Article

Connecting the Dots
Variables of Literary History and Emotions in German-language Poetry

Leonard Konle[1] (iD)
Merten Kröncke[2] (iD)
Simone Winko[2] (iD)
Fotis Jannidis[1] (iD)

1. Institut für Deutsche Philologie, Julius-Maximilians-Universität Würzburg ROR, Würzburg, Germany.
2. Seminar für Deutsche Philologie, Georg-August-Universität Göttingen ROR, Göttingen, Germany.

Abstract. In this study, we attempt new steps towards quantitative literary history, aiming to identify factors relevant to literary history and to model assumptions about the relationships between them. We use a case study of German-language poetry in the transition from realism to early modernism to approach our methodological goal. Using a Bayesian hierarchical generalised linear model, we focus on one aspect relevant to the history of poetry: The emotions represented in the poems, and also include period, author profession, author gender, rhyme, and verse length in the model. We can confirm the important role of thematic genres and find unexpectedly high values for rhyme and author's profession. We also discuss some of the methodological of our attempt to model this tangled web of variables.

1. Introduction

> Any association between an outcome and a predictor can be nullified or reversed when another predictor is added to the model. And the reversal can reveal a true causal influence or rather just be a confound.
> (Richard McElreath 2020, 345)

Most literary histories are based, at least implicitly, on the assumption that the history of literature is determined, on the one hand, by a more or less strong dynamic of its own and, on the other hand, by a number of external factors.[1] These factors, be they the history of the media, the history of reading, the history of social structures, developments in the history of ideas, or the history of mentality, such as the change in ideas about man, etc., are often presented in separate chapters, and their literary-historical relevance may be illustrated by individual texts. However, it is notoriously difficult to show that even one of these factors is relevant to all literary texts, even if there are examples for each of them that make it impressively clear that certain texts can only be understood if one takes the respective focused factor into account. This sometimes

---

1. Here, and throughout the rest of the introduction, we are thinking of works such as those mentioned in section 6, e.g. (for German literary history) *Hansers Sozialgeschichte der Literatur* or *Geschichte der deutschen Literatur von den Anfängen bis zur Gegenwart*.

leads to the description of literary history as a whole as an autonomous process. This view is further supported by the fact that the connection between factors external to literature and the literary text and its specific form can only be plausibly modeled in a few exceptional cases. In addition, literary history hardly has access to methods for the systematic aggregation of individual observations in such a way that the frequency of a phenomenon is taken into account and exceptions are not regarded as refutations. This is where literary-historical projects that use digital corpora and statistical methods come in. They can make digitally based statements about relationships between factors modeled in the analysis and literary texts. These statements are limited to the texts and factors under investigation: One cannot see what one does not see. The complexity of the subject matter complicates the study: There are numerous factors that must be taken into account; we do not know at this point exactly what the role of each of these factors is. Many of the factors clearly interact with each other, making their statistical analysis particularly difficult. Our work should be seen as a step towards a quantitative literary history that attempts to identify the relevant factors.[2] As noted above, this can only be done on a specific corpus. Whether the results are transferable to other corpora representing other genres and the literature of other periods, languages, and cultures is a separate research problem that can only be addressed as more studies of this kind become available. Maybe the important factors can only be identified in the context of a genre or a period or other groups.

At the same time, the timing is particularly favorable: Sufficient numbers of literary texts are accessible in digital form and methods for analyzing complex relations are available. On the other hand many aspects, especially extra-literary factors, cannot yet be integrated into the model because the data is not yet available or not accessible in digital form. This is another reason why these first results are very preliminary. Despite these shortcomings, we believe that the attempt made in this paper is beneficial to the field, if only to learn from it and lay a foundation for further research.

We conduct our experiments and methodological considerations on a case study: On the transition from the poetry of realism to early modernism. Our corpus consists of about 6000 poems from these periods. In what follows, we focus on the question of whether, and to what extent, factors contribute to one aspect of the texts, namely the emotion that is thematized. Our focus is on emotion because, firstly, we have already gathered relevant data on emotions in previous studies, and, secondly (and more importantly) there is ample evidence in literary history and contemporary accounts that the difference between the periods is evident in the literary representation of emotions. In this paper, we will look only at a small set of factors and how they are related to the emotions of poetry in a regular way. We have chosen the factors we focus on because they have been deemed relevant by research and because we can represent them as data. Some of them are external factors like gender or profession of the author, others are text aspects like rhyme or the thematic genre. We do not assume that the same kind of real-world causality underlies all these factors. The gender of the author may have an influence on

2. While the first large wave of studies in CLS was mainly exploratory, we saw in the second wave more studies which systematically tested hypotheses. Only very recently did this include the study of factors in literary history (Underwood et al. 2022). As far as we can see, our work is one of the first attempts to identify a number of different factors for a particular part and aspect of the history of literature – emotions in German poetry 1850-1920 – and determine their varying degrees of relevance.

the selection of emotions depicted in a poem, but not the other way around. The same is probably not true in the same way for thematic genre or rhyme.

2. Research

In literary studies, the relationships between literary texts and other literary or extra-literary phenomena are often discussed under the term "context". We highlight two related aspects of the debate on context that are relevant to our work:

1. One strand of discussion focuses on the extent to which context is relevant to understanding and/or explaining literature and literary change in general. Some approaches, e.g., Marxism, argue for strong contextual influences or even determinacy, while other approaches assume that literature is either autonomous and independent of external factors or should at least be treated as such (Kalliney 2019; Ladegaard and Nielsen 2019).

2. For those approaches assuming that contexts play at least some role for literary texts, another much-debated question is how to select and weigh particular contexts (King and Reiling 2014; Borkowski 2015; Engel 2018; Thomsen 2019, 207) – a question of great importance for this paper. The discussions on context selection and context weighting are also related to the issue of how to explain literary change (see Gittel 2016 for a distinction between different types of explanations). Some positions, often influenced by particular literary theories, name specific contexts they consider relevant. For example, feminist literary theory typically assumes that the gender of the author and/or the role and status of women in society, culture, economics, and politics are important contexts for understanding literature and literary change. Other positions do not focus on specific contexts but on *criteria* for contexts. For example, King and Reiling (2014) propose "relevance", "representativeness", and "usefulness" as criteria for context selection.

When discussing the relevance of context in general or the selection and weighting of specific contexts, literary scholars mostly rely on theoretical considerations, on examples from literary history, or on individual text analysis. While these types of arguments are certainly valuable, what seems to be missing is a practical method for comparatively analyzing the relevance of context(s) that is suitable for our purposes. At the very least, we are not aware of any method proposed in literary studies that is (a) not limited to very specific text corpora or time periods, (b) somewhat independent of theory-specific presuppositions (such as "X is always the most important context"), and, most importantly, (c) computationally operationalizable in a reasonably clear, practicable, and intersubjective way. For example, the criteria of King and Reiling (2014) are so abstract that they immediately raise the question of how to figure out what is "relevant", "representative", or "useful". There are no obvious answers. King and Reiling (2014, 21) themselves write that their suggestions are rather a "heuristic framework" and "rough criteria" that complement the "methodological evidence of the individual case".

It seems possible to use the context selection criteria mentioned in literary studies as helpful heuristics that can narrow down the set of potential contexts worth investigating. Beyond that, however, we need to draw on other resources to achieve the methodological

goals of this paper. At best, this will also allow us to contribute to the ongoing debates about "context" in literary studies.

