

ARTICLE

Examining the representation of landscape and its emotional value in German-Swiss fiction between 1840 and 1940

Giulia Grisot¹, Berenike Herrmann¹¹ Bielefeld University

Keywords: fictional space, German-Swiss fictional prose, quantitative analysis, affective literary studies, urban and rural space

<https://doi.org/10.22148/001c.84475>

Journal of Cultural Analytics

Vol. 8, Issue 1, 2023

This paper presents a quantitative analysis of the representation and affective encoding of fictional space in a corpus of 125 Swiss literary prose texts of the 19th and early 20th Century written in German, offering a contribution to both spatial and affective literary studies.

Motivated by questions about the iconic dichotomy between ‘urban’ and ‘rural/natural’ space in literary works (Sengle; Fournier; Nell and Weiland) – and in Swiss literature around 1900 in particular (Rehm) – we use computational methods to detect and examine how different types of space are distributed and affectively encoded in German-Swiss literature. Taking into account the complexity of cultural perceptions and representations of space across history, we examine the presence of ‘urban’ and ‘rural/natural’ fictional spaces and their potential role in constructing a ‘Swiss’ national literature (Böhler; Zimmer), and their affective encoding.

In order to do this, we first compiled a comprehensive dictionary of named and non-named spatial entities in the broad spatial categories RURAL and URBAN, and examined the presence of sentiment and emotions (valence and discrete emotions) and their ‘strength’ (arousal) in relation to these. We used current state-of-the-art sentiment lexicons for German available to the digital humanities community. Similarly to Heuser et al., we mapped the spatial entities and the sentiment lexicons onto our corpus, and focused on spans of +/-50 words around the detected entities, in order to examine the specific sentiment and emotions related to space.

In an exploratory analysis, we offer here a first-time data-driven perspective on rural and urban fictional space, incorporating the dimension of affective encoding of space systematically.

Introduction¹

The representation of fictional space has received increased academic attention over the last decades, especially in light of the so called ‘spatial turn’ in the humanities. Franco Moretti is generally regarded as the scholar who marked the beginning of a new era in literary geography, with maps being used as tools to examine and interpret literary texts and their cultural

¹ The data used for this paper and some additional resources are available on Zenodo (European Organization For Nuclear Research and OpenAIRE) at the following link: <https://doi.org/10.5281/zenodo.7024595>

significance (Moretti, *Atlas of the European Novel: 1800-1900*). His approach, as that of most of the numerous scholars following his lead (Piatti, Bär, Reuschel, and Hurni; Heuser et al.; Cooper et al.; Taylor et al.; Evans and Wilkens, to name only a few), focuses however mainly on geographical and topographical representations, looking at *where* and *when* certain literatures developed, how these are connected, and how these are temporally characterised – culturally and linguistically (Moretti, *Graphs, Maps, Trees: Abstract Models for Literary History*). Even when stretching the boundaries of geography to include non-specified locations (see Piatti, Bär, Reuschel, Hurni, et al.), the main attempt of these explorations has appeared to produce *visualisations* of such patterns and spaces, either by applying and overlapping them onto real world maps, or by showing the relations and interrelations between objects (such as novels, spatial objects, or characters) typically using network visualisations (see for example Bushell).

Far less attention has been dedicated to those elements of space representation that are *not* geolocations, and that therefore cannot be located on a map. It is this kind of spatial terms that often effectively create the so called ‘storyworld’ (Tally): whether a story is set in a pirate vessel or on a remote island – and whether or not these may also be locatable somewhere in the real world – the spatial imagery will often be evoked by simple spatial terms such as ‘deck’ and ‘bridge’, ‘beach’ and ‘cave’, and the objects and architectural parts that make them ‘tangible’.

We aim here to fill this gap, looking at geographic, named entities (sometimes referred to as ‘landforms’) as well as at non-named ones, i.e. terms such as ‘mountain’ or ‘alley’, which we will call ‘spatial terms’ – thus analysing the representation of space in our corpus both on a geographical level and on the level of the storyworld. We approach space cultural-semiotically as ‘landscape,’ i.e. as an ‘area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors’ (Council of Europe 2). Considering the idealization, or even mythification, of rural and natural landscapes – especially the Alps – and the role these have played for a Swiss national literature (Gsteiger), we hypothesise for our corpus (1843-1940) that fictionally represented natural and rural space are (a) more prevalent than urban space and (b) likely to show a more positive and also a emotionally ‘richer’ encoding than urban space.

The present study focuses on *space* and *affect* in German-Swiss fictional prose written between 1843 and 1940 – a historical period during which the difference between urban and rural landscape in the European and German-speaking literary system robustly emerged (Rehm) – providing a data-driven approach to observe how space is encoded affectively.

Given the complex nature of the cultural and social construction of landscape, however, we are also keen to observe whether specific perspectives would contribute to contrasting results. In particular (1) realist perspectives

that describe rural landscapes in critical/ambivalent ways (Hawkins; Cosgrove), including incarnations of the popular genre of ‘Dorfgeschichte’ (village tale), and (2) a potential role of the ‘sublime’ in alpine contexts, which incorporates strong negative emotions alongside strong positive ones (Denning; Scaramellini; Donaldson et al.).

Natural/Rural landscape as space of yearning and identification?

In the 19th Century, the rapid industrialization across and beyond Europe intensified rural-urban migration. While in other European literatures the modern big city emerged with tales of Paris and London, in Switzerland big city development and corresponding literary accounts lag behind by decades. In German-speaking Switzerland, not only Rousseau's *dictum* that only the natural is wholesome, while man-made things are degenerate (Rousseau 1) but also the works by von Muralt and others preceded a literary mass production evoking rural *Sehnsuchtsorte* (‘places of yearning’) (Piatti and Streifeneder). This construction, initiated by Albrecht von Haller's *The Alps* (Haller), Salomon Geßner's *Idylls* (Geßner), but also Schiller's *Wilhelm Tell* (Schiller), provided an extremely successful version of the active myth revolving around an alliance between the ‘original’ Swiss people and the alpine landscape. An exemplary case for landscape-anchored narratives of what it meant to be ‘Swiss’ (Rusterholz and Solbach) is represented by *Heidi* (Spyri), where the natural beauty of the mountains is put in direct contrast to the city, a place of confinement and restrictions.

In our corpus, however, ‘rural’ literary genres such as *Bauern*-, *Dorf*-, and *Bergroman* (respectively farmer, village and mountain novels) cover a relatively variable spectrum of aesthetic, poetic, and ideological positions. Many farmers’ and village stories belong in a socially-conscious realistic tradition led by Jeremias Gotthelf and Gottfried Keller, in which *both* rural and urban landscapes are depicted in an ambivalent, ironic, and partially critical way. Through literary history, alpine spaces have been depicted on a spectrum across the ‘horrible’ (locus horribilis), ‘sublime’, ‘picturesque’, all the way to the cliché (Lafond-Kettlitz). The dichotomous schema ‘urban/rural’ in Swiss narratives of the 19th/early 20th Century appears thus both obvious and problematic. What is more, Swiss prose fiction of this period does not represent exclusively ‘Swiss’ settings, but portrays other existing and imaginary countries,² as well as very different social contexts, focusing on the one hand on the *Volk* (people), *Älpler* (‘people of the Alps’) and *Bauern* (‘farmers’), and on the other hand on the (*Bildungs*-)*Bürger*, the *bohème*, and the *nobility*.³

2 We include geographical locations of countries surrounding Switzerland (the current territories of Italy, France, Austria and Germany). This allows gauging how ‘spatially self-referential’ the texts in our corpus are.

3 A typical German-Swiss type of the village story is the double encoding for popular and bourgeois audiences, see Tschopp.

Taking into account these characteristics, we investigate here how prominently the ‘rural/natural’ is featured in contrast to the ‘urban’, and how each is affectively encoded. To this end, we present our approach as *proof of concept*, putting to use existing digital resources in German. The corpus used for our analysis was created by gathering digitised texts, and we aimed at providing a collection of German-Swiss works as diverse as possible with respect to the literary forms described above, their narrative structures and their audiences. It comprises low-brow as well as high-brow literary texts (as well as texts that sit between these extremes), and displays a wide array of spatial settings such as cities, towns, villages, and natural-secluded areas. This collection thus allows for an assessment of the distribution of types of spatial terms and their respective affective encoding. Further philological and historical research will thereby be enabled.

Affect and fictional texts

The emotions and sentiment related to fictional space constitute a long-standing interest of literary and cultural studies (Lehnert; Schumacher). The affective turn in literary studies in the 1990s put a spotlight on the close relationship between sentiment, emotions and fictional texts, as these have long been means of exploring subjective, yet shared, human expression and representation of emotions. Both the textual craft of emotion and the corresponding empathy evoked in readers (Oatley and P. N. Johnson-Laird; Hogan) have received scholarly attention, and current cognitive psychology suggests that almost anything deriving from human consciousness is in some way emotionally grounded (Russell, “Emotion, Core Affect, and Psychological Construction”).

The investigation of affect in relation to fictional texts can refer to the emotions and sentiments *perceived in the reader* (Jacobs et al.; Menninghaus et al.; Schindler et al.), but also to the portrayal of emotion and sentiment *in the text* and the *symbolically encoded* fictional world (Hillebrandt; Winko). This latter approach, which we pursue in the current study, does not investigate how a reader would respond, but focuses on the ‘emotion potential’ encoded in the text (Herrmann and Lüdtke).

Text-based approaches typically classify texts based on the presence of unambiguous ‘affect words’, i.e. words that can be associated to a positive/negative value (called valence or polarity) and or to a specific discrete emotion (joy, sadness etc.). While this approach is relatively simple and does not require enormous computational power, it must be considered that its results largely depend on the resources used to determine the valence and/or the emotions associated with words – in digital humanities, the current standard is the so called ‘sentiment lexicons’.

In the following sections, we will introduce how we applied computational text-based sentiment analysis methodology to explore the emotional and affective encoding of different types of fictional space, comparing the performance of the existing sentiment lexicons for German.

Computational sentiment and emotion analysis

Sentiment analysis (SA) and emotion detection have become an approachable methodology not only for scholars with a background in information technology and natural language processing (Pang and Lee), but also for those in the humanities (Y.-S. G. Kim et al.; Eder et al.; Scherer; Schindler et al.). A growing body of SA studies has determined sentiment scores for whole literary texts or computed affective variation throughout a narrative (E. Kim et al.; Reagan et al.; Jacobs; Mohammad, *Sentiment Analysis of Mail and Books*; Klinger et al.; Zehe et al.). Under investigation are different types of texts (Kakkonen and Kakkonen; Jockers and Underwood) and genres such as fairy tales (Alm and Sproat; Mohammad, “From Once upon a Time to Happily Ever after: Tracking Emotions in Mail and Books”), and historic plays (Mohammad, “From Once upon a Time to Happily Ever after: Tracking Emotions in Mail and Books”; Schmidt et al.). Results are however often not suited for generalisations: on the one hand, because of the lack of standard procedures recognised by the academic community, and on the other, because of the idiosyncratic characteristics of corpora representing particular literary domains, as well as because of the specificity of sentiment and emotions lexicons, often built to address specific research questions.

As for many other complex phenomena, literary studies have imported theories of *emotions* and *sentiment* from other disciplines, in particular cognitive psychology, evolutionary biology and philosophy. These differ in their conceptualisation of such phenomena as well as in their application of these, and it is therefore not surprising that several different concepts of sentiments and emotion are used in literary studies.

We conceptualise here textual *sentiment* as the underlying feeling, attitude or evaluation associated with a proposition that can be analysed in terms of *valence* and *arousal* (Liu). The term valence reflects the sentiment orientation, or polarity, ranging from pleasant to neutral to unpleasant. One can think of it as an indication of ‘positivity,’ i.e. whether the general sentiment of the text is negative, neutral or positive. Sentiment can also be evaluated in terms of arousal, which is defined as the degree to which a particular stimulus ‘activates’ the experiencer for action (Russell, “A Circumplex Model of Affect.”). Arousal ranges from high (or active) to low (or passive), and can be substantially understood as the level of intensity of the sentiment expressed.

