# Shared Emotions in Reading Pirandello. An Experiment with Sentiment Analysis

**Simone Rebora**
University of Verona
University of Basel
simone.rebora@univr.it

## Abstract

**English.** The paper reports on an experiment conducted with a group of students, aimed at verifying the effectiveness of Sentiment Analysis on Italian literary texts. Students were asked to annotate each paragraph of the short story "Ciàula scopre la luna" (1907) by Luigi Pirandello with a numeric evaluation of the sentiment and a free comment. Analysis of the annotations shows how, while inter-annotator agreement is still low, (a) emotional shifts in the story heighten the agreement in sentiment detection; (b) Sentiment Analysis works better for the comments than for the text, thus confirming its efficiency in reader response studies.

**Italiano.** L'articolo riporta un esperimento condotto con un gruppo di studenti, volto a verificare l'efficacia della Sentiment Analysis sui testi letterari in lingua italiana. Agli studenti è stato chiesto di annotare ogni paragrafo del racconto "Ciàula scopre la luna" (1907) di Luigi Pirandello con una valutazione numerica del Sentiment e un commento libero. L'analisi delle annotazioni mostra come, mentre l'Inter-Annotator Agreement resta basso, (a) le variazioni emotive nella storia aumentano l'accordo nella rilevazione del Sentiment; (b) la Sentiment Analysis funziona meglio per i commenti che per il testo, confermando così la sua efficienza negli studi sulla ricezione.

## 1 Introduction

Sentiment Analysis (SA) has recently grown in relevance in Digital Humanities. This computational technique, originally developed with the goal of analyzing "people's opinions, sentiments, appraisals, attitudes, and emotions towards entities and their attributes" (Liu, 2015), has found multiple applications in literary studies. From the widely discussed "shapes of stories" by Jockers (2014) and Reagan et al. (2016), to the study of fairy tales (Mohammad, 2012; Rotari, 2018), literary criticism (Rebora, 2017; Mellmann and Du, 2018), genre (Kim et al., 2017; Henny-Krahmer, 2018), and narrative structure (Zehe et al., 2016), SA has become one of the key methodologies in computational literary studies. For an extensive survey, see (Kim and Klinger, 2018). However, criticisms abound, both in theoretical (Ciotti, 2017) and practical (Sprugnoli et al., 2016) terms.
With this paper, I will report on an experiment aimed at verifying the efficiency of the approach in the study of two related phenomena: the narratological structure of a story and its associated reader response.

## 2 The Experiment: Bringing Research and Didactics Together

The experiment was conducted during the Digital Humanities course (*Informatica per gli studi umanistici*) held at the University of Verona in the academic year 2018/2019. Students were asked to read the short story (novella) "Ciàula scopre la luna" (1907) by Luigi Pirandello and were provided with an XML file with the following structure:

- the <novella> root tag;
- the child <frase>, containing one paragraph from the short story;

- the child <sentiment>, which the students were asked to fill with a numeric evaluation of the sentiment of the paragraph, ranging between -5 and +5;

- the child <commento>, where the students could write a free comment on the effects produced by reading the passage.

Full text of the novella was downloaded from LiberLiber and based on the 1986 Mondadori edition (Pirandello, 1986). Sentences were automatically split using the SA software Syuzhet, that was adopted as a groundwork for the entire experiment[1].
At the end of the annotation process, a total of 51 students wrote at least one comment or sentiment evaluation, for a total of 1,884 comments (36.94 per student) and 1,401 sentiment evaluations (27.47 per student). The 51 XML documents were then anonymized and merged into a single file, available for consultation (together with the R scripts for its analysis) on Github.
The experiment had the didactic purpose of letting students familiarize with the XML markup language and with SA computational techniques (both, in a very simplified form). In terms of research purposes, their annotations proved precious for a verification of the efficiency of SA approaches.