Identifying relevant factors in historical processes is a comparatively new perspective in Computational Literary Studies, but there is field which has produced a series of studies with this goal, and its unifying characteristic is a view on culture as an evolutionary process. In this context, single works are not important but historical trends and shifts are (see for example section VII in Barrett and Dunbar 2007). In recent years more and more studies appeared using the framework to analyze larger data collections, for example Sobchuk and Tinits (2020), who analyze the development of anachronies in mystery film between 1970 and 2009. But many of them, like Sobchuk and Tinits, describe a trend in the data quantitatively and discuss at the end on a qualitative level possible predictor variables. This is also the case in most studies from Computational Literary Studies. Usually, they identify an interesting trend in their data and propose reasons for the observed pattern. Discussion and controversies concentrate on the question, whether the trends are really there (Langer et al. 2021; Piper 2022). But this slowly begins to change. Using corpora in three languages, Šeļa et al. (2021) can show that metrical forms and the semantic features of corresponding poems are systematically linked. Underwood et. al. (2022) seem to be the first ones to discuss a causal factor in literary history based on the analysis of a large corpus. They find that cohort succession plays an important role in literary history.

3. Resources

3.1 Corpus

Our corpus includes poems from two periods: From realism and the period 'around 1900' (or early modernism). The poems in question were published in those anthologies that aim to provide their audience with an overview of contemporary poetry (in part also of 'the best' contemporary poetry). Thus, the corpus texts have been labeled, so to speak, by contemporary poetry experts. The 'realism' sub-corpus consists of eight anthologies published between 1859 and 1882, the 'around 1900' sub-corpus of 12 collections published between 1885 and 1911. Thus, the entire corpus contains 6,249 poems from 20 anthologies (for a description of the corpus, see Winko et al. 2022).

3.2 Emotion

Before analyzing the association of different features with the distribution of emotions in poetry, it is necessary to determine how often which poems represent which kind of emotions in the first place. For this purpose, we rely on a machine learning setup that has been trained and evaluated on manual annotations, which we have already described (Konle et al. 2022).

We annotated the representation of emotions in 1,352 corpus poems. The goal was not to annotate readers' emotions, but rather the emotions represented in the text itself, e.g., whether the speaker or a character is happy, sad, in love, etc. The annotators used a list of 40 discrete emotions (e.g., love, joy, surprise, envy, regret, fright), the selection is based on existing emotion models (e.g. Ekman 1992, Ekman 1999; Plutchik 1980a,

Emotion	Joy	Love	Sadness	Anger	Fear	Agitation
f1 (macro)	0.73	0.77	0.74	0.71	0.79	0.62

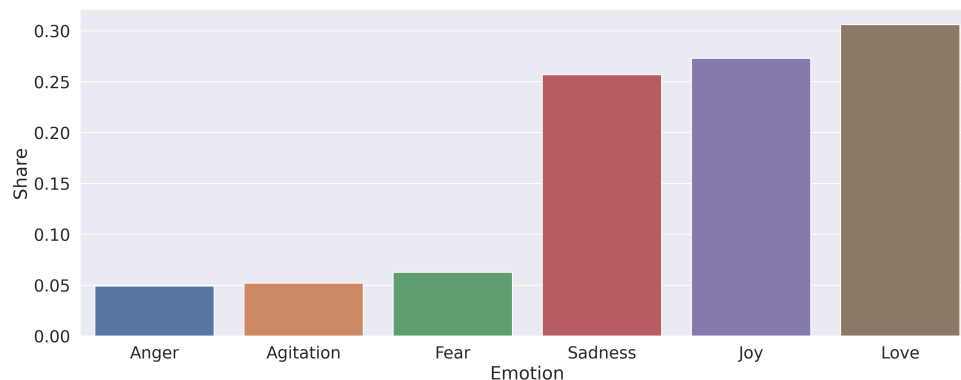**Table 1:** Quality of emotion classification.

Figure 1: Share of predicted emotions.

Plutchik 1980b, Plutchik 2001) and on the emotions that were regularly represented in the poems of our corpus. Because a substantial number of emotions were only infrequently annotated (e.g., disgust), we categorized the emotions after annotation into 6 major groups, inspired by the emotion hierarchy in Shaver et al. (1987): agitation, anger, fear, love, joy, sadness. First, each poem was annotated independently by two annotators, who then manually merged the annotations into a consensus annotation. Their agreement, measured with γ (Mathet et al. 2015), was 0.6445 for individual emotions and 0.7491 for the emotion groups (annotation guidelines: Kröncke et al. 2022), where 0 indicates agreement no better than chance and 1 indicates perfect agreement.

In order to detect emotions automatically, we defined an emotion classification task as a series of binary classifications to avoid the complexity of multi-labeling. The basis of our classification is the German BERT (Devlin et al. 2018) model gbert-large (Chan et al. 2023). Because gbert is trained on contemporary webtext, we continue its pre-training[3] with poetry to adapt the model to our target domain. Subsequently we perform fine-tuning on the binary emotion classification tasks. To overcome the class imbalance we apply undersampling by randomly sampling examples from the majority class in every epoch. While the classification of single emotions leads to a large spread in predictive quality,[4] the grouped emotions lead to more stable performance at an acceptable level of uncertainty (Table 1).

For all further analyses, we focus on the six emotion groups, using consensus annotations where possible and relying on model predictions otherwise.

The poems represent love, joy, and sadness most often; anger, fear, and agitation occur much less frequently.

---

3. Hyperparameters: 500 steps, batchsize 30, learningrate 2e-5 (Konle and Jannidis 2020, Gururangan et al. 2020).
4. Very frequent emotions like longing (f1: 0.73) or suffering (f1: 0.72) yield sufficient classifiers, but less frequent ones like calmness or desire lead to results similar to a random baseline.

3.3 Features

We cannot rule out the possibility that a myriad of features affect the distribution of emotions in poems, including various aspects of the author's biography (e.g., their gender or attitude toward religion) and the poem's content, form, and style (e.g., its themes or stylistic register). However, for reasons of data availability alone, it is not possible to include all conceivable features in our study. Therefore, we confine our analysis to a limited set of features selected according to several criteria: (1) Other researchers consider the features to be relevant to the literary representation of emotions and/or the features appear to be important given the context selection heuristics mentioned in literary studies (section 2) and/or studies from our working group indicate that there is a relationship between the features and the emotions represented in our corpus; (2) the features cover a wide range of textual (e.g., theme) and extratextual phenomena (e.g., author gender); and (3) respective data is available for the entire corpus. As indicated, the restriction to a limited set of features means that our analysis certainly cannot encompass all features that might affect the representation of emotions in poetry. Nevertheless, we are confident that by applying the criteria mentioned above, we have selected features that, taken together, can provide relevant insights.

For each feature, the following sections justify why we include it in our analysis and how we collected the data.

Period Literary periods like realism or modernism have traditionally been among the most important categories for organizing and characterizing literary texts. When examining various aspects of literature, the literary period is almost always considered a possible factor of difference. Likewise, researchers assume that the literary representation of emotions differs significantly depending on the period (e.g. Andreotti 2014, 319). Previous studies by our own working group on the relationship between emotions and periods point in a similar direction and indicate that poems from early modernism represent fewer and less positive emotions than poems from realism, even though the differences are not huge (Konle et al. 2022).

An alternative to including literary change via period as a categorical variable (realism, modernism) would be to use the year of writing/publication as a numerical variable in order to allow for finer-grained analyses of time series. However, since we have not yet collected the writing/publication dates for all poems, we will use the periods for now.

For categorizing the poems into periods, we use the corpus anthologies and the division into subcorpora described above. The 3,367 texts from the anthologies published between 1859 and 1882 are assigned to the period 'realism' and the 2,882 texts from the anthologies published between 1885 and 1911 to the period 'modernism'. Again, it is important to keep in mind that the period labels are provided by the anthologists, and that their views on what counts as 'realist' or 'modernist' do not always align with today's research.

