

ARTICLE

Beyond Plot: How Sentiment Analysis Reshapes Our Understanding of Narrative Structure

Katherine Elkins¹¹ Integrated Program in Humane Studies, Kenyon College

Keywords: computational narratology, digital humanities methods, large language models, diachronic sentiment analysis, quantitative literary analysis, AI-assisted literary criticism, narrative structure, cross-cultural narrative analysis

<https://doi.org/10.22148/001c.143671>

Journal of Cultural Analytics

Vol. 10, Issue 3, 2025

Sentiment analysis, particularly with the advent of large language models, is reshaping our understanding of narrative. Emotional arcs in both fictional and non-fictional narratives surface latent structures that challenge traditional notions of plot and character. One reason is that feature extraction correlates with passages often selected for close reading, suggesting the role of emotion in passage selection has been undertheorized. Case studies, moreover, show that common critiques of sentiment analysis are misplaced, and the method can identify and contrast cultural differences, locate translation effects, and illuminate collective emotional experiences. Both the strengths and limitations of sentiment analysis in understanding affect and emotion in literature are detailed to help address misconceptions. The method, as it turns out, now offers a powerful tool for highlighting emotional structure in narrative, complementing traditional literary analysis and reshaping our understanding of what constitutes a story. Sentiment analysis for literary studies asks us to reconsider not just how we analyze individual texts, but how we conceptualize the very nature of narrative itself.

1. Introduction

Sentiment analysis (SA) is finally having its heyday. "After years of disinterest and neglect," writes Simone Rebora in a 2023 survey (862), sentiment analysis "has recently become one of the most discussed topics in computational literary studies." For the uninitiated, SA is a method for mining opinion or sentiment in text (Liu 23), and the approach has long been used in industry. Literary critics have adapted the method to our field, applying it diachronically over narrative to surface emotional arcs.

While the computational approach is relatively new, the concept has deeper roots in literary theory and criticism. The idea of being able to visualize the organizational structure of a story on a graph with an x and y axis dates back to Kurt Vonnegut, who proposed it for a Ph.D. thesis and was summarily rejected ("The Sexual Revolution" 285–6). He reprises the idea in a lecture which can be seen on YouTube (<https://www.youtube.com/watch?v=oP3c1h8v2ZQ>). Drawing on Cinderella as an example, he tracks the story over time along the x-axis with fluctuations of Cinderella's fortune undulating up and down on the y-axis.

Vonnegut describes the y-axis as “fortune,” referring to the plot of the story, but scholars of sentiment analysis are more careful to describe it in terms that are in keeping with the limitations of the computational tool. In my 2022 *The Shapes of Stories*, I prefer the term “emotional arc” over “plot arc” (21). Technically speaking, the method only surfaces plot to the extent that it surfaces a language of sentiment that surrounds events.

Diachronic sentiment analysis was first popularized by Matthew Jockers, and he named his library after the Russian formalists’ concept of *syuzhet*—the way a story’s raw material (*fabula*) is organized (<https://www.matthewjockers.net/2015/02/02/syuzhet/>). A brief flurry of interest following Matthew Jockers’ blog posts and introduction of the tool was then followed by a cooling in interest. One issue I first identified is the wide variety of methods for surfacing sentiment, with no clear frontrunner for any particular narrative given the wide diversity of narratives across time and culture. While methods vary greatly, the basic pipeline is the same. Text is segmented into chunks, a sentiment value is assigned to each chunk, values are then smoothed, and feature extraction allows for human-in-the-loop evaluation.

Chunk size can vary depending on the task, and smoothing can be done. Methods for assigning a sentiment score are also numerous and quite varied in how they work, although they can be classified broadly into several categories. The simplest methods are lexical—they use a dictionary to assign each word or sentence a score. Most words are neutral, and values can scale along a gradation of intensities (e.g. “ecstatic” would score as higher in intensity than “happy”). Some methods add heuristics to lexical approaches, offering ways to take into account negation and intensification. Machine learning methods are more sophisticated, and they rely on training a classifier on an annotated dataset using one of a number of different approaches like an SVM (Support Vector Machine) or Naive Bayes classifier. Transformer models allow for contextual assessment of linguistic usage and range from smaller models (BERT) to the most recent large language models (LLMs).

