The Course of Emotion in Three Centuries of German Text—A Methodological Framework

Sven Buechel[1] Johannes Hellrich[1,2] Udo Hahn[1]

{sven.buechel,johannes.hellrich,udo.hahn}@uni-jena.de

[1] Jena University Language & Information Engineering (JULIE) Lab
http://www.julielab.de

[2] Graduate School "The Romantic Model"
http://www.modellromantik.uni-jena.de

Friedrich Schiller University Jena,
Jena, Germany

Introduction

Texts not only carry factual, but also affective information, such as expressions of joy or grief. In the humanities, especially literary studies, emotion expression and elicitation (often in texts from prior language stages) have been focused on in many contributions (see, e.g., Carroll and Gibson (2011), Poppe (2012), Hillebrand (2011)).

Similarly, in natural language processing (NLP), emotion analytics have developed into an active area of research (Liu, 2015). Nevertheless, there is little previous work explicitly addressing emotion in historical language and the specific methodological problems this raises. Hamilton et al. (2016) as well as Cook and Stevenson (2010) presented methods for identifying amelioration and pejoration of words. Acerbi et al. (2013) and Bentley et al. (2014) demonstrated the potential of emotion analysis for the Digital Humanities (DH) by linking temporal emotion patterns in texts to major sociopolitical events and trends in the 20th century.

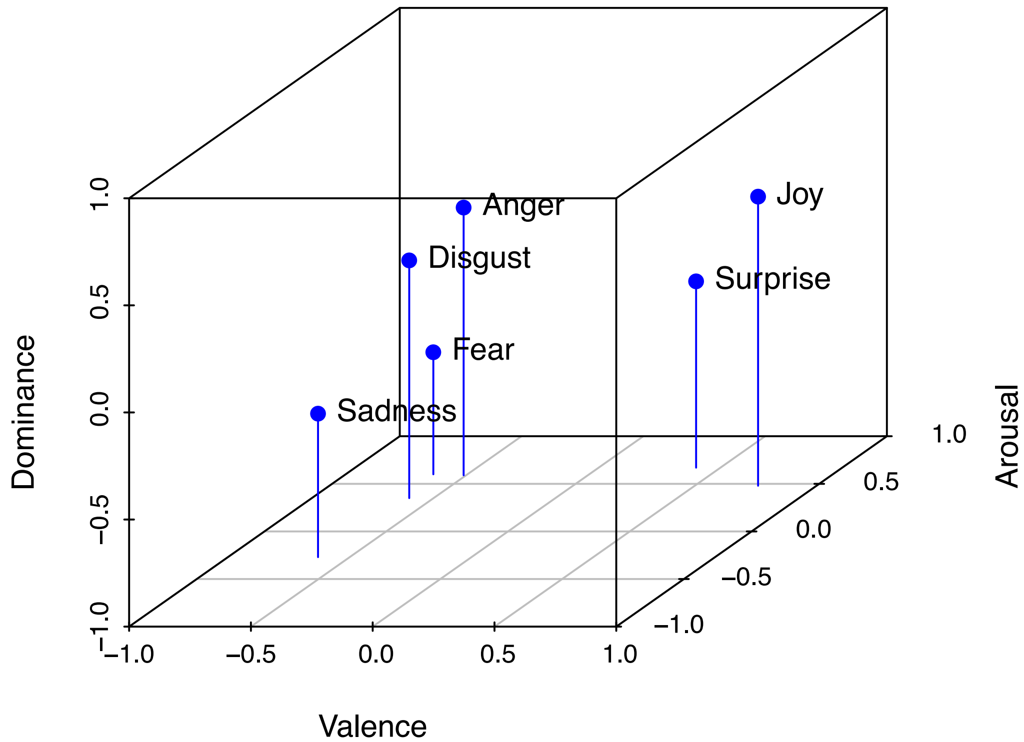
Our work goes beyond these studies in two ways: we claim to be more adequate as we combine these two approaches to analyze non-contemporary **text** based on time-specific **lexical** resources. We also claim to be more informative as we employ the Valence-Arousal-Dominance (VAD) model of emotion (Bradley and Lang, 1994) instead of simple **polarity** (positiveness/negativeness) alone. We have already shown the latter to be beneficial in DH applications (Buechel et al., 2016a). We hope that our work will be a step towards a new set of tools especially beneficial for areas such as literary studies (e.g., in distant reading (Moretti, 2013)) or history of mind.