The concept of *emotion* is somewhat more complex, and it has been object of numerous debates, especially with regards of *what is perceived and defined* as ‘emotion’ and *what can be empirically observed* (Scherer). While the open questions on this topic are certainly worth further research, we will treat ‘emotions’ here as understood by Ortony and Turner and Clore and Ortony, i.e. as ‘affective reactions made specific by the different social and psychological situations in which they arise. [...] their identity depends on the context. As a result, emotions can be as varied in nature and number as the situations they represent [...] Emotions, then, are simply affective states that are about something. (Fox et al.)’

According to several psychological models of emotions, humans are thought to have a set of so-called ‘basic emotions,’ particular affective states that are assumed to be cross-culturally recognizable. Basic emotions are described as ‘discrete’ because they are believed to be idiosyncratic and distinguishable by an individual’s facial expression and/or biological processes. Such ‘basic’ emotions can combine to form more complex ones to cover the whole emotional spectrum. Some of the existing theoretical models for basic emotions are summarised in [Table 1](#) below:

Table 1. Basic emotions taxonomies.

Source	Basic emotions
Arnold	Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness
Ekman et al.	Anger, disgust, fear, joy, sadness, surprise
Gray and MacNaughton	Anxiety, joy, rage, terror
Izard	Anger, contempt, disgust, distress, fear guilt, interest, joy, shame, surprise
James	Fear, grief, love, rage
McDougall	Anger, disgust, elation, fear, subjection, tender-emotion, wonder
Mowrer and N.	Pain, pleasure
Oatley and P. N. Johnson-Laird	Anger, disgust, anxiety, happiness, sadness
Panksepp	Expectancy, fear, rage, panic
Parrott	Anger, fear, joy, love, sadness, surprise
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise
Tomkins	Anger, interest, contempt, disgust, distress, fear, joy, shame, surprise
Watson	Fear, love, rage
Weiner and Graham	Happiness, sadness

Sentiment and emotion lexicons

As readers, it is generally possible to intuitively understand the sentiment and the emotions conveyed by a text, thanks to linguistic devices as well as to co-textual and contextual information, including individual or shared socio-cultural experiences.

What we are interested in examining in this paper is how these values are *encoded* in the text, rather than how readers respond to them. In order to express emotions and sentiment in writing, several linguistic strategies can

be used, such as grammatical and lexical expressions, intensifiers, superlatives, or special punctuation (Mohammad, “Imagisaurus: An Interactive Visualizer of Valence and Emotion in the Roget’s Thesaurus”). The sentiment analysis approach used in the present paper focuses on the detection of so-called *emotion* or *mood words*, i.e. words that are understood as possessing an emotional and/or sentiment value. This approach makes use of a definite list of words – ideally compiled by means of a theory-driven and controlled procedure – associated with one or more values, which typically reflect whether a word embeds sentiment and/or emotions, and how intensely. Within literary studies, the theoretical frameworks most commonly used for the investigation of sentiment and emotions are respectively Russell’s two-dimensional sentiment model (“A circumplex model of affect”) – with *valence* as the degree of positivity and negativity expressed, and *arousal* representing the strength of the sentiment expressed – and Ekman et al.’s atlas of emotions (Ekman et al.) – featuring six basic emotions: *anger*, *disgust*, *fear*, *joy*, *sadness*, *surprise*. These word lists can be then mapped onto a corpus, providing a measure of sentiment and emotions. This approach is generally referred to as *lexicon-based*, and it is one of three broad classes of computational tools used for sentiment analysis across disciplines, which can be distinguished according to whether they use word lists (e.g., Jockers), vector-space models (e.g., Turney and Littman), or a combination of both (for reviews see Taboada et al.; Jacobs and Kinder). In line with other scholars (Jacobs et al.), we rely here on the lexical coverage of lexicons – i.e. the percentage of words of a given text or corpus covered by a given sentiment lexicon – to establish the validity of our results.⁴

Some limitations have to be kept in mind when working with lexicon-based approaches: the finite amount of words in the lexicons (on which the coverage often depends); the differences between one lexicon and another; and the sometime lack of transparency in the way the lexicons were developed. Moreover, sentiment lexicons are built on contemporary language usage, and might not be comprehensive of historical forms or archaic meanings of words that are pertinent to a specific historical corpus such as the one we are examining here. The two alternative approaches – creating (and validating) a custom sentiment lexicon, or using machine learning – are however extremely resource and time intensive. Since tests in the current historical project showed a word coverage comparable to contemporary language, we decided to utilise out-of-the box resources.⁵ Lexicon-based approaches remain among the most easily accessible and practicable methods

⁴ In general terms, it is assumed that a low coverage decreases the validity of a sentiment lexicon. Some scholars from experimental psychology have even suggested that a coverage below 50% cannot be considered reliable (Jacobs et al.).

⁵ Elsewhere, with a different aim, we report on ongoing domain adaptation of sentiment analysis using machine learning (Grisot et al.).

for sentiment analysis (Jockers and Thalken; Flüh), and as these represent the essential building ground for more complex unsupervised sentiment models (Nakov et al.; Barbieri et al.), it is crucial to keep using and improving them.

The application and further development of sentiment analysis techniques has been impacted by industrial interest in customer research, with most of the research done on the English language. Not as much attention has been given to the investigation of literary and other creative texts, and up to now only few resources are available for languages other than English. Nonetheless, in recent years, with the expansion of the digital humanities, new research has resulted in valuable lexicons also for other languages, including German.

We chose to examine emotions and sentiments in our corpus using all the open source lexicons for German available to this date, as suggested by recent research in the field (Kern et al.; Schmidt, Dangel, et al.). We hope therewith to provide a ground for gauging not only how these lexicons compare to each other, but, most crucially, how valid they are for analysing the representation of emotions in relation to space in historical German-Swiss fiction (~1840-1940).

It is not the scope of the present paper to provide a comprehensive comparison of the existing sentiment lexicons for German, nor to evaluate in depth their coverage of the historical dimension of affective language. However, some of their characteristics – in particular regarding the number of words included in the lexicons and their coverage of our corpus – do help us evaluating the potential differences in their performance. A minimal summary of the lexicons used and their composition is shown in [Table 2](#).

Table 2. Main features of the sentiment and emotion lexicons used in the present paper (highest to lowest coverage).

Lexicon	Swiss-corpus coverage	Value type	Word N
Sentiart ^a (Jacobs)	81.8%	Emotions (anger, disgust, fear, happiness, sadness, surprise)	116,313
BAWL-R (Vö et al.)	33.6%	Sentiment (valence, arousal, imageability) ^b	2,902
GermanPolarityClues (Waltinger)	32.1%	Sentiment (polarity)	9,561
Germanlex (Clematide et al.)	15.6%	Sentiment (polarity)	8,681
SentiWS (Remus et al.)	14.9%	Sentiment (polarity)	3,468
Klinger et al. (Klinger et al.)	10.7%	Emotions (anger, contempt, disgust, fear, joy, sadness, surprise)	4,735
LANG (Kanske and Kotz)	7.6%	Sentiment (valence, arousal, concreteness)	1,000
Stamm (Stamm)	3.4%	Emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, trust)	821

^a <https://github.com/matinho13/SentiArt>

^b For our analysis, we decided to focus on comparable measures only, ignoring therefore measures unique to individual lexicons. The ‘imageability’ value from the BAWL-R lexicon and the ‘concreteness’ value from the LANG lexicon were therefore not taken into account in our analysis.

Method

Objectives

From the prominent role of a Swiss-national framework that marked literary production during our time frame, and that putatively sees an opposition between rural/natural and urban spaces, we formulate here two main hypotheses: (1) natural and rural landscapes will be overall more present in our corpus in comparison to towns and cities, and (2), natural and rural landscapes will be encoded overall as more ‘positive.’ Aware that realist writing does depict problematic aspects of life in rural surroundings, however, we retain that the difference in emotional encoding between urban and rural spaces may not be as neat as we hypothesise.

Our exploration will start with (1) a description of the distribution of different types of spatial entities detected in our corpus, and move on to (2) investigating their affective encoding. Importantly, this extends to questions about how ‘spatially self-referential’ the texts in our corpus are, comparing the proportion of mentions of Swiss geographical locations with those of countries surrounding Switzerland (the current territories of Italy, France, Austria and Germany).

We are not aware of any research investigating the presence and distribution of non-named spatial entities along with geolocations in fictional texts, nor systematically comparing rural and urban environments (real or fictional) in terms of the sentiments and emotions related to these. It is this gap that the present paper addresses. Using on the one hand lists of spatial entities that can be matched to our historical literary corpus, and on the other sentiment lexicons that capture the sentiment and emotional value of words, we examine here how the rural and natural landscape and geographical locations are represented, in opposition to urban space and settlements.

Corpus

The compilation of a corpus of Swiss fictional narratives written between the years 1843 and 1940 presented several challenges, in particular due to the lack of a comprehensive catalogue of Swiss fictional works for those years, and to the limited amount of works already existing in a digital form. With the collaboration of experts from the University Library Basel, a comprehensive list of the authors and works of interest to our project was produced,⁶ featuring more than 7,000 titles of fictional narratives written by Swiss authors in German between the years 1880 and 1930.⁷

⁶ See details in Herrmann et al.

⁷ Authors born elsewhere who completed their education in Switzerland were also included, considering that a sufficiently long period of time for them to have been influenced by its culture and become aware of its geography.

Using the list of authors and titles obtained, we aimed to gather as many of the listed works as possible in their digital version. A corpus of 125 Swiss texts was thus compiled, combining 79 texts which already existed in digital form and matched our list of Swiss fiction texts – from [Gutenberg.de](https://www.gutenberg.de) (Projekt Gutenberg DE) – with 46 additional texts, which were newly digitised by the University Library Basel as part of the research project ‘*High Mountains Low Arousal? Distant Reading Topographies of Sentiment in German Swiss Novels in the early 20th Century*’ as a contribution to the *European Literary Text Collection (ELTeC)* (Schöch et al.). The corpus contains narrative texts by 42 different authors, with a maximum of $n=14$ works attributed to a single author (Heinrich Federer) and a minimum of $n=1$ (several authors). More details about the final corpus are available on our online repository.⁸

All of the texts were stored as plain text files, and were then processed with the German language model of the R package ‘Udpipe’ (R Core Team; Wijkfjels), to tokenise and lemmatise the texts while preserving information about paragraphing and sentence structure.

Spatial entities

In terms of the operationalisation of literary landscape, our categorisation of spatial entities overlaps in part with what Wartmann et al. (1980) identify as *biophysical landscape elements* – i.e. ‘terms relating to geology, landforms, soil, land cover, flora, fauna, and climate,’ as distinct from *cultural landscape elements* – ‘terms referring to land use, settlements, infrastructure, domesticated animals, and anthropogenic objects.’⁹ Central to this definition is the notion of landscape as being something which is perceived and defined by the ways in which it is described by people (Scott, “Assessing Public Perception of Landscape: The LANDMAP Experience”; Scott, “Assessing Public Perception of Landscape: From Practice to Policy”).