Gender The gender of the author provides another basic category for differentiating poems. When literary scholars analyze text corpora, gender is regularly brought into view as a possible factor of difference. Hypotheses that assume gender-specific differences also exist with regard to the representation of emotions. For example, one

study argues that until the second half of the twentieth century, sexual desire was represented more frequently, or at least more openly, by male than by female authors (Härle 2007, 118, 135). And an early twentieth-century anthologist named Margarete Huch, to give another example, claims that in her time, "love for man takes up by far the largest space in all women's poetry" (Huch 1911, 12), which might lead to the expectation that female authors are especially likely to represent love, i.e. more likely than men.

We obtained data on the author's gender from the database of the German National Library,[5] by manual research, e.g. in biographical dictionaries, and/or on the basis of the author's first name. Since the multiplicity of possible gender positions was not yet generally recognized during the period under study, and the resources mentioned above identify the corpus authors exclusively as "male" or "female", the feature 'gender' is limited to these two options. We were able to collect data for 78% of the corpus authors (who wrote 85% of the poems). 85% of the identified authors are male, while 15% are female. We cannot say whether the 15% share of female authors is representative, as we have no information on the proportion of women publishing poetry during the time period. If we assume that anthologies can function as an indicator, then we are most likely overestimating the share of female authors due to the inclusion of a female author only anthology (Huch 1911).

Profession   As another feature, we include the profession of the authors in our analysis. Since numerous corpus authors are not only writers, but also, for example, politicians, historians, or philosophers, the feature 'profession' allows to differentiate the authors and their poems in a meaningful way. Certainly, there are not many previous studies that demonstrate a general correlation between profession and emotion in poetry. However, there is at least some research on individual texts or authors that assumes such a connection, and in this respect suggests that a more systematic study of the relationship between profession and emotion might be informative. For example, with respect to Gottfried Benn, it has been argued that there is a link between his profession as a physician and his sober and unemotional style in early poems such as *Kleine Aster* (e.g., Hiebel 2005, 218–221). Moreover, the inclusion of profession allows us to integrate aspects of the author's socioeconomic background in the analysis which would otherwise not be considered at all. Still, including profession as a feature must be understood as an experiment that may or may not prove relevant to the representation of emotions.

To obtain data on the author's profession, we relied on the authority files of the German National Library, the GND, which stores this type of information for each person in its database and which takes into account that a person may have more than one profession at the same time.[6] However, the resulting labels are very diverse: According to the database, the corpus authors practice 346 different professions, with 190 professions being assigned to only one author. It was obvious that the number of classes had to be reduced. Therefore, we divided the professions into 8 groups based on the thematic genres (people working in the field of history, people working in the field of politics, etc.), supplemented by a group for 'other professions' (Table 2). Certainly, using thematic genres to categorize professions is only one of many possibilities, and it may be useful

---

5. See: https://dnb.de/.
6. See: https://www.dnb.de/gnd.

Profession	Examples	Count
Love	–	0
Nature	natural scientist, botanist	18
Poetology	writer, poet	1,497
History	historian, archivist	54
Politics	politician, diplomat	101
Philosophy	philosopher	24
Religion	priest, bishop	87
Culture	actor, painter	251
Other	lawyer, physician	899

Table 2: Number of authors per profession.

	Love	Nature	Poetology	History	Politics	Philosopy	Religion	Culture
Acc.	0.88	0.833	0.632	0.872	-	0.732	0.815	0.470
Std.	0.022	0.024	0.224	0.026	-	0.028	0.059	0.169

Table 3: Quality of genre classification.

to experiment with other categorizations (e.g. socio-economic status, although data availability is an issue in this case) in the future.[7]

That a large number of authors have been assigned to the field of 'poetology' is understandable, since the category includes all those who are designated by the German National Library as 'writers', 'poets', etc., which is true of most corpus authors.

Thematic Genre   Just like period or author gender, genre is a fundamental category for characterizing literary texts. Among other things, the attribution of thematic genres such as love poetry or nature poetry provides basic information about the content of the poems. Furthermore, studies by our own working group have shown that there is a relevant relationship between thematic genre and the representation of emotions in poems (Kröncke et al. 2023). For all these reasons, we include thematic genre as a feature in our analysis.

As described in more detail in Kröncke et al. (2023), we manually annotated 8 thematic genres in 1,412 poems (the themes were love, nature, philosophy, religion, poetology, politics, culture, and history). It was possible to assign exactly one, but also none or several genres to a poem. The annotators reached an agreement of 0.69 (krippendorff's alpha). We continued with training binary classifiers for each thematic genre with the exception of political poetry, for which we had too few annotations. As with the automatic detection of emotions, the classification is based on the gbert-large language model adapted to our corpus.

The performance of the classifiers seems to be sufficient (Table 3), with only cultural poetry and poetological poetry being detected less reliably than the other genres. For

---

7. Furthermore, there is still some work to be done on the data drawn from the German National Library, as samples have shown that they are not always reliable (e.g., the professions 'writer, lawyer, librettist' are assigned to Franz Kafka, https://d-nb.info/gnd/118559230).

this reason, we exclude cultural and poetological poetry (as well as political poetry) from the analysis i.e. treat texts of these three genres in the same way as texts without a genre label.

Form Including formal aspects in the analysis seems instructive, since the term 'form' encompasses basic features of general importance, for poems especially meter and rhyme. Moreover, several studies have indicated that there is a connection not only between form and content (e.g., Carper and Attridge 2003, and, using a computational approach, Šeļa et al. 2021), but also between form and emotion (e.g., Obermeier et al. 2013; Tsur 2017; Haider 2021).

To collect data on formal aspects automatically, we used the Metricalizer tool,[8] which we evaluated with manually annotated poems from our corpus. Three annotators analyzed 100 poems and annotated for each verse, among other things, the number of stresses and the binary distinction of whether the verse is rhymed or unrhymed. Regarding the number of stresses, the annotators reached an agreement of 0.91; for rhyme, an agreement of 0.95. After the initial annotation, the annotators discussed all instances of disagreement and created a consensus annotation that we used as the basis for evaluating the Metricalizer. The tool achieved an accuracy score of 0.91 or an alpha agreement of 0.96 for the number of stresses and an F1 score of 0.97 for the binary distinction 'rhymed/unrhymed'.[9]

Since the performance of the Metricalizer seems good enough, we let the tool analyze all corpus texts and extracted the following features for every poem to be used in our analysis:

- the *proportion of rhymed verses*, since texts with few or no rhymes may represent different emotions than poems that are fully rhymed (Obermeier et al. 2013). According to the Metricalizer, 5% of the corpus poems contain no rhymes at all; 40% include both rhymed and unrhymed verses, while in the remaining 55%, all verses are part of a rhyme;

- the *average verse length* of the poem, measured by the number of characters per line, a feature that is not directly based on the results of the Metricalizer, but is strongly correlated with what the Metricalizer yields as the number of stresses ($r = 0.9$). Research suggests that in certain contexts, longer verses tend to convey a more serious mood, while shorter verses are associated with lighter themes (Šeļa et al. 2021). If this is true, then verse length might also be related to the representation of emotions. In our corpus, the average verse length is 34 characters ($SD = 12$).

In addition to the formal features mentioned above, it would have been instructive to include the verse foot (iambic, trochaic …) as well. However, due to technical difficulties and limitations associated with the Metricalizer, we have to postpone the addition of the verse foot for the time being.