## 3 Analysis

### 3.1 Agreement on Sentiment Annotations

Figure 1 shows all sentiment annotations by the students. As evident, annotations are widely spread throughout the 111 paragraphs of "Ciàula", with a dominance of the central levels of emotionality and a deviation towards the most positive levels only at the end of the novella.

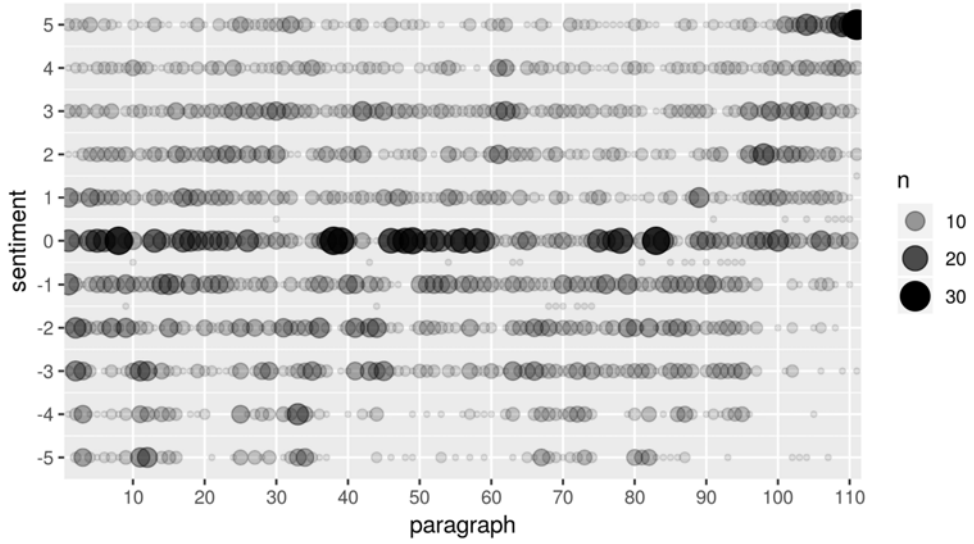


Figure 1: Sentiment annotations on "Ciàula scopre la Luna" (51 annotators: -5/+5)

For a more detailed understanding of the level of agreement, Krippendorff's Alpha (Krippendorff, 2018) was adopted. To maximize the possible agreement, annotations were reduced to a binary selection:

- annotation value < 0: "negative" tag;

- annotation value = 0: no tag;

- annotation value > 0: "positive" tag.

[1]Note that Syuzhet was designed to work on a sentence level (in fact, repeated words do not count towards the total sentiment). This is why annotation was performed on a sentence (and not paragraph) level.

However, Krippendorff's Alpha was still substantially low (0.19), thus confirming the result of Sprugnoli et al. (2016), who showed how inter-annotator agreement advises against the application of SA to historical (or literary) texts.
Given this acknowledgment, still, some interesting outcomes can be derived from the experiment.