We found however that this taxonomy only takes into account the distinction between ‘natural’ and ‘human-made,’ but not that between ‘rural’ and ‘urban.’ Because both ‘rural’ and ‘urban’ are related to human practices, we decided to use a broader umbrella notion of RURAL for all terms related to ‘nature’ and proper names of natural places, as well as terms referring to rural landscapes and proper names of rural settlements. Analogously, URBAN is

⁸ The data used for this paper and some additional resources are available on Zenodo (European Organization For Nuclear Research and OpenAIRE) at the following link: <https://doi.org/10.5281/zenodo.7024595>

⁹ We use the term ‘landscape’ as defined by the European Landscape Convention (Council of Europe), namely as ‘an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors.’

our umbrella term for both proper names of larger settlements – cities and towns – and for terms describing – or characteristic of – the urban landscape as generally conceived in the European culture.¹⁰

In order to further differentiate RURAL and URBAN space we identified four ‘entity’ subcategories for the broad category RURAL, namely ‘rural’ and ‘natural’ terms, ‘rural geolocations,’ and ‘natural geolocations;’ as well as two entities subcategories for URBAN: ‘urban geolocations’ and ‘urban’ terms. We collected as many terms as we could for each subcategory, including, where possible, terms that are characteristic of Swiss landscapes/architecture, such as *Hütte* or *Dörfli*. The six categories are described in [Table 3](#) below:

Table 3. Six entities categories identified under the umbrella terms RURAL and URBAN.

RURAL	URBAN
‘rural entities’: spatial terms relating to – or characteristic of – the countryside, in particular related to human settlements or infrastructures, as opposed to those of the city (for example <i>Wanderweg</i> : footpath, <i>Feld</i> : field, <i>Hütte</i> : hut, shack)	‘urban entities’: spatial terms relating to the city, its buildings and infrastructures (for example <i>Bahnhof</i> : station, <i>Kreuzung</i> : cross, <i>Palast</i> : palace)
‘natural entities’: terms describing spatial elements as found in nature, not involving anything made or done by people (for example <i>Berg</i> : mountain, <i>Bach</i> : brook, <i>Wald</i> : forest)	‘geolocation-urban entities’ (geoloc_urb): proper names of cities and towns in Switzerland, Austria, Germany, France and Italy (for example <i>Berlin</i> , <i>Paris</i> , <i>Rom</i>)
‘geolocation-natural (geoloc_nat) entities’: proper names of natural locations such as mountains, rivers, valley, lakes (for example <i>Matterhorn</i> , <i>Mont Blanc</i> , <i>Donau</i>)	
‘geolocation-rural (geoloc_rur) entities’: proper names of villages and small towns – with less than 5,000 inhabitants – in Switzerland, Austria, Germany, France and Italy (for example <i>Zumdorf</i> , <i>Enggwil</i> , <i>Huttwil</i>)	

Because of the morphology of the German language with its abundance of compound words, the identification of verbs, adverbs and adjectives potentially belonging to the spatial categories revealed complex and often ambiguous (e.g., directional adverbs such as *talabwärts*, *bergabwärts*: ‘down the valley,’ ‘downhill,’ or *berghoch*: ‘as tall as a mountain / ‘uphill;’ all containing ‘Tal’ – valley – and ‘Berg’ – mountain). It was therefore decided to focus on substantives only.

Rural, *urban* and *natural* terms were initially collected from openthesaurus.de (Naber), Wikidata (Wiki), and Swiss Idiotikon (Hunger et al.). These lists were then enriched with the natural terms gathered in the Text+Berg project Bubenhofer et al. as presented by Derungs and Purves and with the urban terms collected by Bologna, translated into German. The lists were then manually checked and improved with the help of a student assistant (L1 German and Swiss German), who removed redundancies and

¹⁰ While designations vary in different countries and regions, the International Statistics Conference of 1887 defined different sizes of German cities based on their population size, as follows: *Landstadt* (‘country town’; under 5,000), *Kleinstadt* (‘small town’; 5,000 to 20,000), *Mittelstadt* (‘middle-sized town’; between 20,000 and 100,000) and *Großstadt* (‘large town’; 100,000 or more) (Baumgart et al.). We apply this taxonomy to our geographical locations independently of the country, considering all settlements with a population smaller than 5,000 as villages – thus belonging to the RURAL entity-group – and all settlements with a population above 5,000 as cities.

Table 4. Distribution of different spatial entity categories in entity list.

	Spatial entities in entity list (N=173,913)
type	
geoloc_nat	42,554 (24.5%)
geoloc_rur	124,812 (71.8%)
geoloc_urb	5,487 (3.2%)
natural	531 (0.3%)
rural	271 (0.2%)
urban	258 (0.1%)
type_grouped	
RURAL	168,168 (96.7%)
URBAN	5,745 (3.3%)

cross-category terms. While we discarded texts fully written in German-Swiss dialect from the corpus because of problems with language processing – especially in tokenization and sentiment analysis – some relevant dialectal words (e.g. *Hütte*, *Dörfli*) and as many Swiss-specific terms as possible were added for each list.

For *geoloc_urb*, *geoloc_rur* and *geoloc_nat* locations of today's territories of Switzerland, Germany, Austria, Italy and France, open data available from [geonames.org](https://www.geonames.org) (Wick and Boutreux) and Wikidata (Wiki) were used. We decided to include geographical locations of countries surrounding Switzerland as these are likely to play a part in the representation of the RURAL and URBAN landscape in German-Swiss literature in the historical time frame considered. These sources allowed us to store also more specific information, in particular about the category *geoloc_nat*, enabling us to distinguish whether these were *streams* or *lakes*, *valleys*, *mountains* or *forests*.

After a process of selection, removal of duplicates, hyphenated location names, cross-category terms (for example *Kirche* – ‘church’) and names of places that correspond to common nouns (for example *Zug*, which also denotes ‘train’) or proper names,¹¹ we obtained a final list of N=173,913 spatial entities belonging to the five subcategories of RURAL and URBAN. Details of these are summarised in [Table 4](#).

Sentiment and discrete emotions

After removing German stopwords, symbols and punctuation, we matched our lists of spatial entities onto the corpus. We created therewith a new corpus composed by the spans defined by the matched spatial terms, where each span comprised 101 words: one matched entity, the 50 words preceding it and the 50 words following it. The resulting corpus counted N=6,776,378

11 A list of German first names was obtained from the open archive of the city of Cologne <https://offenedaten-koeln.de/dataset/vornamen>.

tokens. We then merged each sentiment lexicon onto our corpus by lower case lemmas – having also lemmatised each token – and calculated aggregated values for all the sentiment/emotions for each of text span defined by the presence of a spatial entity. Sentiment and emotion values were aggregated by span. For lexicons measuring discrete emotions, we gave a score of 1 to each word in the span with an emotion value, and counted the sum per emotion, per span, thus obtaining counts for each discrete emotion. Also for lexicons measuring sentiment (arousal, valence) we aggregated the values by span: where values were continuously scaled, we calculated the mean values for arousal and valence. Since mixed models for lexicons were computed independently, we did not need to normalise values across lexicons. If the valence value was categorical (negative vs. positive), we converted these respectively into -1 and 1 and calculated the mean value by span.

As each span was defined by the presence of a single entity, we were able to label each span as either RURAL or URBAN. This allowed us to examine the corpus as a whole, analysing statistically the two broad spatial categories RURAL and URBAN for differences between encoded emotions and sentiment values.

To establish beforehand whether the entities in our lists possessed themselves an emotional value (as this could affect the values for sentiment and emotions of the spans) we found that spatial terms covered together only the 1.1% of the entities. This shows on the one hand that the sentiment lexicons used contain very few locations or spatial terms, and on the other hand that the entities themselves – necessarily present in each text span in our corpus – would be unlikely to bias the sentiment and emotion value of the surrounding text.

Statistical analyses

For our statistical analyses, we used linear mixed-effect models (Kuznetsova et al.; Bates et al.; Baayen), with the broader space type (URBAN/RURAL) as fixed factor, and the author and title of each of the spans as random factors. We ran mixed models for each lexicon independently.

For spatial entities we report descriptive statistics in terms of raw and relative frequencies.

Results

Spatial entities

[Figure 1](#) compares the distribution of spatial entities (*rural*, *urban*, *natural*, *geoloc_rur*, *geoloc_urb*, *geoloc_nat*) in the list (on the left side) with that of the corpus (on the right side). Colors distinguish between the main categories

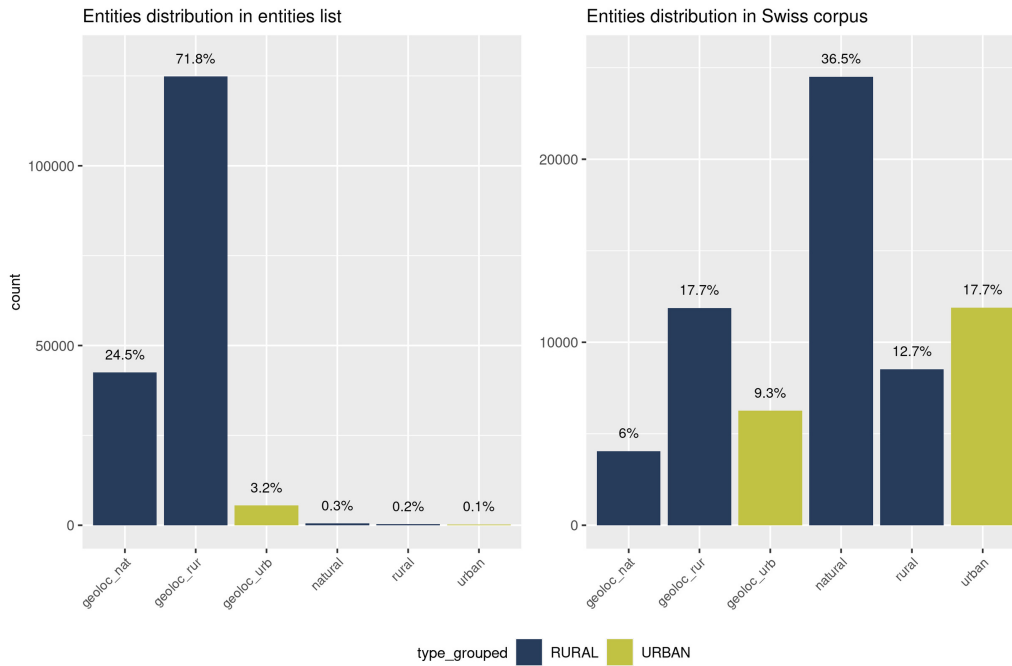


Figure 1. Spatial entities distribution.

RURAL and URBAN. We observe a pronounced skewness for our entities list, where combined RURAL entities amount to 96.7% in total, a number that largely surpasses that of URBAN ones (3.3% in total).

Here, it is important to note the categorical difference between the two bar charts: In the left bar chart, we see the entity list, which was compiled from multiple resources. It is a proxy for the distribution of factual geographical locations in Switzerland, France, Germany, Italy and Austria: Here, the total number of villages, mountains, rivers, lakes and valleys – i.e. *geoloc_rur* and *geoloc_nat* entities – greatly surpasses the number of large towns and cities (*geoloc_urb*). In this bar chart, the number of named-entity terms is also by far higher than the number of *rural*, *urban* and *natural* non-named entity terms.

By contrast, the bar chart on the right shows the text corpus. Searching for the list's terms in the N=125 Swiss literary texts (N=3,539,090 tokens), we find a total of N=67,093 (meaning that 1.9% of all tokens in our corpus are either RURAL or URBAN spatial entities).

Here, the difference between the macro categories URBAN (green) and RURAL (blue) is much less pronounced, but still remarkable: around three quarters of all detected spatial entities fall into the RURAL umbrella category (73%, vs. 27% URBAN).

Details of the distribution of entities in the Swiss corpus can be observed in [Table 5](#) (for ease of comparison, the distribution of the entities list (left column) is repeated from [Table 4](#)).

Table 5. Distribution of different spatial entity categories in entity list and in corpus.

type	Spatial entities in entity list	Spatial entities from entities lists found in Swiss Corpus	Unique spatial entities/tokens in corpus and coverage
geoloc_nat	42,554 (24.5 %)	4,055 (6 %)	865 / 2.03 %
geoloc_rur	124,812 (71.8 %)	11,864 (17.7 %)	1,978 / 1.58 %
geoloc_urb	5,487 (3.2 %)	6,254 (9.3 %)	449 / 8.18 %
natural	531 (0.3 %)	24,500 (36.5 %)	306 / 57.63 %
rural	271 (0.2 %)	8,531 (12.7 %)	150 / 55.35 %
urban	258 (0.1 %)	11,889 (17.7 %)	122 / 47.29 %

In the corpus, the overall prevalence of RURAL entities (n=48,950) in comparison to URBAN ones (n=18,143) is particularly due to *natural* terms, the biggest category overall, with 36.5% of all spatial entities detected (n=24,500 tokens). It is followed by the *urban* category, with n=11,889 tokens, and the *geoloc_rur* (n=11,864 tokens) category, which each contribute about half the proportion of *natural*.