Jockers was the first to suggest that the peaks and troughs identified by sentiment analysis often correspond to passages literary critics select for close reading. This insight has opened new avenues for combining computational and traditional literary analysis methods. Recent scholarship has sought to establish a robust methodology confirming his findings. Nonetheless, with so many choices, it can be difficult to determine the “best” approach for any given literary text.

In commercial settings, the process of training a model on a particular dataset (e.g. yelp or movie reviews) is clear cut. By contrast, there is no “ideal” literary text to use as training data, nor is there a particular lexicon that is likely to work in all circumstances. I offer several recommendations for how to match a model to a text, notably through the use of an ensemble model. When a

wide variety of models agree, one can have greater confidence in the arc. In areas where the models disagree, feature extraction and analysis allows for a careful approach to selecting the most accurate model. I also offer suggestions for smoothing, which favors non-parametric methods that don't presuppose a particular shape, for example a simple moving average (SMA) or LOESS (Locally Estimated Scatterplot Smoothing), a non-parametric method that creates smooth curves by fitting weighted regressions to local subsets of data. These create smooth curves through the data without assuming any global functional form.¹ Finally, I built on Jockers' earlier insights by offering a method of what I call "middle reading"—extracting the peaks and valleys and analyzing them for how well they comport with major shifts in sentiment.

While these methodological advances are promising, skepticism still remains for some critics like Wendy Hui Kyong Chun. Sentiment analysis, she argues, reduces everything to polarities like "love" and "hate" (Chun and Elkins 427). While her analysis of the method is accurate in certain cases (more on that in a moment), her love-it-or-hate-it scenario is an excellent indicator of its reception, at least until relatively recently. It's easy to understand how sentiment analysis could be so polarizing. Andrew Franta and Sean Silva,² writing in the same issue of *Critical Inquiry*, are careful to acknowledge the negative reactions humanists might have. How could a method of adding up and averaging sentiment scores over time possibly reveal anything useful for a literary critic, whose work attends to subtlety and nuance?

Rebora laments in his 2023 survey that to date there has been inadequate attention to operationalizing the robust tradition of thinking about narrative and emotion by scholars like Keith Oatley (39–69) and Patrick Colm Hogan (*The Mind and Its Stories* 20–25). What's unique to sentiment analysis is that the method has the possibility of operationalizing these existing theories in a way that many distant readings have not. While "a 'theory' for distant reading is still lacking," Rebora summarizes, "it is precisely through theoretical reasoning that SA (and other methods along with it) can actually meet the needs of literary scholars." Since then Hogan has begun making connections between his work and sentiment analysis ("Emotional Impact"),

1 While I still recommend non-parametric smoothing methods for longer-form narrative, recent experiments with Savitzky-Golay filtering have shown promising results for shorter texts. Savitzky-Golay, though parametric (using polynomial fitting), preserves peak characteristics better than LOESS because it's designed to maintain signal features while smoothing noise. This peak preservation is crucial for sentiment analysis where emotional cruxes are key interpretative moments. However, its effectiveness may vary with text length. The fixed-window polynomial fitting that works well for short texts might struggle with longer narratives where emotional patterns occur at multiple scales. Longer texts may contain both local emotional fluctuations and broader structural arcs that a single polynomial window size cannot capture optimally. Further testing is needed to determine optimal window sizes for different narrative lengths.

2 Franta and Silva are clearly writing for an audience unfamiliar with the approach and spend a great deal of time anticipating objections, even going so far as to refer to a humanist "reflex of aversion"(417). They are absolutely right to highlight the degree to which traditional approaches to the sentimental novel rely all too heavily on a few words rather than the more holistic understanding afforded by sentiment analysis.

but more work is needed. The following is an attempt to build on Hogan's recent work and further explicate points of agreement and disagreement between theorists and applied practitioners of narrative and emotion.

2. Structure Beyond Plot and Character

Building upon the discussion of sentiment analysis and its reception in literary studies, I now turn to a more fundamental question: how does this computational approach challenge our traditional understanding of narrative structure?