---

8. See: https://metricalizer.de/de/.

9. In addition, we annotated and evaluated the role of the verse in the rhyme structure, using a letter system (a, b, c …) and (for the first two verses of each poem) the verse foot (iambic, trochaic …). Ultimately, however, we did not use these data in our analysis.
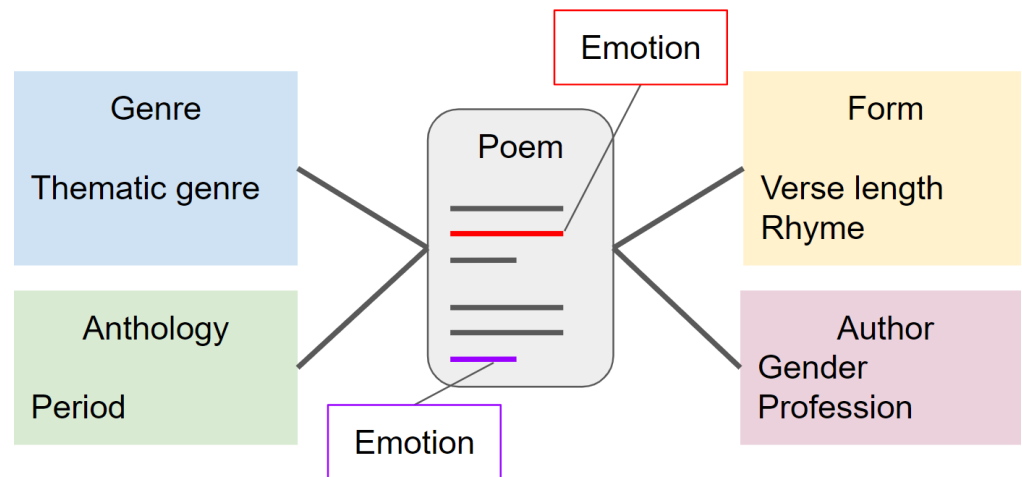
Figure 2: Data model.

4. Methods

4.1 Formal Modeling of the Factors

Our analyses focus on emotions in poems. Both the annotations and the machine learning models for emotion detection operate on spans within poems. As a result, several emotions are attributed to a poem in different, partly overlapping passages of text. The other factors, however, are at the document level (Figure 2). It would facilitate the analysis if all factors were on the same level together with the target variable (emotion). This can be achieved by transforming the data. For example, the emotions could be weighted according to the length of the annotation span. We decided not to use this or any other transformation, as they require assumptions that we do not take for granted. In order to weigh emotions according to the length of their passage, we would have to assume that this actually allows us to make a statement about the relevance of the emotion in the whole poem. It is quite conceivable that an emotion that is only present in the last verse of a poem, for example, could be at the center of the spectrum of emotions. The disadvantage of not transforming, and thus working at the level of emotions rather than poems, is that a poem is modeled as multiple data points. Or, to put it another way, we create an additional layer of inter-correlation in our data as we take multiple 'measurements' from one poem.

In summary, we model individual emotions in poems with the global properties of *thematic genre*, *proportion of rhymed verses*, and *average verse length*. These poems come from anthologies that can be assigned to a *literary period*, and were written by authors of different *genders* whose *professions* we fetched and grouped.

4.2 Bayesian Generalised Hierarchical Linear Model

For the statistical analysis of our dataset, we decided to use a Bayesian hierarchical[10] generalized linear model. The basis for this decision is the nested structure of our data. To determine whether an emotion is overrepresented within a thematic genre, we could simply count how often we find it there and whether the probability is greater than

---

10. Also called: mixed, mixed effects or multi-level.

for poems that do not belong to that group. However, this approach does not take into account the other characteristics of the poems within the genre. If we, for example, assume that they tend not to rhyme and that this favors the occurrence of the same emotion, we cannot say whether our findings are valid or not. The use of a generalized linear model (GLM) allows us to take these contingencies into account, at least to the extent that we control for the variables (factors) present in our data set. This is an improvement, but still not enough.

This will become clearer when we look at an example: Given we observe that the probability of an emotion in religious poetry is about as high as in the rest of the corpus. On further analysis, we find that the probability of the emotion decreases in modern religious poetry and increases in religious poetry from realism. This exciting information escaped us, because we averaged over the genre independently of its literary-historical context. There is the possibility, and literary history gives us reason to believe that this is not unlikely, that genres evolve with the change of literary periods (in our case exclusively with regard to the representation of emotions). To take this possibility into account, we model hierarchy in our GLM. From a purely statistically-motivated point of view without subject-related theoretical assumptions, we would have to assume that this kind of relationship can exist between all factors. This would result in groups such as unrhymed poems by natural scientists or historical poetry with low verse length, all of which must be treated individually. To avoid this scenario and an unnecessarily complex model, we should ask which of the combinations do not have a purely additive relationship in terms of their effect on the distribution of emotions. For example, it is not very plausible to assume that the effect of rhymes on emotions changes or even reverses in different thematic genres. The factor with the greatest plausibility for producing non-additive effects with other factors is the literary period. The genre is also a valid option, but we have decided against using it as a grouping factor, since the greatest variation can be assumed in the interplay with the period and we already cover this by grouping by period itself. The combination of profession and genre also seems plausible, for example when a natural scientist writes a poem about the subject matter of his or her discipline, but our dataset is too small here to allow for certain conclusions anyway.

The hierarchical GLM follows the notation below:[11]

$$h_f \sim Normal(0,1) \qquad [for\ each\ factor\ f]$$

$$c_f \sim Normal(0,1) \qquad [for\ each\ factor\ f]$$

$$s_{f,e} \sim Normal(0,1) \qquad [for\ each\ factor\ f\ and\ period\ e]$$

$$logit(p) = x_f(s_{f,e} + h_f) + c_f$$

$$y \sim Bernoulli(p) \qquad [prob.\ of\ emotion\ is\ present]$$

First, we define slopes h and intercepts c for each factor (gender, rhyme, etc.) individually. Then, a second slope parameter s is added, this time for each factor and each literary period, resulting in three parameters for each factor: global intercept c, global slope

---

11. We use an offset approach instead of classic hyperparameter to improve computational efficiency (Wiecki 2017; Betancourt and Girolami 2015).
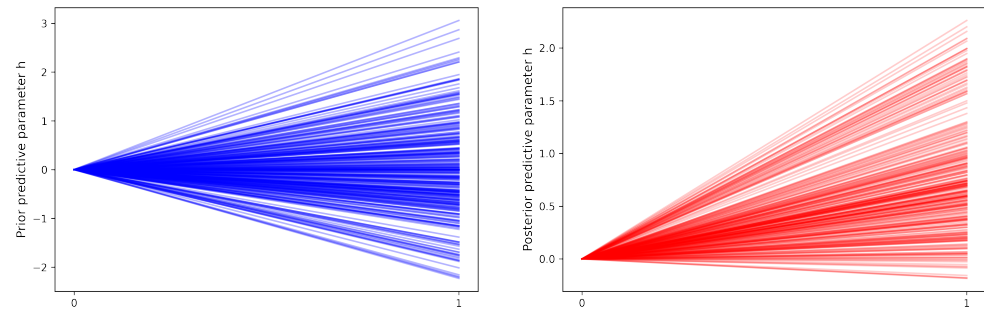
Figure 3: 300 random samples of parameter h_{love} from prior and posterior distributions.

c and grouped slope s. The slopes are summed and multiplied by the actual factor from our data set and added to the global intercept. This way, we model the slope of a factor as being normally distributed around a mean and then corrected by an offset with respect to the literary period. This has the advantage over modeling the effect of, say, modernist religious poetry on emotions directly, because we allow the model to 'share' information between periods through h and c while allowing for variation by s. The limitation of the hierarchisation to slopes (free intercept as well as free slope and free intercept models are also conceivable) follows from results of a preliminary study with a free intercept and free slope model, which shows only negligible variance in intercepts.

Since we use a Bayesian model, we need to set priors according to our expectations. But as can be seen from the notation of the model, we initialize all parameters with so-called flat priors. Flat here means that we do not make any strong assumptions about the effect of factors on emotion or, in other words, we assume that the parameters follow a normal distribution that has its mean at zero and a standard deviation of one and leave the rest to the data.