Relative to the number of terms in the entities lists, we note the highest coverage for non-named entities (*natural*, *urban* and *rural*). By coverage we understand the percentage of unique terms (corpus-linguistic *types*) per category actually found as tokens in the corpus).

This can be explained by the relatively small number of terms included in these three categories, and the fact that we are dealing with common terms, as opposed to proper names of locations. However, it is notable that the relatively small categories *natural*, *urban* and *rural* (n=1060 in total in the entities list) represent about two thirds of all the spatial entities found in the corpus (n=44,920, 67%).

This is quantitative support for the observation that general and common terms (our *rural*, *urban* and *natural* categories) are used across narratives to build the so called ‘storyworld’ (Gavins; Troscianko): creating settings, landscapes, natural phenomena and objects in the fictional world (Nünning). Geolocations (*geoloc_urb*, *geoloc_rur* and *geoloc_nat*), on the other hand, while making up an overpowering proportion in the spatial entities list (see [Figure 1](#)), are comparably less frequent in the corpus (33%). Such low frequency can be explained by their semantic specificity, and by the fact that fictional texts do not necessarily mention (m)any existing geographical locations. Given the sheer number of locations listed in our entities list, it is not surprising that only a small fraction would make it into literary texts, and fewer still into a relatively small corpus such as the one analysed here.

Nonetheless, when considering the percentages of geolocations (*geoloc_rur* 17.7%, *geoloc_urb* 9.3%, *geoloc_nat* 6%), the dominance of RURAL is again notable: *geoloc_rur* and *geoloc_nat* together amount to 23.7% of all spatial entities. Together, their substantial proportion (one fourth of all spatial

entities) may be interpreted as reflecting a type of realistic writing that is anchored not only in (imaginary) references to ‘meadow,’ ‘river,’ ‘hill’, but which can be mapped onto factual geography.

NATIONAL SELF-REFERENTIALITY: ARE SWISS GEOLOCATIONS MORE PROMINENT IN THE CORPUS IN COMPARISON TO OTHER COUNTRIES?

With regard to the dimension of national identity construction discussed above, we still need to answer an open question: do the authors in our corpus – all Swiss – favour the representation of Swiss geographical locations in comparison to locations in other countries?

As discussed above, the list of $N = 173,913$ spatial entities compiled for the present paper was inclusive of six categories: *geoloc_nat* (natural geolocations), *geoloc_rur* (rural geolocations), *geoloc_urb* (urban geolocations), *natural*, *rural* and *urban*. In addition to this categorization, we were able to classify geolocations by country (Austria, France, Germany, Italy, Switzerland), and for the *geoloc_nat* category distinguish among specific sub-classes, namely *streams*, *lakes*, *valleys*, *mountains* or *forests*.

In view of the aforementioned strengthening of nationalism in Switzerland during the period 1843-1940, we assumed that it would exhibit more national self-references as opposed to references to other countries.¹² In operational terms, we expected the corpus to feature, proportionally, more Swiss geographical locations than non-Swiss locations, and in particular we hypothesised that rural and natural geolocations would be prominent in comparison to urban ones, especially when observing Swiss geolocations.

Using identical tools to gather spatial entities and geographical locations for the five countries considered, the raw account of entities per country in our entity list broadly mirrors the country size (see upper part of [Table 6](#)), with Germany at the top counting $n=74,135$ geographical locations, and Switzerland last, counting $n=13,784$.¹³

At first glance, when focusing on the figures printed in the lower half [Table 6](#), it appears that our corpus of German-Swiss texts contains more German geolocations than Swiss ones, across all three categories (*geoloc_nat*, *geoloc_rur*, *geoloc_urb*). This is in fact true in raw counts, and may be taken as support for the basic tenet that German-written literature is a transnational literature, transcending boundaries between the German Empire (as of 1871), Austria, and Switzerland – a consideration that can be extended to

¹² For a similar comparison concerning American vs British literature, see Evans and Wilkens and Wilkens.

¹³ For details of the unique token (coprus-linguistic type) count per category and the relative coverage with respect to the entities lists see additional material on Zenodo.^[^zenodo]

geolocations of France and Italy, as shown by our data. However, a closer comparative look at the upper half (list) and the lower half (corpus) of the table reveals an interesting pattern: in comparison to the proportions of unique tokens of geographical entities in our list (upper half) we observe a consistently and significantly higher percentage of Swiss geographical locations in the lower half,¹⁴ while most other percentages drop between the list and the corpus. This is summarised in [Figure 2](#) below.

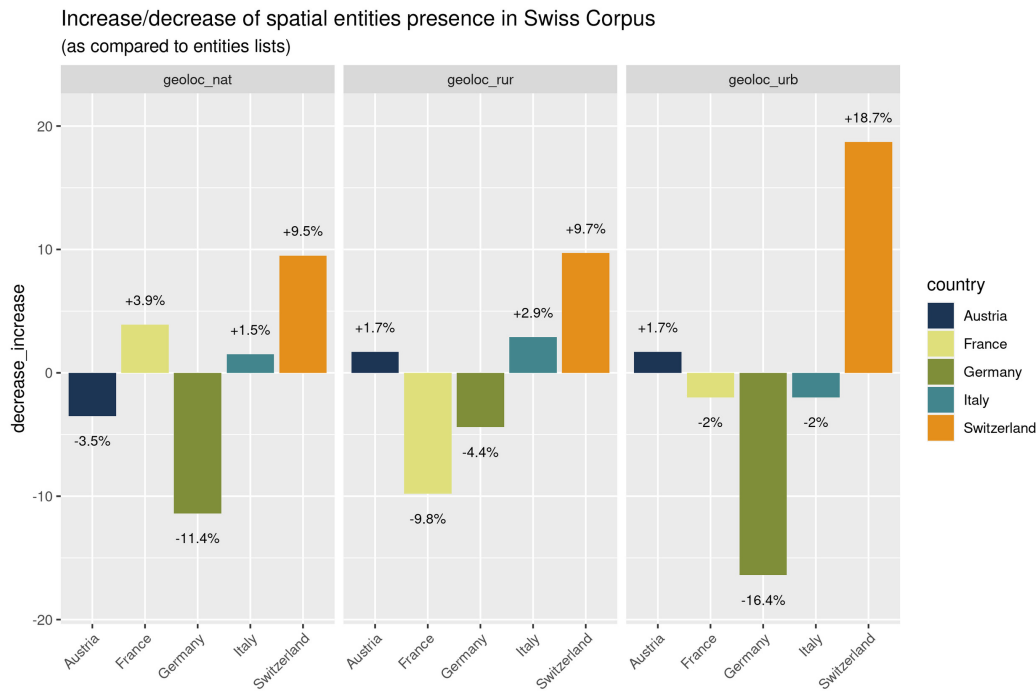


Figure 2. Barplot showing increase and decrease for the three geolocations subcategories by country in Swiss corpus.

Table 6. Overall distribution of geographical locations in the entities list and in the corpus.

DF		geoloc_nat	geoloc_rur	geoloc_urb
Entities list		(N=44,730)	(N=125,038)	(N=5,604)
	country			
	Austria	10,267 (23.0%)	11,493 (9.2%)	242 (4.3%)
	France	2,799 (6.3%)	32,761 (26.2%)	1,145 (20.4%)
	Germany	23,142 (51.7%)	48,387 (38.7%)	2,606 (46.5%)
	Italy	1,847 (4.1%)	25,590 (20.5%)	1,309 (23.4%)
	Switzerland	6,675 (14.9%)	6,807 (5.4%)	302 (5.4%)
Swiss corpus		(N=5,292)	(N=12,005)	(N=7,265)
	country			
	Austria	1,030 (19.5%)	1,303 (10.9%)	437 (6.0%)
	France	542 (10.2%)	1,967 (16.4%)	1,336 (18.4%)
	Germany	2,133 (40.3%)	4,114 (34.3%)	2,187 (30.1%)
	Italy	296 (5.6%)	2,813 (23.4%)	1,554 (21.4%)
	Switzerland	1,291 (24.4%)	1,808 (15.1%)	1,751 (24.1%)

¹⁴ The significance between proportions was calculated with a 2-sample test for equality of proportions with continuity correction. Detailed results can be downloaded from our Zenodo repository.[[^]zenodo]

Importantly, for all three geolocation categories – *geoloc_nat*, *geoloc_rur* and *geoloc_urb* – the proportion of Swiss locations features a substantial increase in the corpus in comparison to the presence of the same entities in Switzerland's neighbour countries, with natural geolocations going from 14.9 to 24.4 (+9.5%), rural settlements from 5.4 to 15.1 (+9.7%) and Swiss cities rising from 5.4 to 24.1 (+18.7%).

While we do not receive support for an extreme national focus we can thus observe a notable tendency for Swiss self-reference in the corpus, especially when comparing the *proportional numbers* for the different entities categories. This is shown in [Table 7](#) and [Figure 3](#) below, which zoom in on the geolocations subcategories *city*, *forest*, *mountain*, *stream/lake*, *valley* and *village*. They show an increase of detected Swiss geographical locations specifically for 'valleys' (+57.6% in comparison to the entity list) and cities (+18.7%), followed by villages (+9.7%), mountains (+9%) and streams/lakes (+4.4%). This increase is counterbalanced particularly by a decrease of references to German and Austrian 'valleys' and German cities.

Again, we find a complex transnational picture, with a relative dominance of German geolocations for all categories, except valleys. However, the comparison between list and corpus shows a notable difference in the categories indicative of the idealised Swiss landscape: valleys, villages, mountains and streams and lakes. In terms of absolute frequencies for the corpus in general (regardless of national attribution), subcategories show the following rank order: village (n=11,864), city (n=6,254), mountain (n=2,762), stream_lake (n=952), forest (n=250) and valley (n=91)).

This rank order mirrors the known setting of the stories in quantitative terms.

The analysis of proportions of the different spatial categories in our corpus provides insight into the depiction of space and landscape in Swiss fiction between 1840 and 1940 – a time frame that in German literary history covers epoch-constructions such as 'Biedermeier'/restoration, realism, and modernism. Working towards an Swiss-specific literary historiography, we observe a dominance of RURAL space that may be less pronounced in other German-language literatures, a use of natural and rural geolocations that may indicate a specific type of realistic anchoring in non-fictional and often Swiss geography, as well as a dominance of landscape elements that match 'idealised Swissness'. Meanwhile, the observed comparatively high recurrence of non-named spatial terms appears a universal feature of narratives. In all, these data-based patterns underscore the assumption of a 'Swiss-specific' encoding of the rural and natural environment.

Table 7. Distribution of geographical locations in the entity list and in the corpus, by category.