The history of theorizing structure in narrative has a long and venerable tradition dating back to Aristotle's theorization of tragedy in the *Poetics* (13-15). After Aristotle introduced the tripartite structure, critics continued to build on this work by identifying plot with a series of events or incidents. Sometimes this structure was theorized as a shape (like Freytag's pyramid); other times, plot was described as a set of events that always occur in a particular order, as in Propp's study of Russian folktales (42-45). While the study of character is a different aspect of narrative, it's also typically the case that plot or narrative structure has often been described in terms of a single individual whether it's the hero or anti-hero of tragedy, the main character of Propp's folktale or Vonnegut's "Cinderella," or the "hero's journey" of Joseph Campbell (23-25). These approaches, while varied, share a common assumption: that plot and character are primary structural elements of narrative.

Sentiment analysis, however, paints a more complicated picture of narrative. Our first published foray in the field looked at a counterintuitive case study: Virginia Woolf's *To the Lighthouse*, which is generally considered "plotless" given the lack of events. In fact, they found, it had a very strong and clear structure independent of plot (Elkins and Chun 8). This finding is not to suggest that plot might not also give structure to narrative. But, at the very least, we can conclude that at least one aspect of narrative structure is indirectly tied to plot. In a series of fortune reversals, a narrative might exhibit an emotional arc that is closely tied to the arc of the plot. But in the case of Woolf's novels—or, for that matter, James Joyce's novels, which Jockers investigates—the narrative can have relatively little plot structure while still evincing a very strong arc. This suggests that emotional structure may be more fundamental than plot structure and can exist independently of the latter. And while most stories exhibit a correlation between plot structure and emotional structure, a powerful narrative can retain its power through an emotional arc even when plot is minimal. The implications of this discovery are far-reaching, forcing us to reconsider what we mean by "structure" in narrative. It's not merely a matter of events or their ordering, but of emotional resonance and flow.

The second takeaway from those early investigations with sentiment analysis, however, pertains to the role of character. The Woolf case study exhibits what I call a “distributed heroine” (Elkins, *Shapes* 27). In other words, the emotional arc, with dramatic rises and falls, has peaks and valleys that highlight different characters’ perspectives. There is not just one “hero’s” journey that unites the events in a single plot arc. Instead, the emotional arc unites the separate journeys of an extended cast of characters. This finding suggests that emotional arc is more foundational than character arc—it exists independently of a single character.³ In Woolf, we rarely follow the entire arc of any one character, instead moving in and out of different perspectives. Rather than character centering the emotional arc, Woolf’s emotional arc brings these many perspectives into a single unifying narrative. While simpler stories typically follow a single protagonist with a single emotional arc, Woolf demonstrates how emotional structure can transcend individual character trajectories. This reverses traditional assumptions about narrative foundations. We typically think of character and plot as the bedrock upon which emotional complexity is built. But sentiment analysis reveals the opposite: emotional trajectory may be the foundation that unites disparate character perspectives and plot events. The relational nature of emotional arc—how shifts in sentiment create peaks and valleys that define key narrative moments—suggests narratives function as emotional architectures and not just as vehicles for character development or plot progression.

Sentiment analysis operationalizes some aspects of what we’ve previously theorized, most notably the importance of emotion in narrative from Aristotle that has been developed thereafter, from Tolstoy to more recent writers like Oatley and Hogan. It’s also the case that emotion and structure may be more fundamental than these other aspects. That is, it remains even in cases in which plot and more traditional character depiction are absent, as is the case in Woolf. This persistence of emotional structure in the absence of traditional narrative elements is particularly striking, suggesting that what we perceive as the “bones” of a story might in fact be its emotional flesh.

If we turn the lens on Woolf’s novel and modernist novels more generally, which are often seen as breaking away from traditional plot structures, we can conclude that as is the case for Woolf, while these novels may be relatively plotless, they can still exhibit a very traditional emotional arc, which complicates our understanding of modernism. This observation invites us to reconsider our understanding of literary movements and periods. Perhaps what we perceive as radical breaks in form are, in fact, continuities on a

³ While there have been attempts to map individual character arc (Vishnubhotla, et al. 2024), significant methodological challenges remain. Character utterances are too sparse and intermittent to meet time-series analysis requirements—they provide isolated data points separated by lengthy gaps rather than the continuous signal needed for sentiment analysis. Additionally, character dialogue often fails to explicitly address emotional states. From an information theory perspective, subdividing already noisy literary data into character-specific segments risks producing datasets too small to extract reliable signals. These limitations make character-level emotional arcs difficult to validate using human-in-the-loop methods and nearly impossible to compare meaningfully with narrative-level emotional arcs.

deeper, emotional level. In light of these findings, we might ask: how does this new understanding of narrative structure through the lens of sentiment analysis change our approach to literary analysis? What new questions does it open up, and what traditional assumptions does it challenge?

