Unhappy Texts?
A Gendered and Computational Rereading of The Modern Breakthrough
Nielsen Degn, Kirstine; Bjerring-Hansen, Jens; Al-Laith, Ali Mohammed Ali; Hershcovich, Daniel

Published in:
Scandinavian Studies

Publication date:
2025

Document version
Peer reviewed version

Citation for published version (APA):
Unhappy Texts? A Gendered and Computational Rereading of The Modern Breakthrough

Kirstine Nielsen Degn, Jens Bjerring-Hansen, Ali Al-Laith, and Daniel Hershcovich

Abstract

Our article discusses and contends a well-rooted hypothesis that the texts of women writers of the Modern Breakthrough in Scandinavia were particularly unhappy. We examine this claim, along with some of the quantitative and qualitative issues it raises: Does this correlation of gender and affect hold true for the entire spectrum of female literary production from the era? What about male authorship and its affectivity? And what does ‘unhappy’ even mean? The hypothesis and the accompanying questions are confronted through two interventions. The first is a quantification made possible by new digital archives and methodologies which allow for a radical upscaling of the investigation’s empirical foundation. Our second intervention is to approach the nineteenth century texts with a theoretical framework from the fields of gender studies and affect theory. Our findings are the following: (1) The thesis of the unhappy text appears partially true, but, importantly, women are even more overrepresented among the positive texts. (2) The affect category of neutrality, however, is more significant. Neutrality turns out to be a male, canonical enterprise, while low neutrality is primarily associated with forgotten or neglected female authors. The most crucial gender bias in the affective economy of the texts is the lack of neutrality in literature by women. (3) Finally, this and other biases point to clear intersectional dynamics between the author’s gender, the affective qualities and quantities of the texts, and their social status.

1. Introduction

Were the texts of women writers of the Modern Breakthrough in Scandinavia really steeped in unhappiness? It has been argued that this was the case. In this article, we will examine this hypothesis of the unhappy female text more closely, along with some of the quantitative and qualitative issues it raises: Does this correlation of gender and affect hold true for the entire spectrum of female literary production from the era? What about male authorship and its affectivity? And what does ‘unhappy’ even mean? We will challenge and discuss the hypothesis and the accompanying questions through two interventions. The first is a quantification made possible by new digital archives and methodologies which allow for a radical upscaling of the investigation’s empirical foundation. Whilst the hypothesis of the unhappy texts is based solely on texts authored by women and often on relatively few textual instances, this upscaling will also allow us to include the works authored by men in the investigation. Our second intervention is to approach the nineteenth century texts with a theoretical framework from the fields of gender studies and affect theory. Central to us is the affect-theoretical understanding of affects as a collective and material-discursive phenomenon, as opposed to the psychoanalytic understanding of emotions as a phenomenon taking place in the individual psyche.

In our work, we not only comment on, but also rely on scholarship on women’s literature of the Modern Breakthrough that is both insightful and impactful. All the more
reason why there is a need for new vantage points on shared concerns on a largely overlooked body of texts from a significant literary period.¹

2. The Modern Breakthrough and the Unhappy Text

With *Det moderne Gjennembruds Mænd* (1883; The Men of the Modern Breakthrough), the prominent critic Georg Brandes coined both a term and a literary perspective that has had a long-lasting impact on the historiography of the literary landscape of late nineteenth century Scandinavia. With Brandes’ telling title as the first example, the historiography of the Modern Breakthrough has primarily focused on a few canonical male authors. In line with this, Georg’s brother Edvard Brandes, also an influential critic, believed that the so-called “Damefortællinger” [Ladies’ stories] were of fluctuating quality;

(...) men alle med det samme forsagende Præg. Det maa man lade vore skrivende Damer, at mandhaftige ere de ikke, man kan vride disse Bøger som man vil, man faar ikke andet end Resignation ud af dem. (Brandes 1876, 308).

(but all with the same renouncing character. You must admit it to our writing ladies, they are not mannish, you can wring these books as you like, you get nothing but resignation out of them.)

A generation later, the influential literary scholar Vilhelm Andersen echoed this rendering of ‘women’s literature’ as something distinct from ‘ordinary literature’. Andersen dubbed the period’s female literary production as “en særegen Provins af Kulturen” [a peculiar province of culture] (Andersen 1925, 711). In 2005 the compact and canonical literary history *Hovedsporet* (2005; The main track) brought the marginalization of female authors into the twenty-first century. Here the debates on gender and sexuality of the Modern Breakthrough are depicted under the heading “Kvindelitteratur – skrevet af mænd” [Women’s literature – written by men] building on the notion that primarily male authors produced significant literary works about women’s conditions (Kondrup 2005, 422–423).

