarity clash</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kappa</strong></td>
<td>0.72</td>
<td>0.72</td>
<td>0.67</td>
</tr>
<tr>
<td><strong>Polarity clash</strong></td>
<td>0.64</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Averaged over 3 annotators
Corpus Statistics

Ironic tweets
1.945 370

Possibly ironic tweets
520
139

Ironic tweets
2.038 482
659

Ironic
Possibly ironic
Not ironic

Opposition
Hyperbole
Understatement

Exploring the Realization of Irony in Twitter Data
Corpus Statistics

Ironic tweets

- 780 harsh
- 1,258 not harsh

Ironic tweets

- 1,007 hashtag needed
- 1,031 no hashtag needed
Corpus Statistics

Ironic tweets
- Ironic: 1,445, 319
- Possibly ironic: 594
- Not ironic: 877

Possibly ironic tweets
- Other: 399
- Situational irony: 195

Ironic tweets
- Opposition: 6
- Hyperbole: 1
- Understatement: 1,522
Corpus Statistics

- Ironic tweets: 1,029 total, 500 harsh, 529 not harsh

- Ironic tweets: 775 with hashtag needed, 754 without hashtag needed
Future Work

• A deeper analysis of the dataset:
  – How is irony realized in text when there is no polarity clash perceived?
  – Is there any consistency in what is perceived as harsh and not harsh?
  – …

• Exploring the feasibility to automatically detect ironic utterances based on a polarity clash.
References


Remarks, questions, suggestions?

Cynthia.VanHee@Ugent.be