3. Narrative versus Story: Expanding the Scope of Sentiment

Analysis

Having established that sentiment analysis challenges traditional notions of plot and character, we turn next to an even more fundamental question: what constitutes a narrative or story? The application of sentiment analysis to an increasingly diverse range of texts has blurred the lines between traditional narrative categories and even raised questions about what distinguishes narrative from non-narrative forms.

After early experiments with diachronic sentiment analysis⁴ on well-known literary narratives, the method was quickly adapted to a much wider range of media like movie scripts, which sometimes look very similar to the novels from which they're adapted (Shaheen), but not always. Nonfiction narratives reveal similar emotional patterns; Song's analysis of end-of-life memoirs, for example, identified emotional arcs that correspond to stages of grief (Song).

But diachronic sentiment analysis can even surface well-known shapes in media we don't typically think of as "stories," for example in political speeches, which often evidence the highs and lows we associate with good storytelling (Harris et al.). And when transcripts from *Shark Tank* were mapped using diachronic sentiment analysis (Gow), winning pitches surfaced shapes that move up and down in a "W" shape—an arc that occurs so often in bestsellers that Archer and Jockers refer to the shape as "bestseller curves" (Archer and Jockers 45). Put another way, these instances could be considered "Stories that Win," to borrow the formulation of Sinclair and Acree, who study stories from a transdisciplinary perspective.

Diachronic sentiment analysis has also been used to surface shapes in tweets over time, extending our understanding of "story" even further. In these case studies, a single tweet might be classified as positive or negative, with the larger shape appearing over a succession of thousands of tweets by many

⁴ The term "sentiment analysis" in literary studies generally refers to what is more precisely called diachronic sentiment analysis—analyzing how sentiment evolves across a text's temporal dimension. This differs from synchronic sentiment analysis used in commercial applications, which classifies discrete units (reviews, tweets) as positive or negative without temporal consideration. Much of the literature uses "sentiment analysis" for both types.

For clarity, this paper uses "diachronic" when the temporal dimension needs emphasis. Applying diachronic sentiment analysis to non-narrative texts demonstrates that temporality and emotional fluctuation are not exclusive to narrative. Rather, these supposedly non-narrative forms contain latent temporal structures that become visible when we track sentiment over time.

different users. Like the distributed heroine of Woolf's novel, the arc appears as an emotional "story" of many different perspectives which, over time, give rise to the same rise and fall we might see in a fictional narrative.

One striking example is the real-life "story" of the financial collapse in Sri Lanka (De Silva). Here, the collective emotional arc of a nation undergoing economic turmoil emerges from the aggregation of individual social media posts, each a microscopic fragment of the larger narrative. Similarly, collective emotional arcs can be found in the rise of anti-Asian sentiment during COVID on Twitter (Holben et al.). This proliferation of emotional arcs across diverse textual forms raises a crucial question: are these all stories? And are they the same thing as narrative? To some extent, it depends on whom you ask. Narrative scholar H. Porter Abbott offers a broad definition, suggesting that "simply put, narrative is the representation of an event or a series of events" (Abbott 13). The most common lay definition, by contrast, found in many dictionaries, suggests the act of "telling" or writing a story is essential: a story is only a narrative if there is a teller. In "Towards a Data-Driven Theory of Narrative," Piper and Bagga attempt to operationalize the minimal features needed to detect narrative by using a classifier. Notable features include temporal distance, eventfulness, and the concreteness of worldbuilding (879-901). These confirm, to some degree, these theories, since temporal distance marks the gap between teller and tale, and eventfulness corresponds to Abbott's minimalist definition.

This definition of narrative differs from what we see in the emotional arcs of the "stories"—to the extent that we can call them such—emerging in the diverse forms just discussed. Event and sequence, as well as temporal distance, are not prerequisites. These stories come closer to operationalizing the claim that Abbott makes that narrative is "present in all discourse." And yet, that formulation is so general that it threatens to become virtually meaningless, and it certainly does not account for the "emotional bones" of the stories that concern us.