The left plot in Figure 3 visualizes the initial slope parameter h_{love} for the emotion love in love poems. Each line shows a randomly drawn value from the distribution defined by our priors before the model was fitted to our data set. There are few extreme values indicating a strong negative or positive relationship and more values leaning towards zero. After the model is fitted, the parameter distribution changes. The right plot contains almost only slope values above zero and the highest density of lines in the range between 0.5 and 0.7. The exact mean of this posterior distribution is 0.77; this equals an odds ratio (e^h) of 2.13, stating that if every other factor is fixed, the chance to observe the emotion love in love poems is more than twice as high as in non-love poems.

The mean intercept c_{love} from the posterior distribution is -1.5; this translates to a chance of 0.19 ($e^c/(1+e^c)$) that an emotion span from a non-love poem contains the love emotion. Since we already know the effect of a poem being a love poem, we can calculate the chance of 40% for emotion spans in love poems containing love. Up to this point, we have not yet considered the hierarchical structure of the model. If we want to know whether the influence of the genre love poem changes with the passing of literary periods, we must additionally sample from the posterior distribution of the
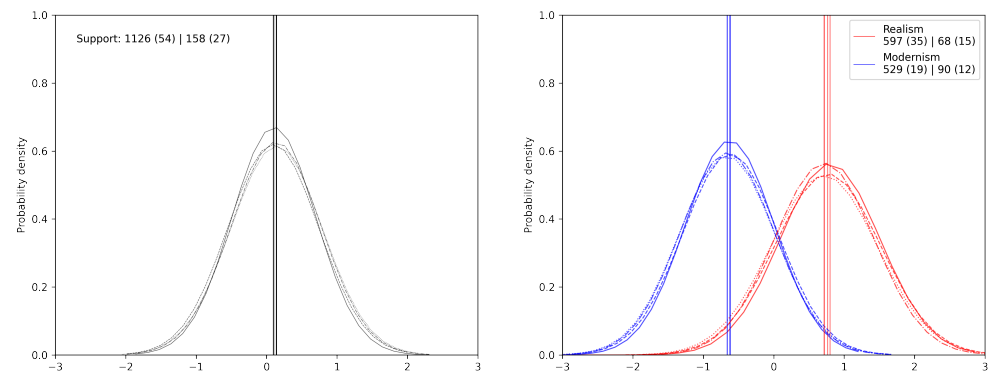
Figure 4: Posterior distribution of $h_{philosopher}$,(on the left side) and $s_{philosopher,realism}$ and $s_{philosopher,modernism}$ (on the right side) for the emotion anger. Support notation: Emotions in group (with anger) | Poems in group (with anger).

parameters $s_{love,realism}$ and $s_{love,modernism}$. Their mean values are 0.252 and 0.249 and their odds ratios 1.286 and 1.282, respectively. Thus, the model indicates a stable effect across both epochs. If we want to know the absolute effect of the genre love poem in realism, we just need to add $s_{love,realism}$ to c_{love} and calculate the odds ratio (2.77).

We estimate the parameters for each emotion individually in separate models. Each model is fitted with 4,000 sampling, 2,000 tuning steps and 4 chains.

5. Results

In the paragraphs that follow, we present four selected findings on the relationships between individual features and emotions, before taking an overarching perspective and analyzing more general aspects. The selected results should be understood as examples that on the one hand illustrate what kind of analyses the model facilitates and on the other hand deepen the understanding of the methodology. Additionally, we selected results that seemed interesting in the light of the assumptions of literary history about these variables.

Figure 4 shows the probability density of slope parameters from posterior distribution of the model fitted to the emotion anger. The bell-shaped curves mark the amount of probability for a certain parameter value. The plot on the left side shows diminishing probability for $h_{philosopher}$ values beyond -2 and 2. The most likely value is located where the curve denotes the highest density. This point (around 0.01) is additionally highlighted with a vertical line for visual convenience. The more pointed the bell is shaped, the more probability is distributed among fewer parameter values. For example, the blue curve for modernist poems in the plot on the right side is flatter than that for realist poems, which means that the parameter $s_{philosopher,modernism}$ can be estimated less confidently. Each plot contains multiple lines; those come from multiple chains or runs of the same model with the same data. If these lines of the same color vary little, we can interpret this as an indicator for model stability.
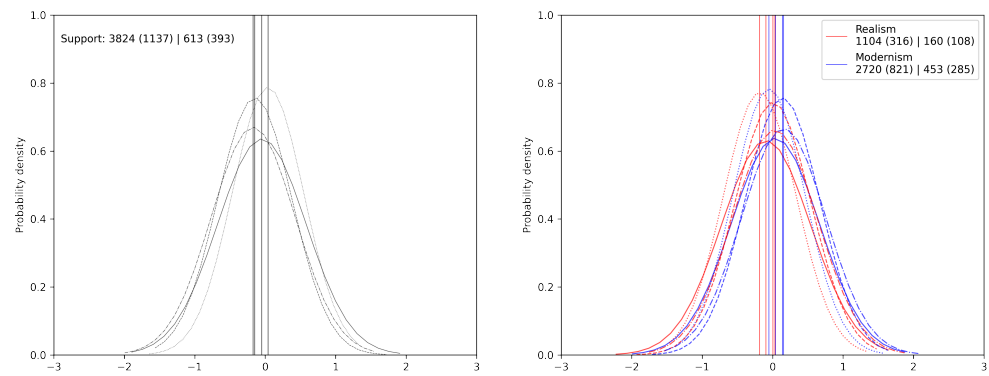
Figure 5: Posterior distribution of h_{gender}, $s_{gender,realism}$ and $s_{gender,modernism}$ for the emotion love. Support notation: Emotions in group (with love) | Poems in group (with love). To belong to the group, a poem needs to be written by a female author.

The support information in the left plot reads as follows: In our dataset, there are 1126 emotion spans (54 contain the emotion anger) from 158 poems written by authors working in the field of philosophy (27 contain the emotion anger at least once). Realism accounts for 597 of these spans and in 35 contain anger, and so on.

The analysis of the two plots allows the following statements: If we know nothing about a poem, except that it is written by a philosopher, we can say that it is just a little more likely (1.1 times), that one of its emotion spans contains anger. But if we get information about the literary period the poem belongs to, that changes drastically. Since the relationship of period grouped slopes s and fixed slope h is additive (subsection 4.2), the odds increase to 2.4 in realism and if it is modernist, they decrease to 0.6. Despite this being a strong effect, this finding should be taken with care due to the low support numbers.

Figure 5 shows that there is no pronounced difference between male and female authors in the probability with which they represent love in their poems. Nor does this change from period to period ($s_{gender,realism}$ and $s_{gender,modernism}$ vary around zero). Note that the support for this finding is stronger than in the case of the "anger/philosopher" example, since our corpus contains a lot more poems and representations of emotions by female authors than by authors working in the field of philosophy. We show this lack of effect because it exemplifies the numerous cases in which we find no substantial relationship between features and emotions. Moreover, this plot shows that the data do not support the assumption that female authors would be more inclined to represent love in their poems (subsection 3.3).

