DH_BUDAPEST_2019
Centre for Digital Humanities – Eötvös Loránd University

Distant Reading Sentiments and Emotions in Historic German Plays

Thomas Schmidt

Media Informatics Group, University of Regensburg, Germany

thomas.schmidt@ur.de

DH_Budapest_2019, Budapest, Hungary

September, 2019

Abstract for the DH_Budapest_2019 Conference. Abstract Booklet available via http://elte-dh.hu/wp-content/uploads/2019/09/DH_BP_2019-Abstract-Booklet.pdf

Please cite as:

Schmidt, T. (2019). Distant Reading Sentiments and Emotions in Historic German Plays. In *Abstract Booklet, DH_Budapest_2019* (pp. 57-60). Budapest, Hungary.

Sentiments and emotions are important parts of the interpretation of literary texts (Winko, 2003; Mellmann, 2015) and are of special interest for literary scholars interested in plays. Furthermore, many famous playwrights focused on the role of emotions to construct a drama theory for their plays e.g. the *Katharsis* model by Aristotle (335 B.C./1994) or the *Affektlehre* by Lessing (cf. Fick, 2016). Therefore, it is not surprising that sentiment analysis, the research area that is concerned with computational methods to analyze sentiment in written text, has found its way into computational literary studies. Sentiment analysis is used to analyze fairy tales (Alm & Sproat, 2005), novels (Jannidis et al., 2016) and historic plays (Mohammad, 2011; Nalisnick & Baird, 2013) oftentimes with a focus on annotation and evaluation of various methods. However, only few studies explore possibilities of integrating *Distant Reading* (Moretti, 2013) and visualizations into sentiment analysis (Kakkonen & Kakkonen, 2011; Mohammad, 2011; Nalisnick & Baird, 2013).

We present a web-based Distant Reading-tool to explore sentiments and emotions in historic German plays. We focus on the analysis of single plays and illustrate our approach on a test corpus of 12 plays by Gotthold Ephraim Lessing (1729-1781). We discuss various use cases on how visualizations of our tool can be used for interpretation and what insights we could gain. The tool is available online[1].

First, we distinguish between three main-concepts of analysis: structural analysis, character analysis, and analysis of character relationships. To calculate metrics we employ a rule-based approach using lexicons with sentiment annotations for words. This is a well-known approach in sentiment analysis and in a previous study we evaluated various lexicons and NLP techniques on a sub corpus of speeches to identify the best performing method (Schmidt & Burghardt, 2018). For those methods, the number of sentiment words in a textual unit are calculated to get an overall value. We calculate values for the polarity (positive/negative) and eight emotion categories. Users can regard either absolute values or values normalized by the number of all words of a text unit.

For the first use case, we analyzed the polarity progression throughout all the acts of Lessing's plays and were able to identify a constant progression to larger amounts of negativity leading up to the fifth act. Figure 1 exemplifies this finding for one play. This result illustrates how the plot becomes more and more negative since disputes and conflicts become more apparent towards the end.

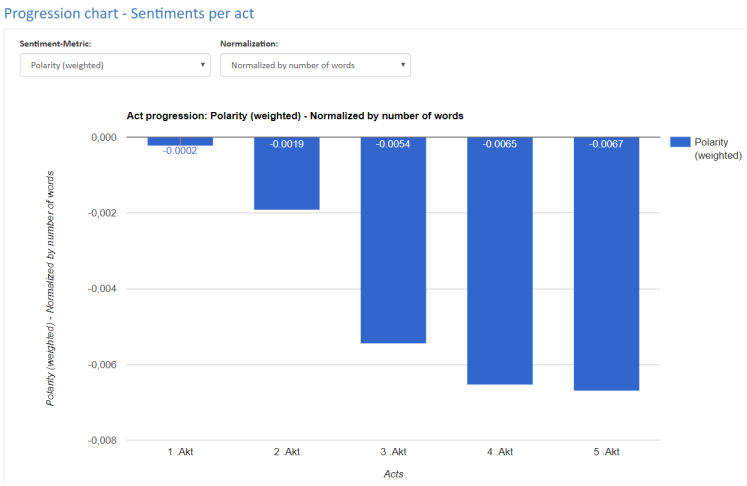

Figure 1. Polarity progression per act throughout the play *Der Freigeist*

[1] https://lauchblatt.github.io/Katharsis/sa_selection.html

Another use case for this tool is to identify emotion driven parts of the play. Figure 2 is a line chart for the scene-based progression of the emotion *fear* throughout the play. Researchers can hover over the graph to identify the scenes that are most connoted with fear.