In large parts of the literary historiography of the Modern Breakthrough, Georg Brandes’ focus on the men of the period has evidently persisted. However, this focus does not reflect the diversity of the period’s literary production, with women being most noticeable among the new groups of authors stepping forward from 1870–1900. In the 1970s and 1980s, several literary scholars made substantial efforts to reinterpret the period and cast light on these female writers, most prominently Pil Dahlerup, who in a direct polemic with Brandes’ malecentred 1883-book, exactly 100 years later, published her doctoral dissertation *Det moderne gennembruds kvinder* (1983; The Women of the Modern Breakthrough). In the dissertation, Dahlerup gives a new account of the period’s female authors, which has been pervasive in many of the subsequent literary histories that have sought to cast light on the women of the Modern Breakthrough. With a theoretical framework drawing on psychoanalysis and structuralism, Dahlerup argues that, in the female-authored works from

¹ We wish to thank The Carlsberg Foundation and The Data+ Pool at the University of Copenhagen for supporting the research underlying this article as well as Tobias Skiveren, Christine Hamm and the anonymous reviewers for valuable comments on earlier versions of the article.
the Modern Breakthrough, the positions of enunciation must be understood as functions with different forms of relations to the patriarchy: daughters, ‘sons’, mothers, mistresses, wives or emancipated women (Dahlerup 1984, 471). According to Dahlerup, the societal circumstances are most noticeable in the period’s female authorships where depressions associated with father figures – either the actual biological father or another representative of the patriarchy as a system – are common features (Dahlerup 1984, 281). Thus, Dahlerup argues, the female authorships of the period centers around a disillusionment of women’s actual living conditions and a resulting depressive state (Dahlerup 1984, 281–283).

In line with Dahlerup’s rendering, the women of the Modern Breakthrough are depicted under the heading “Den Ulykkelige Tekst” [The Unhappy Text] in Nordisk kvindelitteraturhistorie (Møller-Jensen 1993, 328; Nordic women’s literary history), and in Danske kvindelige forfattere (1982; Danish female authors) the female authors’ relation to the new social realist and naturalist literature is described as a ‘female authoritarian masochism’ (Dalager and Mai 1982, 16). This notion of the writings of the women of the Modern Breakthrough as unhappy texts recurs in the most recent historical accounts of the period’s literature. In an essay on ‘hysteria’ in Danish literature Camilla Schwartz describes the Modern Breakthrough’s depiction of female characters as texts that are “mørke og dystre og ofte ender med selvmord” [dark and gloomy and often end in suicide] (Schwartz 2022, 330). Along similar lines, Annegret Heitmann maps “hvordan angsten tematiseres hos en række af periodens kvindelige forfattere” [how angst is thematized by a number of the period’s female writers] in a contribution to a literary history of ‘angst’ in Danish literature (Christensen and Dam 2023, 17).

An underlying premise of the unhappy text as something attached to a distinctive experience of womanhood is that the male authors must be writing if not happy texts, then at least texts that are not unhappy! The premise is articulated explicitly by Dahlerup who argues that whilst the female writers articulate the female experience as a delimited and depressive position, it is in the male-authored works we find the liberated types of women such as J.P. Jacobsen’s Marie Grubbe, Henrik Ibsen’s Mrs. Alving and Henrik Pontoppidan’s Jacobe. According to Dahlerup, we only find these empowered female figures in the period’s male-authored works because they were the ones who themselves had the external and internal opportunities to challenge the norms. They were free themselves and consequently they were able to convincingly portray liberated fictional characters (Dahlerup 1984, 349).

The notion of the female literary production of the Modern Breakthrough as consisting of tendentially unhappy texts—and vice versa the unspoken notion of the male literary production as made up of mostly happy texts—evidently persists well into the most recent literary historiography of the period. We will argue that this notion has a great impact on the way we perceive this significant period in Scandinavian history and thus is worth revisiting. In this article, we will do this with two points of departure, the first concerning scale and empirical selection, and the second concerning theoretical framework and methodology.
3. Measuring Emotions

Reading at Scale

Our explorations of the emotional qualities and their distribution in the literature of the Modern Breakthrough are based on a lot of text. We rely on a corpus comprising all original (i.e., non-translated and non-reissued) novels published in book form in Denmark from the last 30 years of the nineteenth century. This equals 839 novels by Danish and Norwegian authors or more than 50 million words in total (Bjerring-Hansen et al., 2022). As the output statistics in Figure 1 indicates, the Modern Breakthrough was also a breakthrough of the novel in Denmark and Norway, with the yearly output increasing from about 10 volumes to about 50 over a thirty-year period. Throughout the period, around 20 percent of the novels were produced by female authors.

Fig 1. The novel in the Modern Breakthrough. Male and female production over a thirty-year period from 1870 to 1899.

The corpus is the empirical foundation of the MeMo project: [https://nors.ku.dk/english/research/projects/measuring-modernity/](https://nors.ku.dk/english/research/projects/measuring-modernity/). It should be noted that, until 1907, written Norwegian was practically identical to written Danish (Vikør, 2022). In other words, we are dealing with a more or less linguistically coherent body of texts.