Explanations of why narrative is present in all discourse might be helpful, and a few different theories emerge. Fredric Jameson describes narrative as the "central function of the human mind" (Jameson 13), while Jean-François Lyotard calls narration the "quintessential form of customary knowledge" (Lyotard 19). Both definitions are human-centered, whether we ascribe narrative to human nature (a function of our mind) or to human culture (a form of customary knowledge). Those tweeting about the financial crisis in Sri Lanka may see themselves as putting the event into a form of "customary knowledge"—a tweet—but it's not clear that tweeting reveals anything about the central function of their mind, nor are they likely aware that their tweets, when gathered together with a multitude of other tweets, actually demonstrate an emotional arc similar to those we see everywhere. One solution is to distinguish simple units of storytelling as distinct, a move

first made by Hogan, who argues not only for the centrality of emotion to narrative but for the chunking of narrative into smaller beats or stories (Hogan). One could certainly break the tweets in the abovementioned case down into a series of beats, each one driven by emotion.

As I show in the case of both Dickens's *Great Expectations* and Toni Morrison's *Beloved*, a larger and more complex emotional arc can be made of much smaller waves, simpler "hills" and "valleys" that reveal smaller stories within the larger story (Elkins, *Shapes* 77, 84). These story units do not always seem to offer the legibility of narrative as Piper and Bagga define it. Neither "eventfulness" (as in Woolf's novel) nor temporal distance (as is the case for *Shark Tank* transcripts or tweets on X) is necessary to surface emotional arc. Storytelling, in other words, can easily arise in the present, in real time, without the kind of distance or eventfulness that traditional definitions of narrative would apply.

It's important to separate this kind of presentism—stories that unfold in the present rather than a distant past—from the charge of presentism often leveled at the method of sentiment analysis. On the surface, presentism might seem to be a legitimate critique. Since sentiment analysis—at least in its simplest forms—relies on lexicons or trained classifiers that may not correspond to language from earlier time periods, one would imagine they might perform poorly with any but the most recent texts. This argument makes sense intuitively, since some of the lexicons for sentiment were developed using reviews like those on Yelp. And yet, while this may be true for certain lexicons, it holds less true when one uses an ensemble method, and even less so as we develop large language models that have a far more robust sense of language across different time periods and in differing contexts.

This is not to say that one shouldn't treat all sentiment analysis with a certain suspicion that needs to be verified by human evaluation of the results, and we wouldn't expect great results for a model trained on Yelp reviews when analyzing Alexander Pope's rhyming translation of the *Odyssey* published in 1725. But contrary to what we expected, we found that when looking at *Odyssey* translations over several hundred years, the emotional arcs of translations didn't cluster by time period (Elkins, *Shapes* 94–98). Some translations from several hundred years ago looked quite similar to very modern translations, while translations completed within a very narrow time frame sometimes looked the most different. This held true even for translations—like Alexander Pope's—that employed a linguistic register quite different from the modern day. To a certain extent, this may be thanks to smoothing. Smoothing means that there is a certain tolerance for error, and that some misclassifications won't necessarily affect overall arcs unless they are so numerous as to swamp the signal.

While this is only one data point, this case study suggests that sentiment analysis, while it might err for specific localized words, may actually comport fairly well across time periods when smoothed with signal processing. The differences we might expect to find across time and linguistic register don't necessarily hold. This result not only challenges our assumptions about the limitations of sentiment analysis but also invites us to reconsider our understanding of how emotion functions in narratives across different historical periods and linguistic styles as well as across different media, from transcripts to tweets.

As we continue to expand the application of sentiment analysis to diverse forms of text, from traditional literature to social media posts, we are not only refining our methodologies but also deepening our understanding of what constitutes a story. The emotional arcs revealed by these analyses suggest that the essence of storytelling may be more universal and more deeply rooted in human experience than we previously imagined, transcending traditional boundaries of genre, medium, and even time.

4. What Sentiment Analysis Can (and Can't) Teach Us About Affect and Emotion

Practitioners of sentiment analysis, no matter their findings, must still answer the skeptics who ask whether a machine can truly assess emotion. To address this question, we turn next to the complex landscape of emotion research and its intersection with computational approaches.