DF		city	forest	mountain	stream_lake	valley	village
Entities list		(N=5,604)	(N=2,589)	(N=27,339)	(N=14,265)	(N=537)	(N=125,038)
	country						
	Austria	242 (4.3%)	238 (9.2%)	7,532 (27.6%)	2,348 (16.5%)	149 (27.7%)	11,493 (9.2%)
	France	1,145 (20.4%)	11 (0.4%)	768 (2.8%)	2,013 (14.1%)	7 (1.3%)	32,761 (26.2%)
	Germany	2,606 (46.5%)	2,072 (80.0%)	12,513 (45.8%)	8,368 (58.7%)	189 (35.2%)	48,387 (38.7%)
	Italy	1,309 (23.4%)	208 (8.0%)	1,221 (4.5%)	331 (2.3%)	87 (16.2%)	25,590 (20.5%)
	Switzerland	302 (5.4%)	60 (2.3%)	5,305 (19.4%)	1,205 (8.4%)	105 (19.6%)	6,807 (5.4%)
Swiss corpus		(N=7,265)	(N=307)	(N=3,776)	(N=1,117)	(N=92)	(N=12,005)
	country						
	Austria	437 (6.0%)	60 (19.5%)	837 (22.2%)	133 (11.9%)	0 (0%)	1,303 (10.9%)
	France	1,336 (18.4%)	0 (0%)	217 (5.7%)	324 (29.0%)	1 (1.1%)	1,967 (16.4%)
	Germany	2,187 (30.1%)	218 (71.0%)	1,454 (38.5%)	457 (40.9%)	4 (4.3%)	4,114 (34.3%)
	Italy	1,554 (21.4%)	24 (7.8%)	196 (5.2%)	60 (5.4%)	16 (17.4%)	2,813 (23.4%)
	Switzerland	1,751 (24.1%)	5 (1.6%)	1,072 (28.4%)	143 (12.8%)	71 (77.2%)	1,808 (15.1%)

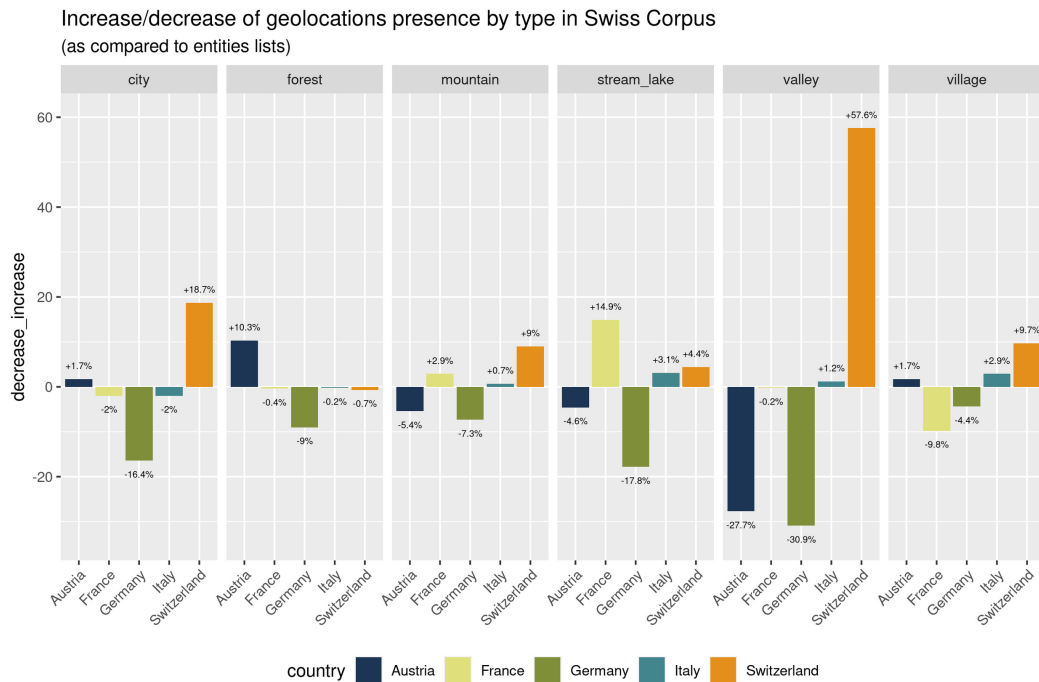


Figure 3. Barplot showing increase and decrease for the three geolocations subcategories by country in Swiss corpus.

Sentiment and Emotions

As mentioned above, for each sentiment and emotion value present in the sentiment lexicons, we fitted a mixed model which looked at the difference between spans defined as RURAL or URBAN based on the spatial entities detected in the corpus. The model had thus RURAL/URBAN as fixed factor, and included title and author as random effects (formula: ‘value ~ type_grouped + (1|author/title)’).¹⁵ All the results of these models are summarised in [Table 8](#) and [9](#) below. The columns *RURAL* and *URBAN* summarise the result of each model, with a ‘+’ and a ‘-’ simplifying the models outcome. When a ‘+’ or a ‘-’ is shown, it indicates that the entity type RURAL/URBAN had significant effect on the sentiment/emotion value. A ‘+’ shows an increase for the value considered, while a ‘-’ indicates a significant decrease. If the cell contains ‘na’, the entity type had no significant effect on that value.

In line with our expectations, we detected a general increase in the emotional ‘richness’ of spans containing RURAL entities, in comparison to URBAN ones. Particularly informative are the results obtained with the SentiArt

¹⁵ Analyses were conducted using the R Statistical language (version 4.3.0; R Core Team 2023) on Ubuntu 22.04.2 LTS, using the packages ggthemes (version 4.2.4; Arnold J, 2021), janitor (version 2.2.0; Firke S, 2023), lubridate (version 1.9.2; Grolemund G, Wickham H, 2011), jcolors (version 0.0.4; Huling J, 2019), ggpubr (version 0.6.0; Kassambara A, 2023), sjPlot (version 2.8.14; Lüdtke D, 2023), tibble (version 3.2.1; Müller K, Wickham H, 2023), patchwork (version 1.1.2; Pedersen T, 2022), table1 (version 1.4.3; Rich B, 2023), ggplot2 (version 3.4.2; Wickham H, 2016), stringr (version 1.5.0; Wickham H, 2022), forcats (version 1.0.0; Wickham H, 2023), tidyverse (version 2.0.0; Wickham H et al., 2019), readxl (version 1.4.2; Wickham H, Bryan J, 2023), dplyr (version 1.1.2; Wickham H et al., 2023), purrr (version 1.0.1; Wickham H, Henry L, 2023), readr (version 2.1.4; Wickham H et al., 2023), tidyr (version 1.3.0; Wickham H et al., 2023), ggsci (version 3.0.0; Xiao N, 2023), knitr (version 1.42; Xie Y, 2023) and kableExtra (version 1.3.4; Zhu H, 2021).

Table 8. Results of Linear Mixed Models: Effect of entity type on sentiment (valence, polarity, arousal) for the different lexicons, with author and title as random factors.

Lexicon	Value	RURAL	URBAN	P_signif	Coefficient	t	p	Std_Coefficient	R2_Condtional
BAWL	valence	+	-	***	-0.010	-4.334	0.000	-0.036	0.137
BAWL	arousal	ns	ns	ns	0.001	1.137	0.255	0.010	0.128
GerPol	polarity	-	+	***	0.017	7.963	0.000	0.067	0.126
Glex	polarity	-	+	***	0.017	6.768	0.000	0.057	0.115
LANG	arousal	ns	ns	ns	-0.008	-1.317	0.188	-0.011	0.088
LANG	valence	+	-	**	-0.009	-2.594	0.009	-0.022	0.067
SentiWS	polarity	-	+	*	0.002	2.284	0.022	0.019	0.091
Sentiart	AAPz	+	-	***	-0.009	-5.552	0.000	-0.045	0.193

Table 9. Results of Linear Mixed Models: Effect of entity type on discrete emotions for the different lexicons, with author and title as random factors.

Lexicon	Value	RURAL	URBAN	P_signif	Coefficient	t	p	Std_Coefficient	R2_Conditional
Klinger et al.	joy	ns	ns	ns	0.022	1.079	0.281	0.009	0.154
Klinger et al.	fear	+	-	***	-0.119	-8.676	0.000	-0.074	0.086
Klinger et al.	contempt	ns	ns	ns	0.024	1.918	0.055	0.017	0.063
Klinger et al.	sadness	+	-	*	-0.040	-2.564	0.010	-0.022	0.062
Klinger et al.	anger	+	-	***	-0.064	-7.169	0.000	-0.062	0.056
Klinger et al.	surprise	+	-	**	-0.037	-3.137	0.002	-0.027	0.045
Klinger et al.	disgust	ns	ns	ns	-0.008	-1.415	0.157	-0.012	0.034
Sentiart	sadness	+	-	***	-0.058	-32.809	0.000	-0.261	0.297
Sentiart	happiness	+	-	***	-0.035	-20.667	0.000	-0.167	0.249
Sentiart	surprise	+	-	***	-0.010	-6.504	0.000	-0.052	0.242
Sentiart	anger	+	-	***	-0.048	-28.978	0.000	-0.237	0.222
Sentiart	disgust	+	-	***	-0.039	-25.759	0.000	-0.212	0.217
Sentiart	fear	+	-	***	-0.029	-16.823	0.000	-0.140	0.157
Stamm	joy	ns	ns	ns	-0.017	-1.543	0.123	-0.013	0.129
Stamm	sadness	+	-	**	-0.020	-3.024	0.002	-0.026	0.063
Stamm	anger	+	-	***	-0.021	-4.144	0.000	-0.036	0.047
Stamm	anticipation	-	+	*	0.011	2.504	0.012	0.022	0.045
Stamm	disgust	-	+	***	0.024	6.337	0.000	0.055	0.042
Stamm	trust	ns	ns	ns	-0.001	-0.139	0.889	-0.001	0.038
Stamm	fear	+	-	*	-0.014	-2.149	0.032	-0.019	0.031
Stamm	surprise	ns	ns	ns	-0.002	-0.510	0.610	-0.004	0.029

lexicon, where this was true for all discrete emotions (positive as well as negative ones). With a lexical coverage of 81.8%, results obtained with the SentiArt lexicon can be considered as comparatively most reliable.

With the caveat of a low coverage, the Stamm and Klinger et al. lexicons show significant differences for discrete emotions between RURAL and URBAN as well: Stamms' *anger*, *sadness*, and *fear* appear to be significantly more present in RURAL spans, as well as Klinger et al.'s *anger*, *sadness*, *fear* and *contempt*. Only Stamm's *anticipation* and *disgust* appeared to be higher in URBAN spans, while models for all other emotions did not show any effect of space categories. Because of the low coverage of these lexicons, these results must be however interpreted with great caution.

When looking at sentiment values, we found again results in line with our hypothesis. Both BAWL-R and LANG lexicons featured a significantly higher valence in spans containing RURAL spatial entities, and, more importantly, the same was true for SentiArt's AAP value (affective-aesthetic potential), which strongly correlates with valence (Jacobs and Kinder, "Computing the Affective-Aesthetic Potential of Literary Texts"). Along with the other discrete emotions detected by SentiArt, the AAP appeared to be strongly affected by the entity type, with spans containing RURAL entities featuring significantly higher AAP than those containing URBAN ones, thus indicating a more 'positive' encoding for the first.¹⁶

In terms of arousal, no difference was found for either the BAWL-R or the LANG lexicons.

Discussion

Exploring the representation and emotional encoding of space in German-Swiss literature between 1840 and 1940, we focused on the one hand on the distribution of specific types of space in our corpus – geographical locations (named entities) as well as non-named spatial terms – and on the other on the difference between the two spatial macro categories RURAL and URBAN with regard to valence/polarity and discrete emotions.

The analysis of proportions of the different types of spatial entities in our corpus showed a clear dominance of RURAL entities, constituted mostly by recurrent 'general' non-named spatial terms (*natural*, *urban*, *rural*).

¹⁶ On the basis of the wide comparison of resources, our analysis detected also some ambiguity with respect to valence, as an increase in this value was found for spans containing URBAN entities in comparison to those containing RURAL ones for the GerPol, Glex and SentiWS lexicons. Again, these lexicons have a lower coverage than the SentiArt and BAWL-R lexicons, but all three surpass LANG. While it is not the scope of the present paper to determine the effectiveness of these tools, a more thorough investigation is needed to understand why these results are not in agreement.