Figure 6 shows the relatively strong relationship of rhymed verses in a poem and the emotion of fear. The mean from the left plot states that, if a poem contains only rhymed verses, the chance to observe the emotion fear is nearly halved (0.56 odds ratio). The right plot shows negative mean values as well, which decreases the odds further for both realism and modernism. This shift is strong in realism, but the difference is small and the variation between chains (varying curves of the same color) indicates a low confidence for this small difference.
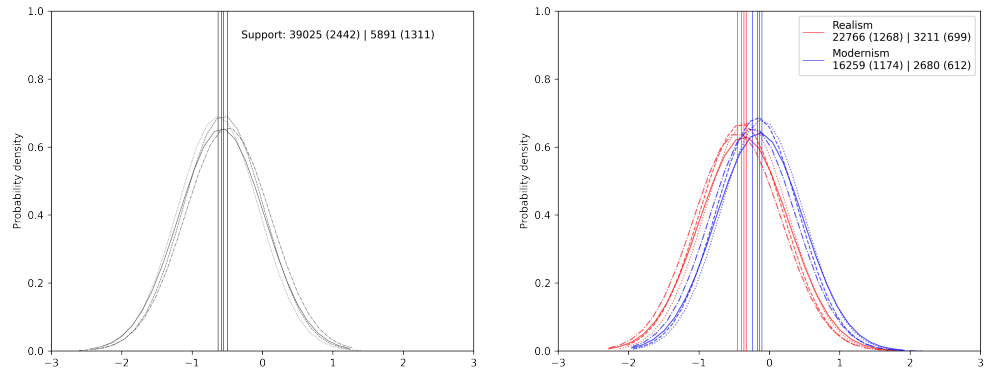
Figure 6: Posterior distribution of h_{rhyme}, $s_{rhyme,realism}$ and $s_{rhyme,modernism}$ for the emotion fear. Support notation: Since proportion of rhyme verses is not a categorical variable, every poem belongs to its group.
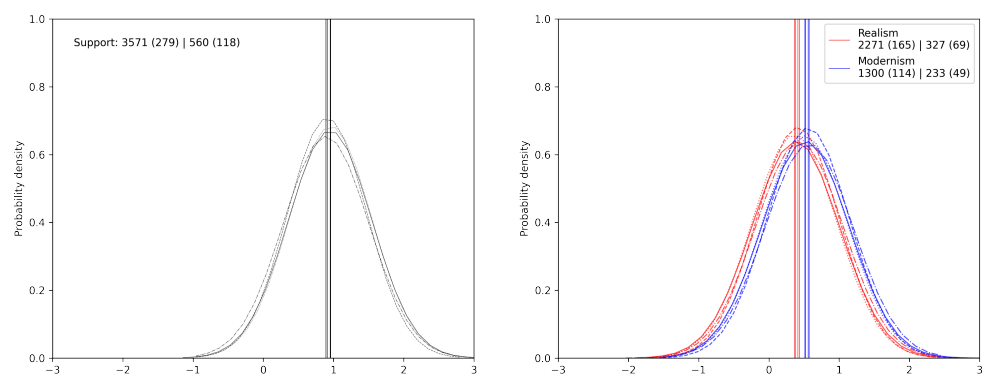

Figure 7: Posterior distribution of $h_{historical}$, $s_{historical,realism}$ and $s_{historical,modernism}$ for the emotion anger.

Feature / Emotion	Love	Sadness	Joy	Anger	Fear	Agitation
Profession	0.16	0.12	0.08	0.33	0.27	0.32
Thematic Genre	0.18	0.19	0.18	0.40	0.42	0.25
Rhyme	0.02	0.24	0.27	0.58	0.56	0.55
Gender	0.08	0.07	0.08	0.08	0.14	0.15
Verse Length	0.07	0.10	0.07	0.03	0.02	0.06

Table 4: Absolute mean value of h from posterior distribution for each factor and emotion.

Our final example (Figure 7) shows that historical poems, e.g., poems about ancient kings or medieval battles, are (2.59 times) more likely to represent anger than poems in other thematic genres. This effect seems to be relatively constant over time, as we see no strong difference between realist and modernist poetry in this regard. With 560 historical poems in our corpus, 118 of which contain the emotion "anger", the support for this relationship is substantial enough to draw the conclusions mentioned.

The result illustrates the relationship between thematic genre and emotion. In this case, a possible explanation might be that historical poetry, as a thematic genre, often depicts (violent) conflicts, such as wars, revolutions, or power struggles, and that these themes make it all the more likely that negative, adversarial emotions, such as anger, are represented (on historical poetry, Detering and Trilcke 2013).

We limit ourselves to the relationships between individual factors and emotions shown to this point and instead provide an overview of the connections in Table 4. Its content provides an overview of the absolute mean values of the h parameter across all emotions and chains (the value which can be obtained from the left plot in Figure 5, Figure 6, Figure 7). Since profession and thematic genre are split into multiple factors, we average their h values.

Although insight into the relationship between individual factors and emotions is informative and, more importantly, useful for forming new hypotheses, We also want to provide an overview of the results of our experiment. However, the simple question of which factor or groups of factors have the greatest influence on emotion raises convoluted methodological and theoretical issues. Let us first make the question more concrete with the following scenario: Given that we find a new poem, which information should we determine first in order to obtain a good guess about its emotions? The mean posterior distribution of the h parameter (or its transformation into odds ratios) allows us to see how much the odds to observe an emotion changes, given the associated factor is actually present (e.g. is a love poem). While the pure rate of change is exciting in its own right, we need to look at the intercept c as well, because quadrupling a small probability may be less informative than doubling a large one. Assume a factor is characterized by a small intercept c and a very large slope h. The information gained would be great if we found that our poem has the properties of this factor; if not, we get not much more certainty about the emotion distribution. That means, we should factor into our consideration how likely our poem is to have a feature before we actually determine it. The only way to get this opportunity, is to look into the distribution of factors in our corpus. By doing so, we assume that the poems outside our corpus follow a similar

Factor group	Profession	Genre	Rhyme	Gender	Verse length
Score	0.16	0.18	0.13	0.12	0.10

Table 5: Score of the factor groups calculated with formula 1.

distribution of factors. Unfortunately, although our corpus is relatively large, its genesis is far from a random drawing from the set of all poems of a literary epoch. However, not including the frequency distribution of the factors in the corpus in the calculation is also not an option, since this would mean assuming an equal distribution of the factors within poetry in general, which is even less likely. The same applies to the distribution of the emotions themselves. A slight tendency against the occurrence of an otherwise very common emotion (e.g. love) can tell us more than a strong correlation with an emotion that is very rare anyway. By taking these considerations into account, we propose the following equation to capture this information:

$$\sum_i (|p(f)h_{if} + c_{if}|)p(em^i) \tag{1}$$

This equation represents the sum of the absolute value of slope h, normalized by the probability that a factor f is present in a poem and intercept c for every emotion i. Afterwards, the result is weighted by the probability of observing an emotion in our corpus $p(em^i)$. The values for $p(em^i)$ are depicted in Figure 1. The values for the factor groups Job and Genre are averaged analogously to the values in Table 4. The result is shown in Table 5, which states that we should first determine a poem's genre to gain the most information about its potential emotions. The profession of the author comes second and after that rhyme, gender and verse length.

6. Discussion

Are the results surprising? What strength of association between features and emotions might we expect according to previous research in (traditional) literary studies? As argued in subsection 3.3, for virtually all features (perhaps except for profession), we find at least some statements claiming that the features are relevant for differentiating literature in general and/or the representation of emotions in particular (while hardly anyone claims that any of the features have no relevance at all). Our results broadly confirm these research statements in that for each feature, at least some association with one or more emotions was found. This is true as well for the extra-literary features 'profession' and 'gender', showing that the representation of emotions in poetry is, at least in part, related to factors outside the literary text itself. On a more specific level, however, some qualifications are in order, since we did not find clear associations between *all* features and *all* individual emotions. For example, Figure 5 has shown that there is no strong connection between gender and the representation of love, although some statements, such as the one by Margarete Huch cited in subsection 3.3, might suggest the opposite.

Beyond that, explicit hypotheses that *rank* or *compare* the relevance of multiple features (e.g., "For literary phenomenon X, genre is more important than gender"), are extremely

rare in literary studies. However, to have at least some point of reference, we can take a look at the *practice* of literary studies and analyze whether our results correlate with it. To this end, we surveyed literary histories (McInnes and Plumpe 1996, Fähnders 1998, Sprengel 1998, Mix 2000, Sprengel 2004, Aust 2006, Ajouri 2009, Stockinger 2010, Beutin et al. 2013, Willems 2014, Willems 2015, Sprengel 2020) as well as all results of a search for the keyword "emotion" in a major bibliography of German literary studies.[12] Mainly on the basis of prefaces and outlines (literary histories) or titles (research on emotion according to the bibliography), we assessed which of the selected features literary scholars most prominently focus on, use for delimiting their objects of study and/or claim to be relevant. It is important to keep in mind that, in addition to assumed relevance, numerous other (reasonable) factors are likely to influence which features literary scholars highlight and/or use to organize or name their studies. It is for this reason, among others, that our purpose is not to 'evaluate' or 'refute' the practice of literary scholars; rather, we use their work as a reference point to frame our findings.