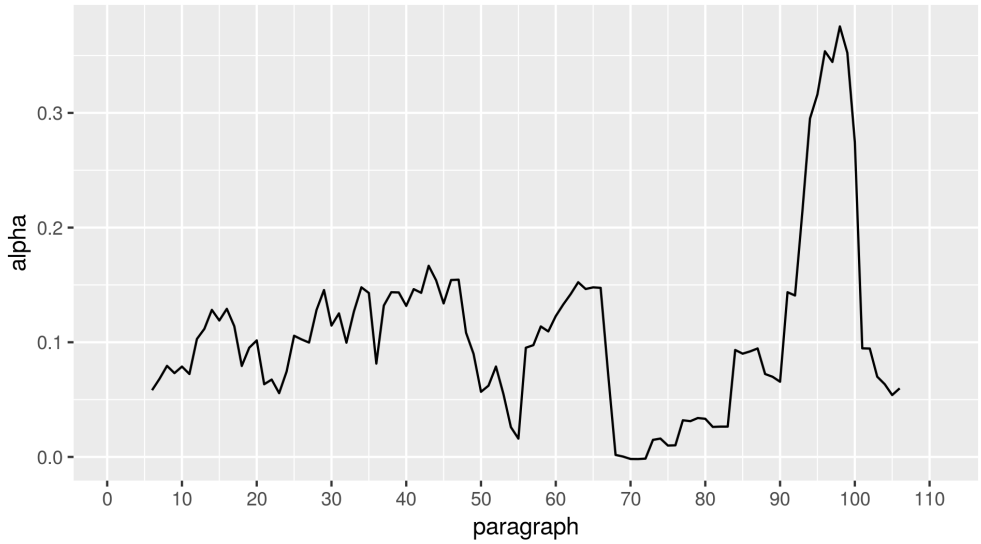


Figure 2: Krippendorff's Alpha for sentiment annotation (51 annotators: POS/NEG; 11-paragraph moving window)

Figure 2 shows the evolution of inter-annotator agreement through a moving window procedure: Krippendorff's Alpha is calculated on just 10% of the text (corresponding to 11 paragraphs), moving from its beginning to its end. The most striking result is in the peak of inter-annotator agreement, that comes not at the very end of the novella (marked by a dominance of positive emotions), but a few paragraphs before. Precisely, it happens around paragraph 98 ("Dapprima, quantunque gli paresse strano, pensò che fossero gli estremi barlumi del giorno"), that signals the beginning of the emotional shift (the "plot twist") in the novella: Ciàula, still fearing the blank darkness of the night, gradually discovers the presence of the moon. This result confirms how an actual sharing of emotions happens not at their climax but with their modification, and transformation—as already noted by Oatley (2012)—is the driving force of narratives.

### 3.2 Correlation in Sentiment Analysis

SA of text and comments was performed using the simplest method (wordcount) implemented by the Syuzhet package. Being Syuzhet designed for the analysis of English language and given the much more inflected nature of Italian language, analysis was performed on lemmatized texts, that were prepared through the UDpipe software. Two Italian sentiment dictionaries were prepared and uploaded in Syuzhet:

- Sentix, where sentiment values were calculated as the product of polarity and intensity;
- OpeNER (Russo et al., 2016), where sentiment values were calculated as the product of sentiment and confidence.

To keep a direct connection between text and comments, for each paragraph of the novella:

- a single sentiment value was calculated for the text;
- the mean of all sentiment values was calculated for the comments.

Figure 3 shows a comparison between the analyses of text and comments with the two sentiment dictionaries. A reference point (the black, dashed line) was set by calculating the means (per paragraph) of

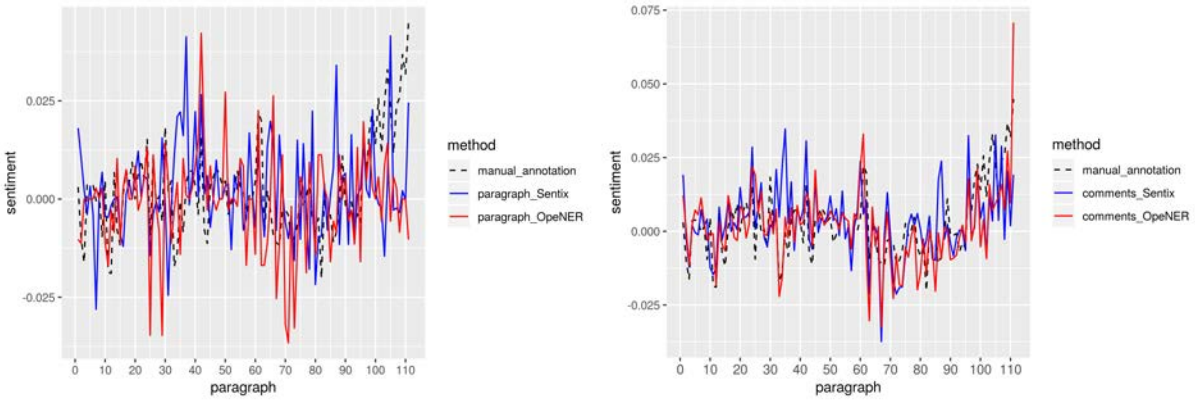


Figure 3: Sentiment analysis of "Ciàula": text (left) and comments (right)

| Focus | Sentix | OpeNER |
|---|---|---|
| paragraph | 0.135 | 0.266* |
| comments | 0.454* | 0.659* |