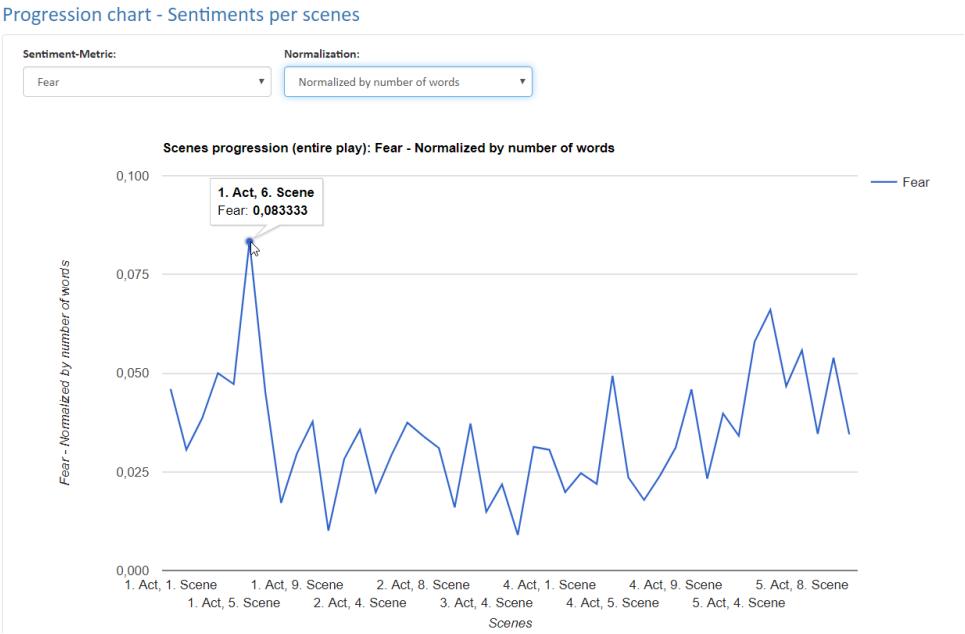

Figure 2. Scene-based progression of the emotion *fear* throughout the play *Miß Sara Sampson*

For a more character-specific use case, we focus on the character *Marinelli*, who is the villain in the play *Emilia Galotti*. Analyzing his speeches in the entire play indeed shows a tendency towards negativity (Figure 3):

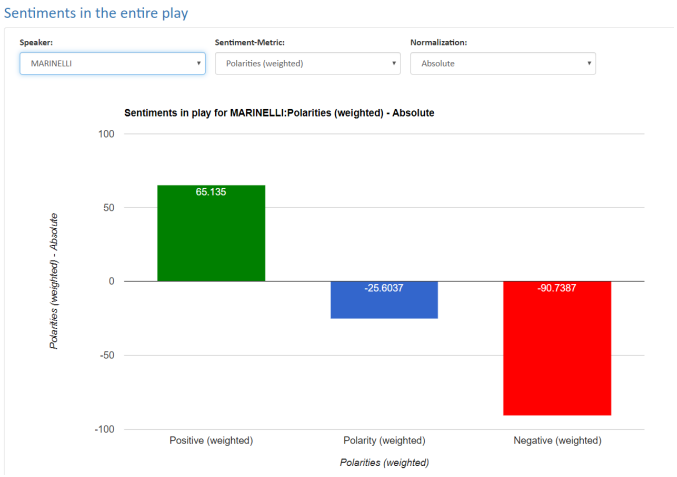

Figure 3. Overall polarity of the character *Marinelli* of the play *Emilia Galotti*

Comparing Marinelli's speeches to other characters, he is ranked as the most negative character in absolute values (see Figure 4).