Methods

The VAD model of emotion assumes that affective states can be described relative to three emotional **dimensions**, i.e., Valence (corresponding to the concept of polarity, see above), Arousal (the degree of excitement or calmness) and Dominance (feeling in or out of control of a social situation). The VAD dimensions allow for a more fine-grained modeling than polarity alone, e.g., words like *orgasm* and *relaxed* have similar Valence but opposing Arousal values (Bradley and Lang, 1999). Formally, the VAD model constitutes a three-dimensional vector space illustrated by Figure 1 (Buechel and Hahn, 2016).

The association of words with a VAD score is captured in **emotion lexicons**. These can either be empirically determined by aggregating subjective judgments of several human subjects; or they can be semi-automatically constructed allowing for greater size but reducing accuracy on individual words. For the semi-automatic construction, the typical approach is to expand an existing lexicon (the **seed lexicon**) based on word similarity (see below). There are several competing expansion algorithms. Cook and Stevenson (2010) were the first to describe expansion algorithms for the induction of the emotion value of words for non-contemporary language stages by using word similarity values determined from historical corpora.

Figure 1: The VAD vector space. For ease of understanding, the positions of six *Basic Emotions* (Ekman, 1992) are given.



Extending this approach, we compared several emotion induction algorithms, viz., those by Turney and Littman (2003), Hamilton et al. (2016), and Bestgen (2008). The former two were slightly modified to make them deal with numerical input values (for a more detailed description of these methods, see Buechel et al. (2016b) and Hamilton et al. (2016)).

We used point-wise mutual information with singular value decomposition (Levy et al, 2015; SVD_{PPMI}) to measure word similarity, since it turned out to be superior for DH applications in previous work (Hellrich and Hahn, 2016). We used the German ANGST lexicon (Schmidtke et al., 2014) as seed. The individual algorithms were evaluated by comparing our induced historical lexicons against judgments of knowledgeable PhD students from the humanities. For this task, we asked them to make their assessments **as if** they were contemporary readers from the respective time period. The Turney-Littman algorithm performed best in this set-up and was thus employed for all subsequent analyses.

Experiments

For demonstration purposes, we here apply our methodology to the core corpus of the *Deutsches Textarchiv*[1] (DTA; Geyken, 2013) [German Text Archive], a well-curated and balanced collection of historical German texts. We analyzed texts created between 1690 and 1899, splitting the resulting corpus into seven slices (each spanning 30 years) to achieve similarly sized and sufficiently large subcorpora for further processing. We computed word similarities within each of these slices and then applied the Turney-Littman expansion algorithm, thus creating seven distinct emotion lexicons, each describing the emotion of words for its specific period. Given these temporally stratified lexicons, we claim that shifts in emotion association of words can be traced over time by comparing the emotion values a word exhibits in different lexicons. To validate this claim, we selected the words for which we could compute similarity scores in each time step[2] and standardized their VAD values for each lexicon and dimension (VAD).

---

[1] TCF version from May 11, 2016, available via `www.deutschestextarchiv.de/download`.
[2] As these methods are more accurate for high-frequency words, rare words were excluded from our study.

Figure 2: Development of the lexical item *Sünde* [sin] during the study period relative to the VAD dimensions.