Of all the features analyzed in our study, literary scholars are by far the most likely to focus on (thematic) *genre*. Literary histories are regularly organized by genre, and research is quite often concerned with how emotions are represented in particular (thematic) subgenres. Consistent with this, we observe that, according to our model, the representation of emotions is most strongly associated with genre (Table 5). The next most common focus of researchers seems to be *gender*. At least some literary histories include individual chapters on gender and there are multiple research contributions that investigate the relationship between gender and the representation of emotions, even though they do not always focus on the gender *of the author*. In any case, gender is a more prominent theme in the surveyed contributions than the other three features 'profession', 'rhyme', and 'verse length'. However, our results show that the relationship between author gender and the represented emotions, while present, is comparatively limited. The association of gender with emotion is about as strong as that of rhyme and verse length, and weaker than that of profession. As already indicated, the features '*profession*', '*rhyme*', and '*verse length*' receive much less attention in literary histories and research on emotions in literature. Although these features (especially rhyme) are occasionally mentioned, e.g. in the context of single-text analyses, there are hardly any studies on literary emotions or chapters in the surveyed literary histories that focus *primarily* on one of these factors. Regarding our results, it is fitting that verse length is relatively weakly related to the representation of emotion. But the comparatively strong association of rhyme, and especially profession, is quite surprising. We certainly did not expect the connection of emotion with rhyme to be as strong as that with thematic genre, and the connection with profession to be as strong as that with gender.

We explained in our introduction that we don't assume a causal relationship between these factors and the emotions in the poems, and that we don't assume that these factors have the same kind of relationship to the emotions. But it is important to understand the limitations of our findings. If we look at the relationship between rhyme and agitation, for example, we see a rather high value of 0.55 in Table 4. This needs to be taken with a grain of salt, because this analysis is based on the automatically labeled verses, and the F1 score for detecting agitation is quite low, only 0.62. In addition, the distribution

---

12. See: https://www.bdsl-online.de.

of support for different values of rhyme is relevant. If we divide this variable into four groups, all lines are rhymed, none, more than 50% of the lines (but less than 100%), less than 50% (but more than none), we see that it is quite unevenly distributed. And even if we accept this result with all these caveats, we are still only talking about a connection where the nature of the connection and the reasons for its existence are unknown. Perhaps, from the contemporaries' point of view, the disruption of the rhyme or its complete suspension was perceived as a textual strategy, which, in the context of an ideal congruence between content and form, was seen as a good option for expressing something like agitation. Or the expression of agitation is often linked to the expression of abstract themes, and the complexity of expressing them precisely is particularly difficult given the limits of rhyme. Or the preference for unrhymed metrical forms of Greek-Roman antiquity is particularly high in relation to the expression of agitation. All of these, and a combination of them, could be causal reasons for this relationship.

These considerations also point in the direction of our future work. We will continue to identify interesting ideas about the connections between external factors and aspects of literary texts in literary studies and try to systematically integrate them into a more formal model. We will try to find and digitize more data about these external factors. And we will try to provide a richer representation of the textual features we include in our model, for example include more features like style, rhetorical devices like figurative speech etc. We are also considering experimenting with other "main phenomena" to focus on (currently emotions), such as topics. Finally, we want to explore the application of models that offer not just descriptions of relationships, but causal explanations – understood in the broadest possible sense.

7. Data Availability

Data can be found here: `10.5281/zenodo.10148449`.

8. Software Availability

Software can be found here: `10.5281/zenodo.10148449`.

9. Acknowledgements

10. Author Contributions

Leonard Konle: Software, Formal Analysis, Visualization, Writing – original draft

Merten Kröncke: Data Curation, Methodology, Writing – original draft

Simone Winko: Data Curation, Conceptualization, Supervision, Writing – original draft, Funding acquisition

References

Ajouri, Philip (2009). *Literatur um 1900: Naturalismus, Fin de Siècle, Expressionismus*. Akademie Studienbücher Literaturwissenschaft. Akademie Verlag. doi:10.1524/9783050049533.

Andreotti, Mario (2014). *Die Struktur der modernen Literatur: Neue Wege in die Textinterpretation: Erzählprosa und Lyrik*. 5th ed. Haupt. 10.36198/9783838540771.

Aust, Hugo (2006). *Realismus*. Lehrbuch Germanistik. J. B. Metzler.

Barrett, Louise and Robin Dunbar, eds. (2007). *Oxford Handbook of Evolutionary Psychology*. 1st ed. Oxford University Press. 10.1093/oxfordhb/9780198568308.001.0001.

Betancourt, Michael and Mark Girolami (2015). "Hamiltonian Monte Carlo for Hierarchical Models". In: *Current Trends in Bayesian Methodology with Applications*. Chapman and Hall/CRC.

Beutin, Wolfgang, Matthias Beilein, Klaus Ehlert, Wolfgang Emmerich, Christine Kanz, Bernd Lutz, Volker Meid, Michael Opitz, Carola Opitz-Wiemers, Ralf Schnell, Peter Stein, and Inge Stephan (2013). *Deutsche Literaturgeschichte*. 8th ed. J.B. Metzler.

Borkowski, Jan (2015). *Literatur und Kontext. Untersuchungen zum Text-Kontext-Problem aus textwissenschaftlicher Sicht*. Mentis.

Carper, Thomas and Derek Attridge (2003). *Meter and Meaning: an Introduction to Rhythm in Poetry*. Routledge. 10.4324/9781003060307.

Chan, Branden, Stefan Schweter, and Timo Möller (2023). "German's Next Language Model". In: *arXiv preprint*. 10.48550/arXiv.2010.10906.

Detering, Heinrich and Peer Trilcke, eds. (2013). *Geschichtslyrik. Ein Kompendium*. 2 vols. Wallstein.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *arXiv preprint*. 10.48550/arXiv.1810.04805.

Ekman, Paul (1992). "An Argument for Basic Emotions". In: *Cognition and Emotion* 6 (3), 169–200. 10.1080/02699939208411068.

— (1999). "Basic Emotions". In: *Handbook of Cognition and Emotion*. Ed. by Tim Dalgleish and Mick J. Power. John Wiley & Sons, Ltd, 45–60. 10.1002/0470013494.ch3.

Engel, Manfred (2018). "Kontexte und Kontextrelevanzen in der Literaturwissenschaft". In: *KulturPoetik* 18 (1), 71–89. https://www.jstor.org/stable/26422521 (visited on 12/11/2023).

Fähnders, Walter (1998). *Avantgarde und Moderne 1890-1933*. Lehrbuch Germanistik. J. B. Metzler.

Gittel, Benjamin (2016). "Lässt sich literarischer Wandel erklären? Struktur, Gültigkeitsbedingungen und Reichweite verschiedener Erklärungstypen in der Literaturgeschichtsschreibung". In: *Journal of Literary Theory* 10 (2), 303–344. 10.1515/jlt-2016-0012.

Gururangan, Suchin, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A. Smith (2020). "Don't Stop Pretraining: Adapt Language Models to Domains and Tasks". In: *Proceedings of the 58th Annual Meeting of the Association*

for Computational Linguistics. Association for Computational Linguistics, 8342–8360. `10.18653/v1/2020.acl-main.740`.