Overall, about two thirds of all spatial entities detected are ‘non-named’, while one third of them are in fact geolocations. This number appears substantial, and supports the assumption of a specific realist mode dominating Swiss literature throughout a long 19th Century – a mode that uses extra-fictional references to anchor the story-worlds quite palpably close to reality. Together with the finding that many of these references are ‘Swiss,’ we take these data-based patterns to underscore the assumption of a ‘Swiss-specific’ idealisation of the rural and natural environment. However, in the absence of more highly or broadly scaled analyses, these interpretations are necessarily provisional - yet they point to important research questions on the specificity and factuality of spatial entities, i.e. space representation in fiction across genres, periods, and (trans-)national literatures.

In terms of fictional reference to factual locations, increased proportions of Swiss geolocations were found for most subcategories (*city, village, mountain, stream/lake, valley*), with the exception of *forest* – the category with the overall lowest number of terms and matches. The increase of percentage from the entity list to our corpus was particularly significant for valleys (57.6%) and cities (18.7%), somewhat in line with our assumption about the relevance of the rural/urban dualism in the conceptualisation of landscape. These findings await further (qualitative) exploration, for example looking into the role of specific cities/towns or valleys, exploring fictional connections to the traditional cultural and political predominance of towns in Switzerland. Much the same holds for ‘valleys’, which may be an often idealised and fictionalised dwelling place for the characters of the stories, but which in fact appear only infrequently in our corpus.

All in all, the findings we gathered on a corpus compiled as a proxy for a representative collection of German-Swiss prose support the view that Swiss authors in the period were indeed influenced by ideas about national identity, remarkably setting apart ‘their German literature’ from other national German literatures. Overall, they tended to make substantial use of real-world locations, thereby referring to their own nation more than to nearby countries. At the same time, we see a relatively strong role of German geolocations in our German-Swiss corpus. This finding can be explained by a strong trans- or pre-national tradition of a literature written in German (Muschg). Further comparative analyses are clearly pending, asking about diachronic change, which may potentially be influenced by WWI and WWII, leading to further increase in own-nation references.

When looking at the emotional representation encoded in our corpus in relation to spatial entities, the quantitative results supported main assumptions of literary history: in Swiss literature 1840-1940, rural and natural landscapes were generally perceived and represented as more positive and emotionally rich, when considered in opposition to the urban environment. The stories take predominantly place in villages, rural, and

natural settings. Some of them are located in rather realistic historical scenarios, going along with a full spectrum of everyday joy and sorrow ('authentic stories'); others depict more clearly idealised places of longing, including idyllic scenarios, tourist-oriented mountaineering adventures, or openly ideologised conservative 'Heimatliteratur' (restaurative and/or popularly-clichéd literature). The nature settings incorporate an ambivalent perception of natural landscape, on the one hand potentially encoding the 'sublime' – including fearful/negative as well as admiring/positive emotions at the same time – on the other, they can be indicative of the type of politically engaged story that depicts poverty and rural despair. Meanwhile, urban settings, which altogether are substantially less often explicitly invoked, nevertheless are significantly less positive, and are also attributed significantly fewer different discrete emotions. Life is less joyful here, and there seems to be less life altogether, if 'life' is to be measured in presence and variance of discrete emotions.

Conclusions

The present study set as its goal a proof-of-concept exploration of landscape representation in German-Swiss literary prose, using computational resources and methods that are both openly accessible and popular within DH literary studies to analyse sentiment and emotion related to these (elsewhere, with a different aim, we report on ongoing domain adaptation of sentiment analysis using machine learning (Grisot et al.).

Our results detected significant differences in the affective encoding of rural and urban spatial entities, with rural/natural space being overall depicted as possessing a richer array of discrete emotions, and, crucially, as being represented overall more positively. Questions might raise from the fact that the richer array of emotions includes 'positive' as well as 'negative' ones. It is to be considered, therefore, that across the different stories collected in the corpus, the representation of rural life possesses negative connotations related to the poor quality of life as well as positive ones depicting an idyllic and bucolic environment. A second, related line of interpretation draws on theories of the sublime, being defined by mixed feelings especially in presence of large natural bodies such as mountain ranges and peaks (Pries).

On this same reasoning, it could be then logical to see why the results for valence/positivity are somewhat ambiguous. Future research is of course necessary, on a larger data base as well as digging into the data with a more qualitative eye to understand which patterns determined our results. Importantly, more research is needed also to understand the different results obtained by different sentiment lexicons, and at the same time to evaluate their performance on historical corpora.

Our results showed that when referencing ‘factual’ locations, Swiss authors 1840-1940 have a general tendency to prefer references to their own country, and that they do so when writing about natural and rural locations, but also when mentioning big cities and urban centres. At the same time, they do make ample references to German geolocations, both of which are important findings in the light of the debate on the character and development of a Swiss national literature. In fact, the ‘paradoxical’ position in between a national Swiss distinction and the adherence to a transnational German literary system is well described in the literature (Haupt). Both strands need further testing, in particular on a larger data base and through a comparison with a similar corpus of German and Austrian literary texts.

Our examination leaves open ground to many other questions, such as the effects of genre types on affective encoding (including constructions such as low-brow vs. mid- and highbrow, or the degree of heteronomy of genre), as well as how plot structure and characters feature in relation to space and affect. Else, irony may be studied as a correlate of affect in the typical village-stories, while a taxonomy of ‘the sublime’ may provide a more stable ground for our still speculative interpretation for the richer emotional range in rural settings. After all, our findings did not measure the proportion of (conflicting) emotions at story level (for an approach to measuring a story affect potential, see. What is more, follow-up studies may run analyses on ‘villages’ and ‘small towns,’ as compared to ‘alpine spaces’ and ‘urban interior space’ in a finer grained spatial analysis (Herrmann and Grisot). These may come to terms with sub-genres of landscape-oriented genres. Particularly relevant appears here a specific Swiss history of technology and industrial infrastructure during industrialisation, with hydroelectric power plants (instead of coal and oil), the construction of infrastructure (tunnels and the railway) all of which acquired topological or even mythical status in Swiss literature (Wege). Finally, our quantitative results support the assumption that while generally marked by the transition from modes of realism to modes of modernism throughout European literatures, the period 1840-1940 appears to function a bit differently in Switzerland. Our data have indeed shown that, next to the rural/urban affective divide, certain ‘Swiss-specific’ features are making themselves visible: the rural, the realistic, and the tension between self-reference and Germany-orientation. We cannot wait to venture deeper into the valleys and among these ranges, with a combination of qualitative and quantitative studies.

Peer reviewers: Matt Erlin, Asko Nivala

Dataverse repository: <https://doi.org/10.7910/DVN/T9QAQV>

Packages

- Arnold J (2021). *ggthemes: Extra Themes, Scales and Geoms for ‘ggplot2’*. R package version 4.2.4, <https://CRAN.R-project.org/package=ggthemes>.

- Firke S (2023). *janitor: Simple Tools for Examining and Cleaning Dirty Data*. R package version 2.2.0, <https://CRAN.R-project.org/package=janitor>.
- Grolemund G, Wickham H (2011). “Dates and Times Made Easy with lubridate.” *Journal of Statistical Software*, 40(3), 1-25. <https://www.jstatsoft.org/v40/i03/>.
- Huling J (2019). *jcolors: Colors Palettes for R and ‘ggplot2’ Additional Themes for ‘ggplot2’*. R package version 0.0.4, <https://CRAN.R-project.org/package=jcolors>.
- Kassambara A (2023). *ggpubr: ‘ggplot2’ Based Publication Ready Plots*. R package version 0.6.0, <https://CRAN.R-project.org/package=ggpubr>.
- Lüdtke D (2023). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.14, <https://CRAN.R-project.org/package=sjPlot>.
- Müller K, Wickham H (2023). *tibble: Simple Data Frames*. R package version 3.2.1, <https://CRAN.R-project.org/package=tibble>.
- Pedersen T (2022). *patchwork: The Composer of Plots*. R package version 1.1.2, <https://CRAN.R-project.org/package=patchwork>.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rich B (2023). *table1: Tables of Descriptive Statistics in HTML*. R package version 1.4.3, <https://CRAN.R-project.org/package=table1>(173196).
- Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
- Wickham H (2022). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.5.0, <https://CRAN.R-project.org/package=stringr>.
- Wickham H (2023). *forcats: Tools for Working with Categorical Variables (Factors)*. R package version 1.0.0, <https://CRAN.R-project.org/package=forcats>.
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K,

Yutani H (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, 4(43), 1686. doi:10.21105/joss.01686
<https://doi.org/10.21105/joss.01686>.

- Wickham H, Bryan J (2023). *readxl: Read Excel Files*. R package version 1.4.2, <https://CRAN.R-project.org/package=readxl>.
- Wickham H, François R, Henry L, Müller K, Vaughan D (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.2, <https://CRAN.R-project.org/package=dplyr>.
- Wickham H, Henry L (2023). *purrr: Functional Programming Tools*. R package version 1.0.1, <https://CRAN.R-project.org/package=purrr>.
- Wickham H, Hester J, Bryan J (2023). *readr: Read Rectangular Text Data*. R package version 2.1.4, <https://CRAN.R-project.org/package=readr>.
- Wickham H, Vaughan D, Girlich M (2023). *tidyr: Tidy Messy Data*. R package version 1.3.0, <https://CRAN.R-project.org/package=tidyr>.
- Xiao N (2023). *ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for 'ggplot2'*. R package version 3.0.0, <https://CRAN.R-project.org/package=ggsci>.
- Xie Y (2023). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.42, <https://yihui.org/knitr/>. Xie Y (2015). *Dynamic Documents with R and knitr*, 2nd edition. Chapman and Hall/CRC, Boca Raton, Florida. ISBN 978-1498716963, <https://yihui.org/knitr/>. Xie Y (2014). “knitr: A Comprehensive Tool for Reproducible Research in R.” In Stodden V, Leisch F, Peng RD (eds.), *Implementing Reproducible Computational Research*. Chapman and Hall/CRC. ISBN 978-1466561595.
- Zhu H (2021). *kableExtra: Construct Complex Table with 'kable' and Pipe Syntax*. R package version 1.3.4, <https://CRAN.R-project.org/package=kableExtra>.

Submitted: February 06, 2023 EDT, Accepted: March 06, 2023 EDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