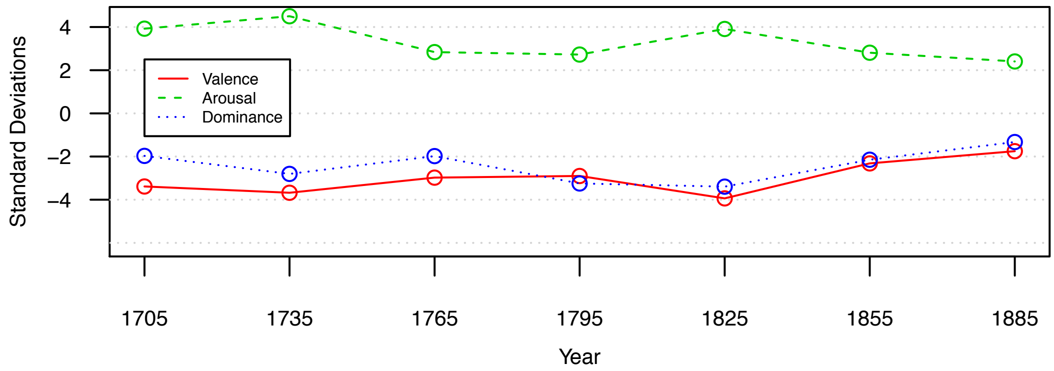


Table 1: Top ten collocations of the lexical item *Sünde* [sin] in the DTA corpus comparing the 1690s and the 1890s using pointwise mutual information for scoring. Source: `http://kaskade.dwds.de/dstar/dta/diacollo/`

Rank	Lemma and Translation			
	1690s		1890s	
1	todt-	(German prefix for ‘death’ as in ‘deadly sins’)	Lamm	lamb
2	Erzürnung	infuriation	hinwegnehmen	to take away
3	läßlich	minor (clerical)	Verzeihung	forgiveness
4	beichten	to confess	Ausschweifung	excess
5	Nachlaß	abatement/ inheritance	Gotte	god
6	Grobheit	crudeness	Schande	shame
7	verschweigen	to conceal	Reue	repentance
8	beweinen	to beweep	Ärgernis	nuisance
9	pichen	to pitch	Laster	vice
10	beichten	to confess	aufrichtig	sincere

We illustrate this approach with an example from Figure 2. It displays an overall amelioration of *Sünde* [sin] whose onset roughly coincides with the age of enlightenment—often understood as the starting point of secularization (Sheehan, 2003).[3] This observation is in line with well-known findings from descriptive lexicography (Grimm and Grimm, 1942). The same semantic shift can also be discovered by a more established method, namely collocation analysis.

Table 1 reveals that *Sünde,* at the end of our examination period, has acquired an additional moral-bourgeois meaning facet (implied, e.g., by *Ausschweifung* [excess], *Ärgernis* [nuisance] and *Laster* [vice]) which was not present in the beginning. There, only the religious sense is traceable.

Going one step further, we then used these lexicons to examine how emotion is distributed over literary texts in the DTA in the course of time. We employed the *Jena Emotion Analysis System* (JEmAS; Buechel and Hahn, 2016),[4] an open-source tool for emotion measurement using a configurable VAD lexicon. We processed each DTA text with the period-aligned lexicon, leading to the main methodological contribution of our work: linking the research areas of automatically inducing historical **word** emotion (e.g., Hamilton et al., 2016) and emotion prediction in historical **text** (e.g., Acerbi et al., 2013).

We scaled the resulting emotion values within each VAD dimension tracing the development of the three principal literary forms—Narrative, Lyric, and Drama—in German literature between 1690 and 1899. For each of the seven 30-year periods (organized in rows), we created three scatterplots (one for each pair of the VAD dimensions; organized in columns) resulting in 21 plots in total (Figure 3). Each data point represents one text—color and shape represent membership to the respective form.

It is evident from the plots how the distinction of the individual forms increases and decreases in emotional terms in the course of time.[5] We find the most distinct emotional patterns between 1780 to 1809 (approximately covering the Weimar Classicism) and between 1870 to 1899 (covering the late German Realism). Drama displays consistently more Arousal than Lyric and Narrative since 1750, whereas Lyric seems to be the most

[3] Note however, that care must be taken when interpreting these word course graphs since noise can be introduced from various sources (such as word similarity and emotion induction algorithms); both strongly depend on the amount of data available for each time step.

[4] `https://github.com/JULIELab/JEmAS`

[5] This application differs from typical stylometric approaches since we employ emotional features instead of word counts.