In line with other digital literary studies dealing with gender issues (for instance Cheng 2020 and Underwood et al. 2018) we employed a coding for the authors based on a tripartition: ‘female’, ‘male’, and, in the case of unknown anonymous authorship, ‘unknown’. Public gender representation in Scandinavia of the late nineteenth century was predominantly organized according to this scheme. Here, we rely on a Butlerian understanding of gender, namely as the conventional roles that individuals were expected to assume to be considered legitimate in social contexts (Butler 2002, 123–4, 178). This approach is used because the article's focus is precisely to explore gender as an affective, cultural, and historical category, not the authors’ personal gender identities or the processes that shape it. We will only address these questions indirectly in our highlighting of the fluid and unstable nature of the signification of public gender signals.
Also, the cultural diversity of the texts is pronounced. From a genre perspective, we have contemporary novels as well as historical novels and other forms of genre fiction such as romance, crime, and war stories (cf. Bjerring-Hansen and Rasmussen, 2023). And from an aesthetic perspective we have both avant-garde forms of realism, including naturalism and impressionism, and more traditional “it was a dark a stormy night” prose with a preference for abstract or generalized over concrete specification (Bjerring-Hansen and Wilkens, 2023).

In other words, the corpus allows – and is indeed intended – for capturing and exploring the “forgotten 99 percent” of literature in circulation, which Franco Moretti famously has drawn attention to in his studies of the British novel of the nineteenth century (Moretti 2000, 208). Beyond this opening up of the canon, a basic critical argument for digital literary studies is that computational techniques open the way for analyses of large textual archives to capture long-term and widespread literary trends. The important strategy here is quantification: the act of counting and measuring. As Matthew Wilkens has pointed to, literary scholars often underestimate “the extent to which their claims are implicitly quantitative, pattern-based, and dependent on reductive models of the texts they treat” (Wilkens 2015, 11). Arguments on literary trends or epochs, e.g., “the unhappy text” of the Modern Breakthrough in Scandinavia are built on what Wilkens calls “an abstractly quantifiable model” of a problem domain and of the texts’ place within it (Wilkens 2015, 11). Even without numbers, statistics, and graphs, they rely on considerations of the distribution of specific features across a range of relevant texts. What digital literary studies does and insists on doing is to make this procedure explicit in order, not to trump tradition, but to bring new perspectives and evidence to the table. In this our ideal is that critical knowledge can be reviewed through computational models and that computational models, in turn, can be evaluated against critical knowledge, as Andrew Piper and others have contended (Piper 2015).

Sentiment Analysis

Sentiment analysis is the quantification of the sentiments associated with the words of a text. The techniques, which have been an important tool in natural language processing, but also have found their way into Digital Humanities work, aim to compute, and analyze “people’s opinions, sentiments, appraisals, attitudes, and emotions towards entities and their attributes” (Liu 2015, 1). In this process, emotional content is automatically extracted from the texts and converted into machine-readable information such as labels (in our case ‘positive’, ‘negative’, and ‘neutral’ indicating sentiment quality or so-called ‘polarity’: what?) and numbers (in our case a score between 0 to 1 indicating sentiment intensity or so-called ‘valence’: how much?). These labels and numbers can be modeled statistically and visualized in graphs and plots. Traditionally, sentiment analysis has been applied to texts with a strong valence and subjectivity (for instance reviews or tweets). In recent years, however, different methods of detecting sentiments have been used on literary texts (see for instance Kim and Klinger 2019; Rebora, 2023).

It goes without saying that an automated analysis of the sentimental qualities of novels poses all kinds of challenges and caveats. We are dealing not only with long texts, but also imaginative literature infused with ambiguity, interpretational confounding, rhetorical
sophistication, and narrative layerings between author, narrator, and characters. Add to this the processes of language change and the dynamic nature of emotions. Some of these issues will be addressed along the way. That said, we believe that it is precisely the reductionism of the metric we use that allow us to see the novels of the Modern Breakthrough in a radical new light by linking—or delinking—our computational efforts to critical understandings of the characteristics of the period and more generally to notions of how literary texts work.

But how did we go about analyzing the sentiments in the novels? While the long answer can be read in a newly published methods paper (Al-Laith et al., 2023), the short answer is that we in a collaboration between computer scientists, linguists, and literary scholars developed a classifier based on a large multilingual language model (“XLM RoBERTa”, released by the Facebook AI team in 2019) after it had been fine-tuned on Danish tweets annotated for sentiment and, not least, a human-annotated dataset of the sentiment of almost 3000 sentences belonging to novels from our corpus. With an accuracy of 80%, meaning that the model in 4 out 5 instances classified the sentiment of the sentences from the annotated dataset in accordance with the (human) domain experts, this new model outperforms other models when it comes to nineteenth century prose in Danish and thus it was used for an automatic prediction of sentiments of the whole corpus, whereby the more than 3,2 million sentences were processed by the algorithm and annotated as either positive, negative, or neutral.