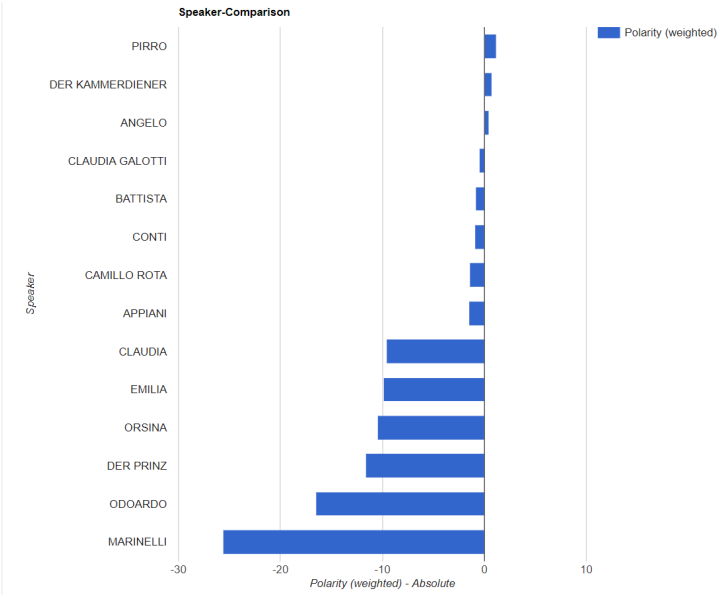

Figure 4. Polarity comparisons of all characters of the play *Emilia Galotti*

Analyzing the character relationships, the results for Marinelli are in line with the plot. To calculate relationships we regard heuristically every speech a character expresses as directed towards the previous speaker (Nalisnick & Baird, 2013) and take all those character-to-character speeches as input for the sentiment analysis of relationships. Figure 5 shows the results for Marinelli in the entire play:

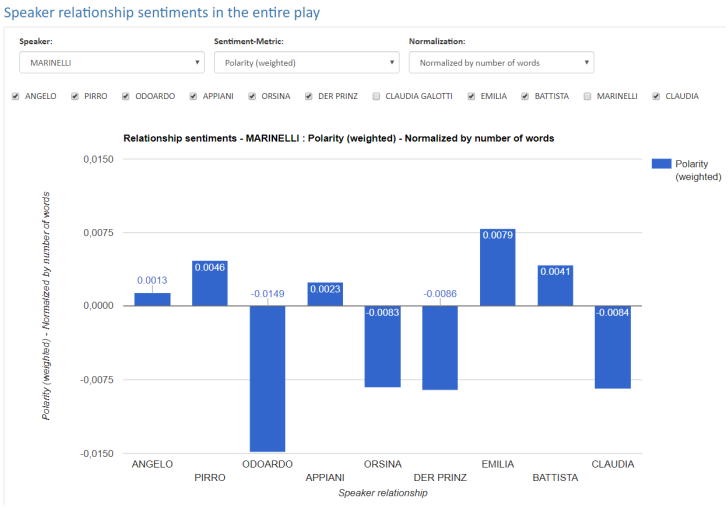

Figure 5. Polarities of the character relationships concerning the character *Marinelli* in the play *Emilia Galotti*

For many of the main characters the relationships prove to be rather negative. However, it is telling that the relationships towards *Angelo, Pirro* and *Battista* are positive since those are Marinelli's allies. The positive relationship between Marinelli and Emilia shows the limitations of sentiment analysis. Although Marinelli plans to damage Emilia throughout the entire play and his behavior is instrumental in her committing suicide the relationship is shown as positive. The reason for this is that when both characters meet in the play, Marinelli acts very nice and polite disguising his true intentions which can't be noticed on a solely textual level.

We would like to present more examples illustrating Distant Reading-possibilities of sentiments but also emotions in plays with our tool and discuss limitations and future work in more detail during a presentation if our contribution gets accepted.

References

Alm, C. O. & Sproat, R. (2005). *Emotional sequencing and development in fairy tales.* In International Conference on Affective Computing and Intelligent Interaction (pp. 668-674). Springer Berlin Heidelberg.

Aristotle. (334 B.C./1994). *Poetik: Griechisch / Deutsch*. Ditzingen: Reclam.

Fick, M. (2016). *Lessing-Handbuch. Leben – Werk – Wirkung*. Stuttgart; Weimar: Verlag J.B. Metzler.

Jannidis, F., Reger, I., Zehe, A., Becker, M., Hettinger, L. & Hotho, A. (2016). *Analyzing Features for the Detection of Happy Endings in German Novels.* arXiv preprint arXiv:1611.09028.

Kakkonen, T. & Kakkonen, G. G. (2011). SentiProfiler: creating comparable visual profiles of sentimental content in texts. In *Proceedings of Language Technologies for Digital Humanities and Cultural Heritage* (pp. 62-69).

Mellmann, K. (2015). Literaturwissenschaftliche Emotionsforschung. In: Rüdiger Zymner (Hg.): *Handbuch Literarische Rhetorik*. Berlin/Boston, 173-192.

Mohammad, S. (2011). From once upon a time to happily ever after: Tracking emotions in novels and fairy tales. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 105-114). Association for Computational Linguistics.

Moretti, F. (2013). *Distant reading.* Verso Books.

Nalisnick, E. T. & Baird, H. S. (2013). Character-to-character sentiment analysis in shakespeare's plays. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 479–483).

Schmidt, T. & Burghardt, M. (2018). An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing. In: *Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 139-149). Santa Fe, New Mexico: Association for Computational Linguistics.

Winko, S. (2003). Über Regeln emotionaler Bedeutung in und von literarischen Texten. In: Fotis Jannidis & Gerhard Lauer & Matias Martinez & SW (eds.): *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte.* Berlin, New York: de Gruyter, 329-348.