WORKS CITED

- Alm, Cecilia Ovesdotter, and Richard Sproat. "Emotional Sequencing and Development in Fairy Tales." *International Conference on Affective Computing and Intelligent Interaction*, Springer, 2005, pp. 668–74, https://doi.org/10.1007/11573548_86.
- Arnold, Magda B. *Emotion and Personality. Vol. 1: Psychological Aspects*. Columbia Univ. Press, 1969.
- Baayen, R. H. "Analyzing Linguistic Data: A Practical Introduction to Statistics Using R." *Sociolinguistic Studies*, vol. 2, no. 3, Dec. 2008, p. 353, <https://doi.org/10.1558/sols.v2i3.471>.
- Barbieri, Francesco, et al. "UPF-Taln: SemEval 2015 Tasks 10 and 11. Sentiment Analysis of Literal and Figurative Language in Twitter." *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 2015, pp. 704–08, <https://doi.org/10.18653/v1/s15-2119>.
- Bates, Douglas, et al. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, vol. 67, no. 1, 2015, pp. 1–48, <https://doi.org/10.18637/jss.v067.i01>.
- Baumgart, Sabine, et al. *Klein- und Mittelstädte*. Verkleinerte Blaupausen der Großstadt?, 2004.
- Böhler, Michael. "Gefängnis Schweiz oder Bergnebel Seldwyla?" *Schweiz Schreiben*, Max Niemeyer Verlag, 2010, pp. 45–64, <https://doi.org/10.1515/9783484970526.1.45>.
- Bologna, Federica. "A Computational Approach to Urban Space in Science Fiction." *Journal of Cultural Analytics*, vol. 5, no. 2, Nov. 2020, <https://doi.org/10.22148/001c.18120>.
- Bubenhof, Noah, et al., editor. "Text+Berg-Korpus." *Institut für Computerlinguistik, Universität Zürich; XML-Format*, 2015.
- Bushell, Sally, editor. "A Shifting Relationship: From Literary Geography to Critical Literary Mapping." *Reading and Mapping Fiction: Spatialising the Literary Text*, Cambridge University Press, 2020, pp. 16–43, <https://doi.org/10.1017/9781108766876.002>.
- Clematide, Simon, et al. *Evaluation and Extension of a Polarity Lexicon for German*. 2010.
- Clore, Gerald L., and Andrew Ortony. "Psychological Construction in the OCC Model of Emotion." *Emotion Review*, vol. 5, no. 4, Sept. 2013, pp. 335–43, <https://doi.org/10.1177/1754073913489751>.
- Cooper, David, et al., editors. *Literary Mapping in the Digital Age*. Routledge, 2016, <https://doi.org/10.4324/9781315592596>.
- Cosgrove, Denis. "Prospect, Perspective and the Evolution of the Landscape Idea." *Transactions of the Institute of British Geographers*, vol. 10, no. 1, 1985, p. 45, <https://doi.org/10.2307/622249>.
- Council of Europe. "European Landscape Convention." *Report and Convention*, 2000.
- Denning, Andrew. "From Sublime Landscapes to 'White Gold': How Skiing Transformed the Alps after 1930." *Environmental History*, vol. 19, no. 1, Jan. 2014, pp. 78–108, <https://doi.org/10.1093/envhis/emt105>.
- Derungs, Curdin, and Ross S. Purves. "From Text to Landscape: Locating, Identifying and Mapping the Use of Landscape Features in a Swiss Alpine Corpus." *International Journal of Geographical Information Science*, vol. 28, no. 6, Apr. 2013, pp. 1272–93, <https://doi.org/10.1080/13658816.2013.772184>.
- Donaldson, Christopher, et al. "Locating the Beautiful, Picturesque, Sublime and Majestic: Spatially Analysing the Application of Aesthetic Terminology in Descriptions of the English Lake District." *Journal of Historical Geography*, vol. 56, Apr. 2017, pp. 43–60, <https://doi.org/10.1016/j.jhg.2017.01.006>.
- Eder, Jens, et al. "Media and Emotion: An Introduction." *NECSUS. European Journal of Media Studies*, vol. 8, no. 1, 2019, pp. 91–104.

- Ekman, Paul, et al. "What Emotion Categories or Dimensions Can Observers Judge from Facial Behavior?" *Emotion in the Human Face*, 1982, pp. 98–110.
- European Organization For Nuclear Research and OpenAIRE. *Zenodo*. CERN, 2013, <https://www.zenodo.org/>.
- Evans, Elizabeth, and Matthew Wilkens. "Nation, Ethnicity, and the Geography of British Fiction, 1880-1940." *Journal of Cultural Analytics*, vol. 3, no. 2, July 2018, p. 11037, <https://doi.org/10.31235/osf.io/psg2z>.
- Flüh, Marie. "Sentimentanalyse." *forTEXT*, 2019, <https://fortext.net/routinen/methoden/sentimentanalyse>.
- Fournier, Mauricette. *Rural Writing: Geographical Imaginary and Expression of a New Regionality*. Cambridge Scholars Publishing, 2019.
- Fox, A. S., et al. *The Nature of Emotion: Fundamental Questions*. Oxford University Press, 2018, <https://books.google.it/books?id=fh5pDwAAQBAJ>.
- Gavins, Joanna. *Text World Theory*. Edinburgh University Press, 2007, <https://doi.org/10.1515/9780748629909>.
- Geßner, Salomon. *Idyllen*. Gessner, 1756, <https://books.google.de/books?id=m1pdAAAAcAAJ>.
- Gray, Jeffrey Alan, and Neil MacNaughton. *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Second, Oxford University Press, 2003.
- Grisot, Giulia, et al. *Sentiment Lexicons or BERT? A Comparison of Sentiment Analysis Approaches and Their Performance*. 2022.
- Gsteiger, Manfred. *Nationalliteratur*. June 2009, <https://hls-dhs-dss.ch/de/articles/011214/2009-09-24/>.
- Haller, Albrecht von. *Die Alpen*. Versuch Schweizerischer Gedichte, Göttingen, 1729.
- Haupt, Sabine. "'Vom Topos kultureller Selbstbehauptung zur Höflichkeitsformel.' Schweizer Literatur 'Und Ihre Diskursgeschichte'." *Caduff/Sorg*, 2004, pp. 191–217.
- Hawkins, Harriet. "Picturing Landscape." *The Routledge Companion to Landscape Studies*, Routledge, 2018, pp. 206–14, <https://doi.org/10.4324/9781315195063-16>.
- Herrmann, Berenike, and Jana Lüdtke. *A Fairy Tale Gold Standard. Annotation and Analysis of Emotions in the Children's and Household Tales by the Brothers Grimm*. 2022, https://doi.org/10.17175/2023_005.
- Herrmann, J. Berenike, et al. "Ein großer Berg Daten? Zur bibliothekswissenschaftlichen Dimension des korpusliteraturwissenschaftlichen Digital Humanities-Projekts 'High Mountains – Deutschschweizer Erzählliteratur 1880–1930.'" *Digital Humanities und wissenschaftliche Bibliotheken*, vol. 8, no. 1, Apr. 2021, <https://doi.org/10.21428/1bfadeb6.6e2feff6>.
- Herrmann, J. Berenike, and Giulia Grisot. "Lieblingsgegenden, Fenster und Mauern." *Zur emotionalen Enkodierung von Raum in Deutschschweizer Prosa zwischen 1850 und 1930*, 2022, <https://zenodo.org/record/6328025>.
- Heuser, Ryan, et al. "Mapping the Emotions of London in Fiction, 1700–1900: A Crowdsourcing Experiment." *Literary Mapping in the Digital Age*, Routledge, 2016, pp. 43–64, <https://doi.org/10.4324/9781315592596-12>.
- Hillebrandt, Claudia. "Das emotionale Wirkungspotenzial von Erzähltexten." *Akademie Verlag*, 2012, <https://doi.org/10.1524/9783050057163>.
- Hogan, Patrick Colm. "What Literature Teaches Us About Emotion: Synthesizing Affective Science and Literary Study." *The Oxford Handbook of Cognitive Literary Studies*, 2015.
- Hunger, Nik, et al. *Schweizerisches Idiotikon*. 2021, <https://www.idiotikon.ch>.
- Izard, Carroll E. *The Face of Emotion*. Irvington, 1982.

- Jacobs, Arthur M. "Sentiment Analysis for Words and Fiction Characters from the Perspective of Computational (Neuro-)Poetics." *Frontiers in Robotics and AI*, vol. 6, July 2019, p. 53, <https://doi.org/10.3389/frobt.2019.00053>.
- . "Sentiment Analysis of Children and Youth Literature: Is There a Pollyanna Effect?" *Frontiers in Psychology*, vol. 11, Sept. 2020, p. 574746, <https://doi.org/10.3389/fpsyg.2020.574746>.
- Jacobs, Arthur M., and Annette Kinder. "Computing the Affective-Aesthetic Potential of Literary Texts." *AI*, vol. 1, no. 1, Sept. 2019, pp. 11–27, <https://doi.org/10.3390/ai1010002>.
- . "What Makes a Metaphor Literary? Answers from Two Computational Studies." *Metaphor and Symbol*, vol. 33, no. 2, Apr. 2018, pp. 85–100, <https://doi.org/10.1080/10926488.2018.1434943>.
- James, William. "What Is an Emotion?" *Mind (Old Series)*, vol. 9, no. 34, 1884, pp. 188–205, <https://doi.org/10.1093/mind/os-ix.34.188>.
- Jockers, M. "Extracts Sentiment and Sentiment-Derived Plot Arcs from Text." *R Package "Syuzhet"*, 2017.
- Jockers, Matthew L., and Rosamund Thalken. *Text Analysis with R for Students of Literature*. Springer International Publishing, 2014, <https://doi.org/10.1007/978-3-319-03164-4>.
- Jockers, Matthew L., and Ted Underwood. "Text-Mining the Humanities." *A New Companion to Digital Humanities*, Nov. 2015, pp. 291–306, <https://doi.org/10.1002/9781118680605.ch20>.
- Kakkonen, Tuomo, and Gordana Galić Kakkonen. "SentiProfiler: Creating Comparable Visual Profiles of Sentimental Content in Texts." *Proceedings of the Workshop on Language Technologies for Digital Humanities and Cultural Heritage*, 2011, pp. 62–69.
- Kanske, Philipp, and Sonja A. Kotz. "Leipzig Affective Norms for German: A Reliability Study." *Behavior Research Methods*, vol. 42, no. 4, Nov. 2010, pp. 987–91, <https://doi.org/10.3758/brm.42.4.987>.
- Kern, Bettina M. J., et al. "A Review and Cluster Analysis of German Polarity Resources for Sentiment Analysis." *3rd Conference on Language, Data and Knowledge (LDK 2021)*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.
- Kim, Evgeny, et al. "Prototypical Emotion Developments in Literary Genres." *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, 2017, pp. 17–26.
- Kim, Young-Suk Grace, et al. "Unpacking Eye Movements during Oral and Silent Reading and Their Relations to Reading Proficiency in Beginning Readers." *Contemporary Educational Psychology*, vol. 58, July 2019, pp. 102–20, <https://doi.org/10.1016/j.cedpsych.2019.03.002>.
- Klinger, Roman, et al. "Automatic Emotion Detection for Quantitative Literary Studies. A Case Study Based on Franz Kafka's 'Das Schloss' and 'Amerika.'" *Proceedings of the Digital Humanities*, 2016.
- Kuznetsova, Alexandra, et al. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, vol. 82, no. 13, 2017, pp. 1–26, <https://doi.org/10.18637/jss.v082.i13>.
- Lafond-Kettlitz, Barbara. "Die Alpen in Literatur und Malerei. Albrecht von Haller, Caspar Wolf, Ludwig Hohl, Ferdinand Hodler." *Études Germaniques*, vol. 256, no. 4, 2009, pp. 933–53, <https://doi.org/10.3917/eger.256.0933>.
- Lehnert, Gertrud. *Raum und Gefühl: der Spatial Turn und die neue Emotionsforschung*. transcript Verlag, 2014.
- Liu, Bing. "Many Facets of Sentiment Analysis." *A Practical Guide to Sentiment Analysis*, Springer, 2017, pp. 11–39, https://doi.org/10.1007/978-3-319-55394-8_2.
- McDougall, William. *An Introduction to Social Psychology*. Psychology Press, 2015, <https://doi.org/10.4324/9781315724256>.