In developing, training, and validating an accurate model, the human touch, in form of annotation, is crucial, while the process from the point-of-view of the three literary scholars acting as annotators, confronted with 3 to 4 randomly picked sentences from each of the 839 novels, also provided broad insight to the diverse make-up of the corpus. We decided to work on sentence-level as a comprise avoiding cruder alternatives: text-level, in which case the polarity of an entire novel should be determined and wrapped up in a single label, or word-level, which would require a sentiment lexicon for nineteenth century Danish and, more importantly, too often miss out on the subtle ways in which sentiment is expressed in language and interacts with sentence structure.4

The annotation task can be illustrated by pointing to two of the central challenges we faced. The first is connected to semantic change as exemplified in one of the snippets from the corpus: “Hun blev strax noget fortumlet over disse uforberedte Kjærtegn; men da hun laa i hans Arme, saa’ hun op paa ham med et besynderligt, ikke fornærmet Blik” (English translation: She was at once somewhat taken aback by these unexpected caresses; but as she lay in his arms, she looked up at him with a curious, not offended, look). Since the adjective ‘besynderlig’ in the nineteenth century also had the meaning ‘special’ or ‘curious,’ besides the dominant modern connotation ‘weird,’ this sample should not be labeled as negative, but rather as positive. Hence, the language model is implicitly trained to deal with issues of semantic change without this explicit expert knowledge, in NLP jargon, word-sense disambiguation (WSD). The second challenge is related to social change and specifically involves gender and gender roles. An awareness of the fact that cultural values and norms change over time is crucial for annotation and the usefulness of sentiment classification of

4 The annotation guidelines as well as an array of interpretational and methodological challenges are documented and discussed in Author, 2023, 327.
historical texts. We paid special attention to text segments involving gendered behavior and dialogue, by relying on fundamental insights from social constructivist gender theory, understanding gender as a category subject to change in form and meaning over time. From this follows that a modern understanding of gender roles could potentially come at odds with the intentions in nineteenth century texts. While on the one hand dominant behavior from a male character, e.g., the use of grammatical imperative (“Stop!”) would tend to be understood more negatively today than in the past, vice versa, female passivity is probably more likely to be conceived as tendentially negative than positive. The latter can be exemplified by a sample involving a female character—passive, dreamful and nostalgic—which we decided to label as positive: “Men allerhelst laa hun dog i Vinteraftenernes Skumring i sin Yndlingsstilling i Armstolen og grublede og drømte og ventede og ventede —ligesom Prinsessen i Eventyret.” (Translation: But most of all, in the twilight of the winter evenings she would lie down in her favorite position in the armchair, pondering and dreaming and waiting and waiting — like the princess in the fairytale.). Awareness of both linguistic and cultural change is pivotal for our classifier (as well as more generally for the development of useful AI models for computational explorations of a low-resource domain such as Scandinavian literary history).

Affect Theory

The successes of sentiment analysis within Digital Humanities in recent years are paralleled by a similar rising interest in emotions in several other disciplines (Keen 2011, 7). In the humanities and social sciences, the increased interest in emotions or affects has even been dubbed the “affective turn” (Clough and Halley 2007). Within literary studies, explorations of the connections between emotions and literature have a long history, not limited to this most recent turn towards feelings and affects. For one, Georg Brandes thought of literature as a depiction of collective emotions: “When a people’s literature is fully developed, it represents the entire history of that population’s thoughts and feelings.” (Brandes 2017 [1872], 699). In more recent times multiple scholars have also argued for this connectedness. In a positivistic vein similar to Brandes, Martha C. Nussbaum claims that literature is a particularly privileged art form because it gives access to people’s inner emotional and intellectual life (Nussbaum 2016, 26). Analogous to this, Sianne Ngai has highlighted the aesthetic domain as “the ideal site to examine the politically ambiguous work of negative emotions.” (Ngai 2005, 6). Similarly, Rita Felski has argued that literary works have a special ability to depict a social phenomenology as literary works are concerned with how particular lifeworlds are felt and experienced (Felski 2008, 89, 92).

When we set out to juxtapose sentiment analysis, literary studies and affect theory it is not an unlikely or ill-fitting endeavor, in other words, but instead one among other ways of trying to get hold of and make it possible to study how emotions are at work in literary works. But what exactly are we looking for when reading for emotions at scale? In the following, we will elaborate on how we try to make use of a conceptual framework from affect theory to answer this question.

---

5 In this, we took crucial inspiration from Joan W. Scott’s middle range theory-approach and foucauldian criticism of essentialist and deterministic notions regarding gender as a category of historical analysis (Scott 1986, 1066–7, 1074).
Affect theory springs out from a turn against the notion of emotions as ‘expressive’. Sara Ahmed has argued that emotions are not a private matter that originates from within the individual psyche and moves outwards (Ahmed 2004, 117). Instead, Ahmed understands emotions as cultural practices rather than psychological states (Ahmed 2014, 194). Emotions are thus regarded as collective phenomena which are embedded in and constituted by a historical, cultural, and social context (Ahmed 2004, 126). It is this embeddedness that we seek to investigate. In Ahmed’s terminology we set out to track “how emotions circulate between bodies, examining how they ‘stick’ as well as move.” (Ahmed 2014, 4). In this respect, we want to examine how and to what extent emotions have ‘stuck’ to both literary and social categories. How do emotions circulate, stick, and move in relation to for instance issues concerning gender, narratology, canonicity, and popular culture?

This cultural understanding of emotions makes literary works an obvious object of interest. Yet, the endeavor to examine “how emotions circulate between bodies” (Ahmed 2014: 4) calls for attention to the difference between the ‘bodies’ found in the extra-textual world and literary ‘bodies’. As Tobias Skiveren has highlighted, literary bodies are embedded in a ‘body of text’, and thus they can only come into view through the sensations that literature’s various “æstetiske kompositioner” [aesthetic compositions] offer (Skiveren 2020, 93).