Haider, Thomas (2021). "Metrical Tagging in the Wild: Building and Annotating Poetry Corpora with Rhythmic Features". In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Association for Computational Linguistics, 3715–3725. `10.18653/v1/2021.eacl-main.325`.

Härle, Gerhard (2007). *Lyrik, Liebe, Leidenschaft: Streifzug durch die Liebeslyrik von Sappho bis Sarah Kirsch*. Vandenhoeck und Ruprecht.

Hiebel, Hans H. (2005). *Das Spektrum der modernen Poesie. 1: 1900 - 1945*. Königshausen & Neumann.

Huch, Margarete, ed. (1911). *Frauenlyrik der Gegenwart. Eine Anthologie*. Fritz Eckardt.

Kalliney, Peter (2019). "Introduction: Literary History after the Nation?" In: *Modern Language Quarterly* 80 (4), 359–377. `10.1215/00267929-7777767`.

King, Martina and Jesko Reiling (2014). "Das Text-Kontext-Problem in der literaturwissenschaftlichen Praxis: Zugänge und Perspektiven". In: *Journal of Literary Theory* 8 (1), 2–30. `10.1515/jlt-2014-0001`.

Konle, Leonard and Fotis Jannidis (2020). "Domain and Task Adaptive Pretraining for Language Models". In: Computational Humanities Research. `http://ceur-ws.org/Vol-2723/short33.pdf` (visited on 11/29/2023).

Konle, Leonard, Fotis Jannidis, Merten Kröncke, and Simone Winko (2022). "Emotions and Literary Periods". In: *DH Conference Abstracts*. Digital Humanities.

Kröncke, Merten, Fotis Jannidis, Leonard Konle, and Simone Winko (2022). "Annotationsrichtlinien Emotionsmarker und Emotionen". In: *Zenodo*. `10.5281/zenodo.6020616`.

Kröncke, Merten, Leonard Konle, Fotis Jannidis, and Simone Winko (2023). "Gattungen und Emotionen in der Lyrik des Realismus und der frühen Moderne". In: *Zenodo*. `10.5281/zenodo.7715402`.

Ladegaard, Jakob and Jakob Gaardbo Nielsen (2019). "Introduction: the Question of Context". In: *Context in Literary and Cultural Studies*. Comparative Literature and Culture. UCL Press, 1–13.

Langer, Lars, Manuel Burghardt, Roland Borgards, Katrin Böhning-Gaese, Ralf Seppelt, and Christian Wirth (2021). "The Rise and Fall of Biodiversity in Literature: A Comprehensive Quantification of Historical Changes in the Use of Vernacular Labels for Biological Taxa in Western Creative Literature". In: *People and Nature* 3 (5), 1093–1109. `10.1002/pan3.10256`.

Mathet, Yann, Antoine Widlöcher, and Jean-Philippe Métivier (2015). "The Unified and Holistic Method Gamma (γ) for Inter-Annotator Agreement Measure and Alignment". In: *Computational Linguistics* 41 (3), 437–479. `10.1162/COLI_a_00227`.

McElreath, Richard (2020). *Statistical Rethinking: a Bayesian Course with Examples in R and Stan*. Second edition. Texts in Statistical Science Series. CRC Press, Taylor & Francis Group.

McInnes, Edward and Gerhard Plumpe, eds. (1996). *Bürgerlicher Realismus und Gründerzeit, 1848–1890*. Hansers Sozialgeschichte der deutschen Literatur vom 16. Jahrhundert bis zur Gegenwart 6. Carl Hanser.

Mix, York-Gothart, ed. (2000). *Naturalismus, Fin de siècle, Expressionismus (1890-1918)*. Hanser.

Obermeier, Christian, Winfried Menninghaus, Martin von Koppenfels, Tim Raettig, Maren Schmidt-Kassow, Sascha Otterbein, and Sonja A. Kotz (2013). "Aesthetic and Emotional Effects of Meter and Rhyme in Poetry". In: *Frontiers in Psychology* 4. `10.3389/fpsyg.2013.00010`.

Piper, Andrew (2022). "Biodiversity is Not Declining in Fiction". In: *Journal of Cultural Analytics* 7 (3). `10.22148/001c.38739`.

Plutchik, Robert (1980a). "A General Psychoevolutionary Theory of Emotion". In: *Theories of Emotion*. Ed. by Robert Plutchik and Henry Kellerman. Academic Press, 3–33. `10.1016/B978-0-12-558701-3.50007-7`.

— (1980b). "A Psychoevolutionary Theory of Emotions". In: *Social Science Information* 21 (4), 529–553. `10.1177/053901882021004003`.

— (2001). "The Nature of Emotions". In: *American Scientist* 89 (4), 344–350. `http://www.jstor.org/stable/27857503` (visited on 12/11/2023).

Šeļa, Artjoms, Petr Plecháč, and Alie Lassche (2021). "Semantics of European Poetry is Shaped by Conservative Forces: The Relationship between Poetic Meter and Meaning in Accentual-Syllabic Verse". In: *arXiv preprint*. `10.48550/arXiv.2109.07148`.

Shaver, Philipp, Judith Schwartz, Donald Kirson, and Cary O'Connor (1987). "Emotion Knowledge: Further Exploration of a Prototype Approach". In: *Journal of Personality and Social Psychology* 52 (6), 1061–1086. `10.1037//0022-3514.52.6.1061`.

Sobchuk, Oleg and Peeter Tinits (2020). "Cultural Attraction in Film Evolution: the Case of Anachronies". In: *Journal of Cognition and Culture* 20 (3), 218–237. `10.1163/15685373-12340082`.

Sprengel, Peter (1998). *Geschichte der deutschsprachigen Literatur 1870-1900. Von der Reichsgründung bis zur Jahrhundertwende*. C. H. Beck.

— (2004). *Geschichte der deutschsprachigen Literatur 1900–1918. Von der Jahrhundertwende bis zum Ende des Ersten Weltkriegs*. Vol. 2. Geschichte der deutschen Literatur von den Anfängen bis zur Gegenwart 9. C. H. Beck.

— (2020). *Geschichte der deutschsprachigen Literatur 1830–1870. Vormärz-Nachmärz*. Geschichte der deutschen Literatur von den Anfängen bis zur Gegenwart 8. C.H. Beck.

Stockinger, Claudia (2010). *Das 19. Jahrhundert. Zeitalter des Realismus*. Akademie Verlag. `doi:10.1524/9783050052908`.

Thomsen, Mads Rosendahl (2019). "From Data to Actual Context". In: *Context in Literary and Cultural Studies*. Comparative Literature and Culture. UCL Press, 190–209.

Tsur, Reuven (2017). "Metre, Rhythm and Emotion in Poetry. A Cognitive Approach". In: *Studia Metrica et Poetica* 4 (1), 7–40. `10.12697/smp.2017.4.1.01`.

Underwood, Ted, Kevin Kiley, Wenyi Shang, and Stephen Vaisey (2022). "Cohort Succession Explains Most Change in Literary Culture". In: *Sociological Science* 9, 184–205. `10.15195/v9.a8`.

Wiecki, Thomas (2017). *Why Hierarchical Models are Awesome, Tricky, and Bayesian*. While My MCMC Gently Samples. `https://twiecki.io/blog/2017/02/08/bayesian-hierchical-non-centered/` (visited on 11/29/2023).

Willems, Gottfried (2014). *Geschichte der deutschen Literatur. Band 4: Vormärz und Realismus*. 1st ed. Böhlau Verlag Köln. `10.36198/9783838538747`.

— (2015). *Geschichte der deutschen Literatur. Band 5: Moderne*. 1st ed. Böhlau Verlag Köln. `10.36198/9783838542492`.

Winko, Simone, Leonard Konle, Merten Kröncke, and Fotis Jannidis (2022). *Korpusbeschreibung der Lyrikanthologien 1850-1910*. `10.5281/zenodo.6053972`.