- Menninghaus, Winfried, et al. "Aesthetic Emotions Are a Key Factor in Aesthetic Evaluation: Reply to Skov and Nadal (2020)." *Psychological Review*, vol. 127, no. 4, July 2020, pp. 650–54, <https://doi.org/10.1037/rev0000213>.
- Mohammad, Saif M. "From Once upon a Time to Happily Ever after: Tracking Emotions in Mail and Books." *Decision Support Systems*, vol. 53, no. 4, Nov. 2012, pp. 730–41, <https://doi.org/10.1016/j.dss.2012.05.030>.
- . "Imagisaurus: An Interactive Visualizer of Valence and Emotion in the Roget's Thesaurus." *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2015, pp. 85–91.
- . *Sentiment Analysis of Mail and Books*. National Research Council Canada, 2011.
- Moretti, Franco. *Atlas of the European Novel: 1800-1900*. Verso Books, 1999, https://books.google.co.uk/books?id=ja2MUXS_YQUC.
- . *Graphs, Maps, Trees: Abstract Models for Literary History*. Verso, 2007.
- Mowrer, Orval, and Hobart C. N. "Learning Theory and Behavior." *R. E. Krieger Pub. Co*, 1973, p. 555.
- Muschg, Adolf. "Gibt es eine schweizerische Nationalliteratur?" *Jahrbuch Der Deutschen Akademie für Sprache Und Dichtung*, Wallstein Verlag, 1980.
- Naber, Daniel. *Openthesaurus.de*. 2020, <https://www.openthesaurus.de/>.
- Nakov, Preslav, et al. "Developing a Successful SemEval Task in Sentiment Analysis of Twitter and Other Social Media Texts." *Language Resources and Evaluation*, vol. 50, no. 1, Jan. 2016, pp. 35–65, <https://doi.org/10.1007/s10579-015-9328-1>.
- Nell, Werner, and Marc Weiland, editors. *Gutes Leben auf dem Land? Imaginationen und Projektionen vom 18. Jahrhundert bis zur Gegenwart*. Transcript, 2021, <https://doi.org/10.1515/9783839454251>.
- Nünning, Ansgar. "Formen und Funktionen literarischer Raumdarstellung: Grundlagen, Ansätze, narratologische Kategorien und neue Perspektiven." *Raum Und Bewegung in Der Literatur*, edited by Wolfgang Hallet and Birgit Neumann, transcript Verlag, 2009, pp. 33–52, <https://doi.org/10.1515/9783839411360-002/html>.
- Oatley, Keith, and P. N. Johnson-Laird. "Towards a Cognitive Theory of Emotions." *Cognition & Emotion*, vol. 1, no. 1, Mar. 1987, pp. 29–50, <https://doi.org/10.1080/02699938708408362>.
- Oatley, Keith, and Philip. N. Johnson-Laird. "Emotion and Reasoning to Consistency." *Advances in Consciousness Research*, vol. 44, 2002, pp. 157–82, <https://doi.org/10.1075/aicr.44.07oat>.
- Ortony, Andrew, and Terence J. Turner. "What's Basic about Basic Emotions?" *Psychological Review*, vol. 97, no. 3, 1990, pp. 315–31, <https://doi.org/10.1037/0033-295x.97.3.315>.
- Pang, Bo, and Lillian Lee. "Opinion Mining and Sentiment Analysis." *Comput. Linguist*, vol. 35, no. 2, 2009, pp. 311–12.
- Panksepp, Jaak. "Toward a General Psychobiological Theory of Emotions." *Behavioral and Brain Sciences*, vol. 5, no. 3, Sept. 1982, pp. 407–22, <https://doi.org/10.1017/s0140525x00012759>.
- Parrott, W. Gerrod, editor. *Emotions in Social Psychology: Essential Readings*. Psychology Press, 2001.
- Piatti, Barbara, Hans Rudolf Bär, Anne-Kathrin Reuschel, and Lorenz Hurni. "Die Geographie der Fiktion — Das Projekt „Ein literarischer Atlas Europas“." *KN - Journal of Cartography and Geographic Information*, vol. 58, no. 6, Nov. 2008, pp. 287–94, <https://doi.org/10.1007/bf03543995>.
- . "Mapping Literature: Towards a Geography of Fiction." *Cartography and Art*, Springer Berlin Heidelberg, 2009, pp. 1–16, https://doi.org/10.1007/978-3-540-68569-2_15.

- Piatti, Barbara, and Thomas Streifeneder. "Rural Criticism: Oder: Welche Erzählungen über das Drama des ländlichen Raums brauchen wir dringend? Ein Plädoyer." *Rurale Topografien*, edited by Werner Nell and Marc Weiland, 1st ed., vol. 12, transcript Verlag, 2020, pp. 199–238, <https://doi.org/10.14361/9783839454251-006>.
- Plutchik, Robert. "A General Psychoevolutionary Theory of Emotion." *Theories of Emotion*, Elsevier, 1980, pp. 3–33, <https://doi.org/10.1016/b978-0-12-558701-3.50007-7>.
- Pries, Christine. *Das Erhabene: Zwischen Grenzerfahrung Und Größenwahn*. Akademie Verlag, 2015, <https://doi.org/10.1515/9783050082899/html>.
- Projekt Gutenberg DE. *Projekt Gutenberg-DE*. 2021, <https://www.projekt-gutenberg.org/>.
- R Core Team. *R: A Language and Environment for Statistical Computing*. 2021, <https://www.r-project.org/>.
- Reagan, Andrew J., et al. "The Emotional Arcs of Stories Are Dominated by Six Basic Shapes." *EPJ Data Science*, vol. 5, no. 1, Nov. 2016, p. 31, <https://doi.org/10.1140/epjds/s13688-016-0093-1>.
- Rehm, Stefan. *Stadt/Land: eine Raumkonfiguration in Literatur und Film der Weimarer Republik*. Ergon-Verl, 2015.
- Remus, Robert, et al. "SentiWS-A Publicly Available German-Language Resource for Sentiment Analysis." *LREC*, 2010.
- Rousseau, J. J. "Émile, ou De l'éducation." *Jean Néaulme*, 1762, <https://books.google.co.uk/books?id=Elu6g0qpyMMC>.
- Russell, James A. "A Circumplex Model of Affect." *Journal of Personality and Social Psychology*, vol. 39, no. 6, Dec. 1980, pp. 1161–78, <https://doi.org/10.1037/h0077714>.
- . "Emotion, Core Affect, and Psychological Construction." *Cognition & Emotion*, vol. 23, no. 7, Nov. 2009, pp. 1259–83, <https://doi.org/10.1080/02699930902809375>.
- Rusterholz, P., and A. Solbach. *Schweizer Literaturgeschichte*. J.B. Metzler, 2016, <https://books.google.co.uk/books?id=BYBJDQAAQBAJ>.
- Scaramellini, Guglielmo. "The Picturesque and the Sublime in Nature and the Landscape: Writing and Iconography in the Romantic Voyaging in the Alps." *GeoJournal*, vol. 38, no. 1, Jan. 1996, pp. 49–57, <https://doi.org/10.1007/bf00209119>.
- Scherer, Klaus R. "What Are Emotions? And How Can They Be Measured?" *Social Science Information*, vol. 44, no. 4, Dec. 2005, pp. 695–729, <https://doi.org/10.1177/0539018405058216>.
- Schiller, Friedrich. *Wilhelm Tell: Schauspiel ... zum Neujahrsgeschenk auf 1805*. Cotta, 1804.
- Schindler, Ines, et al. "Measuring Aesthetic Emotions: A Review of the Literature and a New Assessment Tool." *PloS One*, vol. 12, no. 6, June 2017, p. e0178899, <https://doi.org/10.1371/journal.pone.0178899>.
- Schmidt, Thomas, et al. "Herausforderungen für Sentiment Analysis bei literarischen Texten." *InfDH-Workshop 2018 – Im Spannungsfeld Zwischen Tool-Building Und Forschung Auf Augenhöhe – Informatik Und Die Digital Humanities*, 2018, <https://doi.org/10.18420/INFDH2018-16>.
- Schmidt, Thomas, Johanna Dangel, et al. *SentText: A Tool for Lexicon-Based Sentiment Analysis in Digital Humanities*. 2021.
- Schöch, Christof, et al. "Creating the European Literary Text Collection (ELTeC): Challenges and Perspectives." *Modern Languages Open*, vol. 25, no. 1, 2021, <https://doi.org/10.3828/mlo.v0i0.364>.
- Schumacher, Mareike K. *Orte und Räume im Roman: ein beitrag zur digitalen literaturwissenschaft*. J B METZLER, 2023, <https://doi.org/10.1007/978-3-662-66035-5>.

- Scott, Alister. "Assessing Public Perception of Landscape: From Practice to Policy." *Journal of Environmental Policy & Planning*, vol. 5, no. 2, June 2003, pp. 123–44, <https://doi.org/10.1080/1523908032000121193>.
- . "Assessing Public Perception of Landscape: The LANDMAP Experience." *Landscape Research*, vol. 27, no. 3, July 2002, pp. 271–95, <https://doi.org/10.1080/01426390220149520>.
- Sengle, Friedrich. "Wunschbild Land und Schreckbild Stadt." *Studium Generale*, vol. 16, no. 10, 1963, pp. 619–30.
- Spyri, Johanna. *Heidi's Lehr- und Wanderjahre: eine Geschichte für Kinder und auch für Solche, welche die Kinder lieb haben*. Perthes, 1880.
- Stamm, Nadine. *Klassifikation und Analyse von Emotionswörtern in Tweets für die Sentimentanalyse*. Universität Zürich, 2014.
- Taboada, Maite, et al. "Lexicon-Based Methods for Sentiment Analysis." *Computational Linguistics*, vol. 37, no. 2, June 2011, pp. 267–307, https://doi.org/10.1162/coli_a_00049.
- Tally, Robert T. "The Space of the Novel." *The Cambridge Companion to the Novel*, June 2018, pp. 152–67, <https://doi.org/10.1017/9781316659694.011>.
- Taylor, Joanna E., et al. *Mapping Digitally, Mapping Deep: Exploring Digital Literary Geographies*. 2018.
- Tomkins, Silvan. "Affect Theory." *Approaches to Emotion*, edited by Klaus R. Scherer and Paul Ekman, Erlbaum, 1984, pp. 163–95.
- Troscianko, Emily T. "Cognitive Perspectives on Immersion." *Immersion and the Storyworld*, 2012.
- Tschopp, Silvia S. "Die Geburt der Nation aus dem Geist der Geschichte: historische Dichtung Schweizer Autoren des 19." *Jahrhunderts*, 2004.
- Turney, Peter D., and Michael L. Littman. "Learning Analogies and Semantic Relations." *arXiv*, Nov. 2003, <http://arxiv.org/abs/cs/0307055>.
- Võ, Melissa L. H., et al. "The Berlin Affective Word List Reloaded (BAWL-R)." *Behavior Research Methods*, vol. 41, no. 2, Sept. 2009, pp. 534–38, <https://doi.org/10.3758/brm.41.2.534>.
- Waltinger, Ulli. "GERMANPOLARITYCLUES: A Lexical Resource for German Sentiment Analysis." *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC), Electronic Proceedings*, 2010.
- Wartmann, Flurina M., et al. "Describing and Comparing Landscapes Using Tags, Texts, and Free Lists: An Interdisciplinary Approach." *International Journal of Geographical Information Science*, vol. 32, no. 8, Apr. 2018, pp. 1572–92, <https://doi.org/10.1080/13658816.2018.1445257>.
- Watson, John B. *Behaviorism*. W.W. Norton, Incorporated, 1930, <https://books.google.ch/books?id=2W9OMAEACAAJ>.
- Wege, Carl. *Buchstabe und Maschine: Beschreibung einer Allianz*. Aufl., Suhrkamp, 2000.
- Weiner, Bernard, and Sandra Graham. "An Attributional Approach to Emotional Development." *Emotions, Cognition, and Behavior*, 1984, pp. 167–91.
- Wick, Mark, and Christophe Boutreux. *Geonames.Org*. 2020, <http://www.geonames.org/>.
- Wijffels, Jan. *udpipe: Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the 'UDPipe' 'NLP' Toolkit*. 2020, <https://cran.r-project.org/package=udpipe>.
- Wiki and Wikimedia Foundation Governance. *Wikimedia Foundation Governance Wiki*. 2021, <https://foundation.wikimedia.org/w/index.php?title=Home&oldid=130586>.
- Wilkens, Matthew. "Too Isolated, Too Insular: American Literature and the World." *Journal of Cultural Analytics*, vol. 6, no. 3, June 2021, p. 25273, <https://doi.org/10.22148/001c.25273>.

- Winko, Simone. *Kodierte Gefühle: Zu einer Poetik der Emotionen in lyrischen und poetologischen Texten um 1900*. Erich Schmidt Verlag GmbH & Co KG, 2003.
- Zehe, Albin, et al. "Prediction of Happy Endings in German Novels Based on Sentiment Information." *3rd Workshop on Interactions Between Data Mining and Natural Language Processing*, Riva Del Garda, 2016, p. 9.
- Zimmer, Oliver. "In Search of Natural Identity: Alpine Landscape and the Reconstruction of the Swiss Nation." *Comparative Studies in Society and History*, vol. 40, no. 4, Oct. 1998, pp. 637–65, <https://doi.org/10.1017/s0010417598001686>.