It is therefore the aesthetic compositions by which the affects come into view that we will examine. In other words, we are interested in the affective mode or style of the text. In this respect we are, as outlined above, working with a demonstratively primitive typology of affective styles: positive, negative, and neutral. What we aim to do is not to map in detail how affects are at work in the individual texts, but to give a quantitative overview which hopefully will create new questions and open interesting dimensions of the relations between emotions, gender, and literary history, which will provide ground for new and further investigations.

4. Major Trends

To begin with, let us look at a raw and basic analysis of the sentimental qualities and their quantities in the corpus. In Figure 2, the distribution of sentiment levels is depicted in relation to the author’s gender and the percentage of sentences.
Fig 2. The big picture. The diagram shows the distribution of sentiments juxtaposed with gender. The X-axis identifies sentiment class. The Y-axis is the fraction of sentences per sentiment class.

We are not looking at individual novels, in other words, but at the corpus as a whole dispersed into six bags, according to two variables: sentiment class and gender. Each bag contains sentences that are classified as either positive, negative, or neutral and produced by either a male or a female author. We see that there are significant gender differences in all sentiment classes. Notably, according to our metric, on average, female authors tend to exhibit a higher proportion of both negative and positive sentiments compared to male authors, whereby the hypothesis about female authorships being particularly ‘unhappy’ is both confirmed and contradicted. This is not a result or a conclusion, but merely a modeling of the data, which calls for further tweaks and experiments to understand the dynamics and tensions at play.

An adjustment of the metrics that proves to be somewhat significant is based on formal and narratological considerations. The analysis in Figure 3 is identical to the basic experiment, but it draws only on the last 1% of the novels (e.g., the last three pages of a 300-page novel).
Fig. 3: Closure. This graph only consists of the endings of the novels, defined as the last percent of the text.

The inspiration comes from a long tradition of theoretical emphasis on narrative closure as, – from Aristotle’s poetics, to Frank Kermode’s *Sense of an Ending* (Kermode 1967), to contemporary cognitive narratology (see Klauk et al 2016) – which are often paired with a naive assumption that the emotional register is particularly distinct at the end of a novel. Indeed, a notable shift in polarity can be observed as the emotional polarity of the endings swings from neutrality to positivity. But seemingly there’s no gender effect at play. So, we have decided to work with whole texts rather than chunks. This is also more in line with our affect-theoretical starting points, where concepts such as closure, storyworld and disruption do not play a crucial role. Also, we are not interested in trying to mimic an analogue reading practice and, thereby, making the computational interventions too complicated and opaque.

An attempt to find a temporal logic behind the numbers proves insignificant and inconclusive as we see Figure 4, where year of publication is aligned with sentiment classification and author gender. There are no telling developments in either category. There are no signs of impact from particular contextual events, e.g., the so-called “Sædelighedsfejde” (The Nordic Sexual Morality Debate) in the 1880s, or time-specific aesthetic developments, such as the rise of realism (see Bjerring-Hansen and Wilkens, 2023). Seemingly, there’s no time-related plot in the literary history of emotions in the modern breakthrough. At least not according to our findings.

---

6 Affect-theoretical literary studies with a narratological orientation, tend to focus on text-internal affects at the character, narrative, or composition level rather than on the affective relationship between text and reader (see for instance Andersen 2015). On a more fundamental level Patrick Colm Hogan, for example, distinguishes between ‘Affect theory’ (a philosophical and cultural studies approach) and ‘Affect science’ (a cognitive approach) as two different directions within affect studies (Hogan 2016, 4).
In this iterative process of (re)building, (re)modeling, and (re)visualizing, we found it compelling, instead, to look at all the texts embedded in the same (heterogenous) cultural space or literary system of circulation. In the scatterplot in Figure 5 each dot represents a novel, whose position is defined by positive valence (x-axis) and negative valence (y-axis).
This visualization makes it even more clear how the female texts are characterized by a high degree of affect – and conversely how the male texts are characterized by a low degree of affect. The high affect quadrants are populated and partly dominated by women authors, which we remember only account for about 20% of the novels.

At the same time, together with the underlying data, the graph can be used to identify individual texts. In this regard, sampling of outliers suggests that the canonical texts are oriented away from the positive quadrants and vice versa. On the one hand, two novels by Amalie Skram, praised and criticized for her raw and “manly” naturalism (see Garton 2007, 329) are placed to the far left on the y-axis among novels from, equally non-positive and negative, male authors. On the other hand, to the far right, an adventurous frontier-novel like Adam Dan’s *Prærierosen* (1892, The Prairie Rose) is to be found in an affective domain of literature to a large extent dominated by female authorships.\(^7\) In other words, when identifying individual texts and looking at their relative position within our approximated affect system, we get clear hunches of what Sara Ahmed calls “sticky relations” between affects, gender and aesthetic positions.