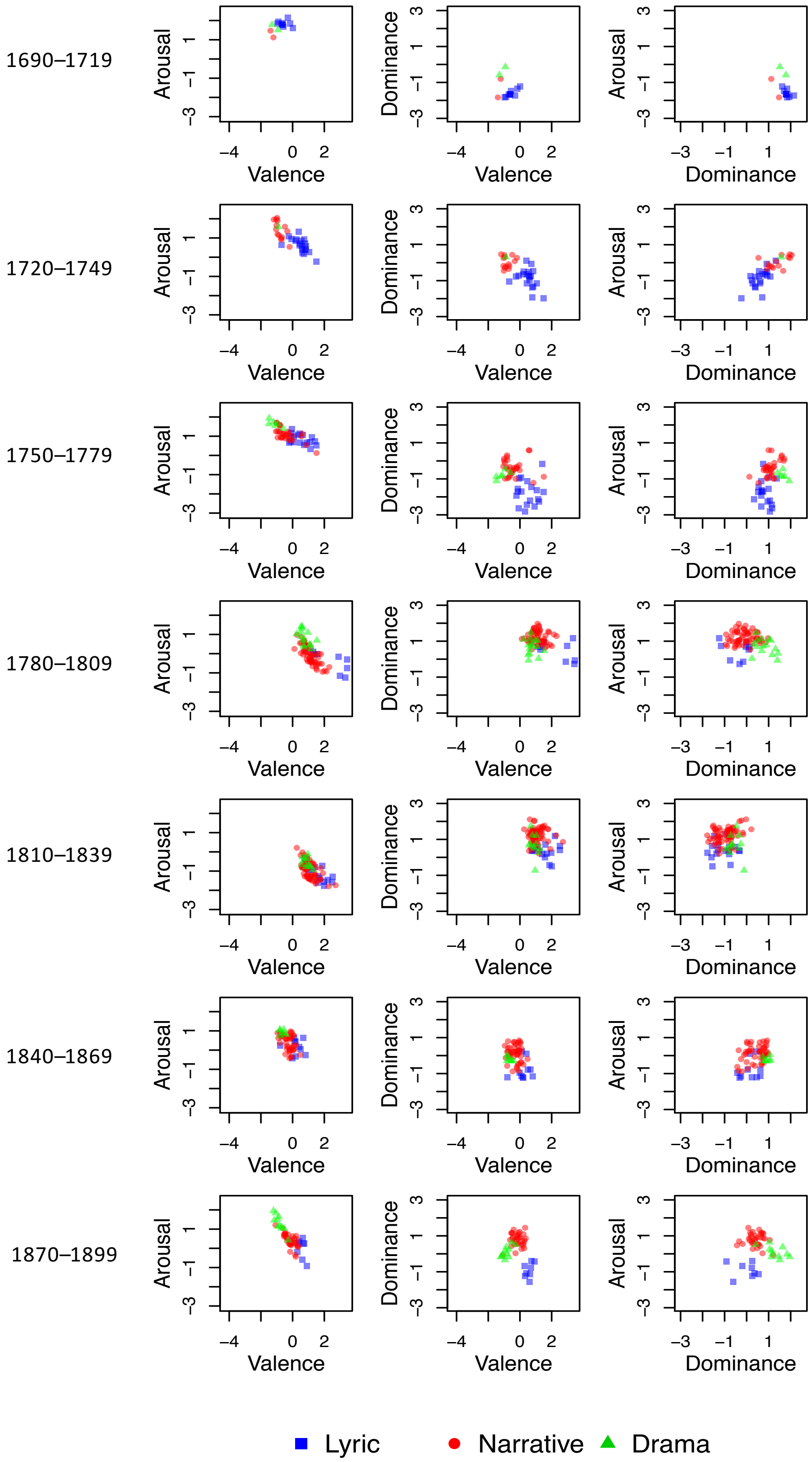
positive class (Valence) expressing the least control (Dominance). Of course, the examination of the DTA offers many more intriguing findings, however, for brevity, we limit ourselves here to presenting examples.