Table 1: Pearson correlations between SA results and mean values of manual annotation. Asterisks indicate significant correlations (p-value < 0.05)

the sentiment values annotated by the students. A mathematical evaluation of the similarity between the plots was provided by Pearson correlation tests. See Table 1 for an overview of the results.
At least two phenomena call for attention. First, OpeNER seems to achieve better results than Sentix. Second, and most importantly, analyses of comments show much higher correlations than analyses of the commented text. This may be considered as a confirmation of the fact that SA is much more effective when studying reader response, than when analyzing narrative structure, as already shown by Rebora and Pianzola (2018). These results become even more striking when applying the "rolling mean" procedure, implemented in Syuzhet to harmonize plots (see Figure 4): here the similarity can be noticed with the naked eye.

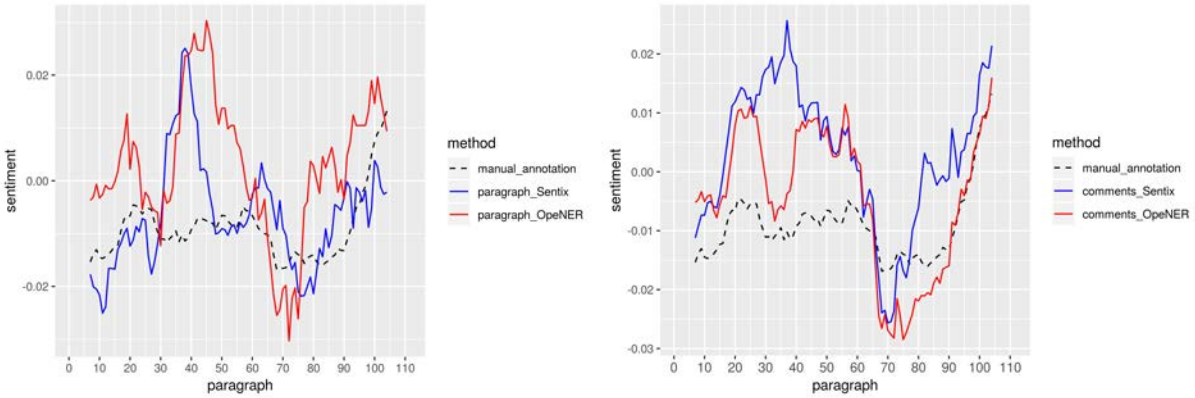


Figure 4: Sentiment analysis of "Ciàula": text (left) and comments (right), normalized with rolling mean

## 4 Conclusion

The small dimensions of the analyzed corpus call for caution when trying to generalize such results. However, they are in line with evidence already presented in previous studies and they call for new research on the topic. In particular, the high correlation in the SA of comments suggests how, notwithstanding the low agreement between readers when trying to evaluate the sentiment of a text, SA is still able to catch

general trends in reader response. At this point, two main lines of enquiry seem advisable: one, that focuses on improving the methodologies further[2]; another, that tries to tighten the connection between computational methods and literary theory.

In conclusion, while being still a very problematic and disputable technique, SA offers multiple stimuli for theoretical and methodological reflection, revealing how, through a direct confrontation with its limitations and imperfections, research in Digital Humanities can still progress towards unexplored grounds.

## 5 Acknowledgments

I thank Tiziana Mancinelli for allowing me to perform this experiment with her students.