5. Closer Readings and Sticky Relations

To further investigate these hunches, we will take a step closer to the individual texts as well as their intersectional positioning, by focusing on groupings of 50 novels at a time, instead of working with all 837 novels. We will show how the gender discrepancies become clearer, when zooming in on the most ‘extreme’ cases. But most importantly, the process allows us to isolate distinct aesthetic and social positions: advanced or highbrow literature on the one hand and popular or traditionally oriented literature on the other.

---

\(^7\) The novels by Amalie Skram are the two first volumes of her naturalist saga of the Hellemyr-family and its decay: *Sjur Gabriel* (1887), scoring 7% positive and 38% negative, and *To Venner* (1887), scoring 8% positive and 35% negative. Adam Dan’s *Prærierosen* scores 37% positive and 35% negative.
Figure 6 reveals several interesting correlations between gender, affects, canonicity, and popular culture. We will highlight four gendered biases: 1) Female authors are, in line with the unhappy text hypothesis, significantly overrepresented among the most negative texts accounting for 28%, compared to the only 17% they constitute of the corpus. 2) However, contrary to the hypothesis, female authors are even more significantly overrepresented among the most positive text accounting for 36% of the top 50. 3) Female authors are highly underrepresented among the most neutral texts, accounting for only 4% of the top 50. 4) Finally female authors are most significantly overrepresented among the least neutral texts accounting for 50% of the top 50, roughly three times as much as they make up in the corpus.

Intersectional Dynamics

Starting with the hypothesis of the unhappy texts, it is partially confirmed by our sentiment analysis. Female authors account for 28% of the texts and are thus overrepresented compared to their 17% representation in the total corpus. However, it is not the canonized authors who are often highlighted as the prime examples of the unhappy, dark, and angst-filled literature written by women who top the list. Looking closer at the works which make up the list, a quite different picture reveals itself than when reading literary histories. Authors such as Amalie Skram and Olivia Levision, which are some of the prime examples of ‘unhappy texts’, are not found in the top 50, and we have to scroll down to positions 24 and 38 to find Erna Juel-Hansen and Adda Ravnikilde. Instead, the most negative novels are religious works such as Theodora Mau’s *Deodata* (1871) and *Solstraalen* (1872; The Ray of Sunshine), and Laura Kieler's *I En Lysengels Skikkelse* (1892; In the Shape of an Angel of Light). Despite
the fact that these works based on a sentiment analysis like ours should be the primary examples in the representation of unhappy female texts, they are hardly mentioned in critical literature, and when they do turn up, they are deemed outside the interest of literary historiography as they are considered to be written in a retrospective and naïvely positive fashion (see, for example, Dahlerup 1984, 270 and Hjordt-Vetlesen and Mortensen 1993, 361). Our distant reading, however, indicates that something beyond naïve positivity and romanticism, is also at work in the religiously informed novels.

Turning to the most positive novels, the hypothesis of the unhappy text is challenged. Even more so than regarding negativity, female authors are overrepresented among the most positive novels, accounting for 36% of the top 50. Positivity was very much a female undertaking, too, so to say. However, even more pronounced than among the most negative texts, the focus on the most positive texts makes up a non-canonical top 50. Amongst the male authors, two canonical writers have found their way to the list: Peter Nansen with Maria (1894) and the fairy-tale writer H.C. Andersen with the late novel, Lykke-Peer (1870, translated to English as Lucky Peter), of whom Nansen is the only one who is usually thought of as a part of the Modern Breakthrough. The rest of the male authors and all of the female authors are either completely unknown or at least non-canonical authors.

Among the unknown authors, we have Camilla Ravn’s Interiører fra vore Bedsteforeldres og Forældres Tid (1899, Interiors from the Time of Our Grandparents and Parents) published under the pseudonym Ulrik Horn, and Emma Larsen’s Skæbnens Veje (1898, The Paths of Fate). Additionally, we find authors such as Adolphine Fogtmann, Christine Mønster, Louise Bjørnsen, Drude Krog Janson, and Theodora Mau, all of whom are mentioned in women’s literary histories but are far from being canonized. The shared attributes among these non-canonical female authors are that they are considered as religious, traditional, and/or naïvely romantic writers. Fogtmann’s form and content are described by Dahlerup as “så ukompliseret traditionelt og så småt, at det ikke kan sætte læseren i bevægelse (...) Epigonlitteratur.” [so uncomplicatedly traditional and so small that it cannot move readers. (...) Epigonal literature] (Dahlerup 1984, 146).8

In this way, the most positive texts are traditionally linked to a retrospective and idealistic style of writing. They are largely thought of as unmodern and rooted in a bygone romanticism and not of interest for literary historiographers studying the Modern Breakthrough. However, as our analysis shows, positivity should not be dismissed when studying the female writers, as the positive works make up a large part of the female literary production, and even may serve to highlight the sticky relations between affects and canonicity.