Conclusion

In this contribution, we described a novel methodological framework for quantifying emotion in non-contemporary text. Applying this approach to a 210-years section of the German DTA corpus, we find clear emotional signals for temporally evolving distinctions between the principal literary forms, viz. Narrative, Lyric, and Drama. All resources and software we developed for this work are publicly available.[6]

6 https://github.com/JULIELab/HistEmo

Figure 3: Distribution und development of the principal literary forms, Lyric (blue), Drama (green) and Narrative (red), relative to each pair of VAD emotions (in columns) between 1690 and 1899 (each row representing a 30-year slice).



Acknowledgments

This research was partially conducted within the Graduate School “The Romantic Model" supported by grant GRK 2041/1 from the Deutsche Forschungsgemeinschaft (DFG).

References

Acerbi, A., Lampos, V., Garnett, P. and Bentley, R.A. (2013). The expression of emotions in 20th century books. *PLoS ONE* 8(3): e59030.

Bentley, R.A., Acerbi, A., Ormerod, P. and Lampos, V. (2014). Books average previous decade of economic misery. *PLoS ONE*, 9(1): e83147.

Bradley, M.M. and Lang, P.J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1): 49–59.

Bradley, M.M. and Lang, P.J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective Ratings. Report C–1, The Center for Research in Psychophysiology, University of Florida, Gainesville.

Bestgen, Y. (2008). Building affective lexicons from specific corpora for automatic sentiment analysis. *Proceedings of the Sixth International Conference on Language Resources and Evaluation,* pp. 496–500.

Buechel, S. and Hahn, U. (2016). Emotion analysis as a regression problem: Dimensional models and their implications on emotion representation and metrical evaluation. *Proceedings of the 22nd European Conference on Artificial Intelligence*, pp. 1114–22.

Buechel, S., Hahn, U., Goldenstein, J., Händschke, S.G.M. and Walgenbach, P. (2016a). Do enterprises have emotions? *Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pp. 147–53.

Buechel, S. Hellrich, J. and Hahn, U. (2016b). Feelings from the past—Adapting affective lexicons for historical emotion analysis. *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities*, pp. 54–61.

Carroll, N., and Gibson, J. (eds) (2011). *Narrative, Emotion, and Insight.* University Park: Pennsylvania State University Press.

Cook, P. and Stevenson, S. (2010). Automatically identifying changes in the semantic orientation of words. *Proceedings of the 7th International Conference on Language Resources and Evaluation*, pp. 28–34.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4): 169–200.

Geyken, A. (2013). Wege zu einem historischen Referenzkorpus des Deutschen: das Projekt Deutsches Textarchiv. *Perspektiven einer corpusbasierten historischen Linguistik und Philologie*, pp. 221– 34.

Grimm, J. and Grimm, W. (eds) (1942). Sünde. In: *Deutsches Wörterbuch*, volume 20.

Hamilton, W.L., Clark K., Leskovec, J. and Jurafsky D. (2016). Inducing domain-specific sentiment lexicons from unlabeled corpora. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 595–605.

Hellrich, J. and Hahn, U. (2016). Bad company—Neighborhoods in neural embedding spaces considered harmful. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics*: Technical Papers, pp. 2785–96.

Hillebrand, C. (2011). *Das emotionale Wirkungspotenzial von Erzähltexten: Mit Fallstudien zu Kafka, Perutz und Werfel*. Berlin: Akademie Verlag.

Levy, O., Goldberg, Y. and Dagan, I. (2015). Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association for Computational Linguistics*, 3: 211–25.

Liu, B. (2015). *Sentiment Analysis: Mining Opinion, Sentiments and Emotions*. New York: Cambridge University Press.

Moretti, F. (2013). *Distant Reading*. London: Verso.

Poppe, S. (2012). *Emotionen in Literatur und Film*. Würzburg: Königshausen & Neumann.

Schmidtke, D.S., Schröder, T., Jacobs, A.M. and Conrad, M. (2014). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior Research Methods*, 46(4): 1108–18.

Sheehan, J. (2003). Enlightenment, religion, and the enigma of secularization: A review essay. *The American Historical Review*, *108*(4): 1061–80.

Turney, P.D. and Littman, M.L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 21(4): 315–46.