## References

Fabio Ciotti. 2017. Modelli e metodi computazionali per la critica letteraria: lo stato dell'arte. In B. Alfonzetti, T. Cancro, V. Di Iasio, and E. Pietrobon, editors, *L'Italianistica oggi*, Adi Editore, Roma, pages 1–11.

Ulrike Edith Gerda Henny-Krahmer. 2018. Exploration of sentiments and genre in spanish american novels. ADHO, Mexico City. https://dh2018.adho.org/exploration-of-sentiments-and-genre-in-spanish-american-novels/

Matthew Jockers. 2014. A Novel Method for Detecting Plot. http://www.matthewjockers.net/2014/06/05/a-novel-method-for-detecting-plot/

Evgeny Kim and Roman Klinger. 2018. A Survey on Sentiment and Emotion Analysis for Computational Literary Studies. *arXiv:1808.03137 [cs]* ArXiv: 1808.03137. http://arxiv.org/abs/1808.03137

Evgeny Kim, Sebastian Padó, and Roman Klinger. 2017. Investigating the Relationship between Literary Genres and Emotional Plot Development. In Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature. Association for Computational Linguistics, Vancouver, Canada, pages 17–26. https://doi.org/10.18653/v1/W17-2203

Klaus Krippendorff. 2018. *Content analysis: an introduction to its methodology*. SAGE, Los Angeles, fourth edition edition.

Bing Liu. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.

Katja Mellmann and Keli Du. 2018. Sentimentanalyse in unstrukturierten Texten (am Bsp. literaturgeschichtlicher-Rezeptionsanalyse). Köln, pages 305–308.

Saif M. Mohammad. 2012. From once upon a time to happily ever after: Tracking emotions in mail and books. *Decision Support Systems* 53(4):730–741. https://doi.org/10.1016/j.dss.2012.05.030.

Keith Oatley. 2012. *The passionate muse: exploring emotion in stories*. Oxford University Press, New York.

Luigi Pirandello. 1986. *Novelle per un anno*. A. Mondadori, Milano. OCLC: 14516480.

Andrew J Reagan, Lewis Mitchell, Dilan Kiley, Christopher M Danforth, and Peter Sheridan Dodds. 2016. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science* 5(1):31.

Simone Rebora. 2017. A Software Pipeline for the Reception of Italian Literature in Nineteenth-Century England. Preliminary Testing. In *Proceedings of the 2nd International Conference on Digital Access to Textual Cultural Heritage (DATeCH)*. ACM, New York, pages 129–134. https://doi.org/10.1145/3078081.3078102

Simone Rebora and Federico Pianzola. 2018. A New Research Programme for Reading Research: Analysing Comments in the Margins on Wattpad. *DigitCult - Scientific Journal on Digital Cultures* 3(2):19–36. https://doi.org/10.4399/97888255181532

Gabriela Rotari. 2018. Digital Analysis of Emotions in the Brothers Grimm's Fairy Tales. EADH, Galway. https://eadh2018.exordo.com/programme/presentation/12

[2]It is undeniable that the method here adopted (that ignores sentiment shifters, does not parse sentences or use machine learning) is quite basic. However, as noted by Kim and Klinger (2018), this is the state of the art for SA in Digital Humanities.

Irene Russo, Francesca Frontini, and Valeria Quochi. 2016. OpeNER Sentiment Lexicon Italian - LMF. http://www.opener-project.eu; https://dspace-clarin-it.ilc.cnr.it/repository/xmlui/handle/20.500.11752/ILC-73

Rachele Sprugnoli, Sara Tonelli, Alessandro Marchetti, and Giovanni Moretti. 2016. Towards sentiment analysis for historical texts. *Digital Scholarship in the Humanities* 31(4):762–772. https://doi.org/10.1093/llc/fqv027

Albin Zehe, Martin Becker, Lena Hettinger, Andreas Hotho, Isabella Reger, and Fotis Jannidis. 2016. Prediction of happy endings in German novels based on sentiment information. In *Proceedings of the workshop on interactions between data mining and natural language processing*. volume 2016, pages 9–16.