The Least Neutral Text

Whilst Figure 5 only indirectly accounts for neutrality, the process of zooming in on groupings of 50 texts at a time as displayed in Figure 6 and Table 1 and 2 allows us to

---

8 Theodora Mau is similarly listed by Dahlerup under the heading “Døtrene: De religiøse” [The Daughters: The religious] (Dahlerup 1984, 270), and Louise Bjørnsen is described as a gothic-inspired writer rooted in romanticism (Busk-Jensen 1993, 217–9). Both Mønster and Janson are also described as authors where religious themes and messages are central (respectively Iversen 1993, 422 and Hjordt-Vetlesen and Mortensen 1993, 367).
incorporate neutrality as an affective and stylistic feature into our analysis. Just from a brief glance at the neutrality-rank displayed in Table 1 and 2, it is clear that we can observe a pronounced gender and canonicity bias regarding neutrality. Female authors account for only 4% of the top 50 most neutral texts, while men account for 84%. Neutrality can consequently be said to be a male undertaking during the period. Note that this should not be understood as an inherent or essential masculine/feminine trait. Instead, it should be viewed as a specific historical and cultural example of how emotions have stuck to particular social categories, shaping conventional gender roles that individuals were expected to embody for social legitimacy. In this context, it is also worth noting that the top 50 primarily consists of canonical authors, particularly authors associated with modern realism. We find novels by Herman Bang in place 2, 3, and 5, while names like Jonas Lie, Bjørnstjerne Bjørnson, and Holger Drachmann are represented with a total of nine works. The only female author in the top 50 is the naturalist Amalie Skram, represented by two works. Contrary to the implicit hypothesis of the unhappy female texts, regarding male-authored works as ‘happy’ texts, it is not amongst the most positive novels, we find the male authors overrepresented, but amongst the most neutral texts.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Author</th>
<th>Title</th>
<th>Year</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kristian Gløersen</td>
<td><em>Fra Mit Friluftsliv</em></td>
<td>1881</td>
<td>69%</td>
</tr>
<tr>
<td>2</td>
<td>Herman Bang</td>
<td><em>Liv og Død</em></td>
<td>1899</td>
<td>65%</td>
</tr>
<tr>
<td>3</td>
<td>Herman Bang</td>
<td><em>Stuk</em></td>
<td>1887</td>
<td>63%</td>
</tr>
<tr>
<td>4</td>
<td>Laurits Johansen</td>
<td><em>Til Nordpolen pr Ballon</em></td>
<td>1894</td>
<td>62%</td>
</tr>
<tr>
<td>5</td>
<td>Unknown</td>
<td><em>Olidencymandens Frieri</em></td>
<td>1879</td>
<td>61%</td>
</tr>
<tr>
<td>6</td>
<td>Herman Bang</td>
<td><em>Stille Eksistenser</em></td>
<td>1886</td>
<td>61%</td>
</tr>
<tr>
<td>7</td>
<td>Alexander Thorsoe</td>
<td><em>Interiører Fra Det Danske Hof</em></td>
<td>1897</td>
<td>61%</td>
</tr>
<tr>
<td>8</td>
<td>Johannes Jørgensen</td>
<td><em>En Fremmed</em></td>
<td>1890</td>
<td>60%</td>
</tr>
<tr>
<td>9</td>
<td>Christian Krohg</td>
<td><em>En Duel</em></td>
<td>1888</td>
<td>60%</td>
</tr>
<tr>
<td>10</td>
<td>Christian Lundgaard</td>
<td><em>Don Juan</em></td>
<td>1896</td>
<td>59%</td>
</tr>
</tbody>
</table>

Table 1. The most neutral texts. The ten novels with the highest percentage of neutral sentiment.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Author</th>
<th>Title</th>
<th>Year</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>828</td>
<td>Rosalia Rosenfeld</td>
<td><em>Hvem af de to</em></td>
<td>1878</td>
<td>22%</td>
</tr>
<tr>
<td>829</td>
<td>Louise Bjørnsen</td>
<td><em>Skibladner</em></td>
<td>1883</td>
<td>22%</td>
</tr>
<tr>
<td>830</td>
<td>Theodora Mau</td>
<td><em>En Formynder</em></td>
<td>1874</td>
<td>22%</td>
</tr>
<tr>
<td>831</td>
<td>Leopold Budde</td>
<td><em>En Historie fra Fattige Seder</em></td>
<td>1874</td>
<td>22%</td>
</tr>
<tr>
<td>832</td>
<td>Drude Krog Janson</td>
<td><em>Mira</em></td>
<td>1897</td>
<td>21%</td>
</tr>
<tr>
<td>833</td>
<td>Otto Martin Møller</td>
<td><em>Af Elskovs Naade</em></td>
<td>1896</td>
<td>20%</td>
</tr>
<tr>
<td>834</td>
<td>Theodora Mau</td>
<td><em>Solstraalen</em></td>
<td>1872</td>
<td>19%</td>
</tr>
<tr>
<td>835</td>
<td>Anders J. Meldgaard</td>
<td><em>Pastor Brandt</em></td>
<td>1888</td>
<td>19%</td>
</tr>
<tr>
<td>836</td>
<td>Louise Bjørnsen</td>
<td><em>En Kvinde</em></td>
<td>1894</td>
<td>18%</td>
</tr>
<tr>
<td>837</td>
<td>Theodora Mau</td>
<td><em>Deodata</em></td>
<td>1871</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 2. The least neutral texts. The ten novels with the lowest percentage of neutral sentiment.
Parallel to this our sentiment analysis finds that female authors are significantly overrepresented among the least neutral novels, accounting for 50% of the top 50 compared to their mere 17% representation in the total corpus.

Again, it is a top 50 that predominantly consists of non-canonical works. We find a range of religious and traditional texts that were also present in the top 50 of the most positive texts, such as Theodora Mau, Adolphine Fogtmann, Louise Bjørnsen, and Drude Krog Janson. Additionally, we find authors like Magdalene Thoresen and Anna Margrethe Krebs, both of whom are described by Lise Busk-Jensen as conservative critics of realism (Busk-Jensen 2009, 58). We also find an author like Massi Bruhn, who despite being a rather central figure in the women’s movement (Mortensen 2023), is dismissed in *Nordisk kvindelitteraturhistorie* (1993; Nordic women’s literary history) as “efterklangslitteratur” [epigonal literature] where “romantikken klæber til både motiver og form” [romanticism clings to both motives and form] (Hjordt-Vetlesen and Mortensen 1993, 384).

So, our analysis points to a correlation between what is regarded as ‘modern’ literature and neutral sentiment. The texts who top the list of the least neutral works are in the historiographies regarded as old-fashioned, conservative epigonal literature. What is important is that this has a significant implication regarding gender. Our analysis shows that if we are to talk about a specific female style of writing in the period, we must highlight their lack of neutral sentiment, rather than their unhappiness (‘lack’ understood in a merely quantitative sense). We could even give it a new heading: ‘The least neutral text’. This outlines an interesting intersectional dynamic between gender, neutrality, and canonization, which raises several questions. In what manner have the sticky relations between these cultural categories influenced literary historiographies? Literary scholars like Jan Rosiek have engaged critically with the notion that the modern part of modern literature is a development away from metaphysics (Rosiek 2003, 9). Our analysis raises the same question. Has our understanding of modern literature equaling non-metaphysical literature made us regard an ‘unsentimental’ or neutral style of writing as the only truly modern literary disposition? And have we in this way solely emphasized a style of writing, which can be said to be a primarily male undertaking?

Our radical upscaling of the empirical foundation has in this way both challenged and confirmed the thesis of the unhappy text. The female authors are overrepresented amongst the most negative texts but are ditto amongst the most positive. Most interestingly a gendered bias has emerged regarding neutrality. The traditional hypothesis of a specifically female unhappiness has in this way shown to be canon-oriented and normative in its emphasis on negative sentiment and its notion of an unsentimental and anti-metaphysical writing style being constitutive of the modernity of modern literature. In this motion, it has largely overlooked and rejected religiously influenced or strongly affective works as uninteresting

---

9 What we are suggesting here is a perspective, which has gained attention in academic contexts outside of Denmark, particularly among feminist scholars. For instance, Christine Hamm has argued that Amalie Skram’s marriage novels can be read as a critique of the patriarchy, precisely due to what Hamm highlights as the emotionally charged, melodramatic features of the novels (Hamm 2006, 8–9). Also, Henrik Ibsen’s plays were infused with melodramatic sequences as among others Toril Moi has shown (Moi 2006). And finally, Margaret Cohen has argued that the nineteenth century realist French novel, often celebrated as the emergence of the modern novel, actually occurred through a ‘hostile takeover’ of the contemporary genre of sentimental novels, primarily written by women (Cohen 1999, 6).
epigonal literature, and in consequence, marginalized large parts of the female literary production. The unknown and non-canonical works have thus shown to pose a significant challenge to the hypothesis of the unhappy female text.

Conclusions

To summarize, the key findings in our rereading of novels from the Modern Breakthrough, based on a vast body of text and our particular metrics, are that: 1) The thesis of the unhappy text appears partially true, but, importantly, women are even more overrepresented among the positive texts. 2) However, the affect category of neutrality is more significant. Neutrality turns out to be a male, canonical enterprise, while low neutrality is primarily associated with forgotten or neglected female authors. The most crucial gender bias in the affective economy of the texts is the lack of neutrality in literature by women. 3) Finally, this and other biases point to clear intersectional dynamics between the author’s gender, the affective qualities and quantities of the texts, and their social status.

These findings are new and raise critical questions regarding a literary historiography, in which an understanding of ‘modern’ and ‘proper’ literature seems to be totally dependent on a type of writing, which first and foremost has been a male undertaking. Also, the findings are of a new kind compared to the close(r) and (more) psychologically charged readings we have been in discussion with. As an immanent result of the digital methodology, our analyses seem to capture emotions on the body rather than the mind of the texts. It can be argued that we have identified particular styles of writings in the texts rather than psychological states, but we chose to say—as programmers sometimes do—that this is not a bug but a feature in our work, because such a ‘superficial’ reading is in line with our affect-theoretical starting point, but also, of course, because there are limits to how much psychological depth and sensitivity you can expect from a language model and artificial intelligence. These boundaries and the division of labor between humans and machines will be negotiated further in the coming decades. In these negotiations, literary data, especially older texts from low-resource languages such as Scandinavian, can play an important role, both in improving the cultural and historical depth of language models and in testing their analytical (non)sense.10

Bibliography


10 For full details of the methods and codes used, including access to our language model sentiment classifier, see the online code appendix to this article [link; to be established].