The debate over emotion analysis can be framed around three key questions:

1. Are emotions universal or culturally constructed?
2. How do sentiment and emotion differ in their computational analysis?
3. What can sentiment analysis reveal about emotion that traditional approaches might miss?

We begin with the question of universality. Kate Crawford takes a strong stand against emotion analysis, suggesting that the research does not support the method (Crawford). This skepticism is echoed by those who point to the culturally constructed nature of emotion, often citing Lisa Feldman Barrett's influential work, *How Emotions Are Made*. By contrast, Paul Ekman is often thought of as a proponent of the opposite camp, arguing that emotion is primarily biological. The main controversy revolves around facial expression. Ekman argues that facial emotions are largely biological, and other researchers have published several studies quantifying the high degree of universality—approximately 70%—shared across cultures (Cordaro et al. 2019, 1293). As many as sixteen facial expressions are shared across cultures, writes Cowen's group (A. S. Cowen et al. 2021, 251).

Both Crawford and Barrett (“Emotional Expressions Reconsidered”), on the other hand, suggest that emotions are too culturally-dependent to allow for reliable facial emotion recognition. The scientific studies upon which the emotion facial recognition is built are inherently flawed. Interestingly, research into AI emotion facial recognition suggests that many facial recognition models are imperfect at best. Jill Noorily, for example, demonstrates that while there were many errors in specific emotion “labels,” the general valence of sentiment (positive or negative) was fairly accurate (Noorily). This finding suggests that while categorizing emotions into discrete labels may indeed be challenging, sentiment—the more general positive or negative valence—seems more consistent. These classification errors could be explained because the science is flawed, and emotions cannot be accurately classified. But they could also be explained by poor model performance, as is the point of view of researchers like Cowen, who are at work developing more performant models.

Whether or not emotion analysis will eventually perform better remains to be seen. In the meantime, it’s worth pointing out that if one believes the science to be flawed, then the critical need to preserve data privacy may be less pressing. If one suspects that emotion analysis may soon work all too well, then we have an urgent need to protect individuals from mass widespread emotional surveillance. Cowen falls into the latter camp and is part of a group developing ethical safeguards around the use of the technology.

Returning to narrative use cases, two points are worth making when it comes to emotional arcs in literary studies. First, and in spite of the headlines, the different scientific camps are not as far apart as they at first seem. Ekman carefully distinguishes between biological and cultural components of emotion. Barrett does the same in Chapter 7 of her book, “The Origin of Feeling.” There, she distinguishes between basic physiological responses to events and our interpretation of those events, arguing that while the former may be universal, the latter is subject to cultural construction.

Second, even if 70% of emotions are shared across many cultures, that still leaves a significant percentage that are not. Given this not an insignificant portion of non-shared emotional experiences, it would be easy to assemble a wide range of examples to make the argument for emotion’s cultural constructedness. Undoubtedly, more research is needed to verify and reproduce the claims being made on either side of the debate and to be more precise about shared versus distinctive emotional experiences.

This brings us to our second question: how do sentiment and emotion differ in their computational analysis? Sentiment analysis remains more in the pre-interpretive realm, at the level of arousal and intensity. When Hui Kyong Chun suggests that sentiment analysis reduces everything to “hate” and “love,” she is actually more accurately describing emotion analysis, which typically uses a limited number of categories, often based on the commonly-

referenced Plutchik's emotion wheel. Sentiment analysis, by contrast, preserves a wider range of gradations, although it is certainly the case that in certain industry use cases (i.e. very short tweets or reviews), a simple binary classification is sometimes used.

Although only loosely analogous, Patrick Colm Hogan distinguishes between two major approaches to literature and emotion: one in the tradition of Martha Nussbaum, which is focused on information processing, and another more aligned with his own work that focuses on embodiment. Information processing implies an interpretative aspect to emotion, whereas embodiment pertains to what he calls interpersonal stance: "one's sense of a target's emotional state and one's own emotional state insofar as it derives from the former" (Hogan, "Emotional Impact" 465). Sentiment analysis is more in keeping with embodiment than information processing, since interpretation often comes later. Some language of sentiment may directly interpret events according to a specific emotion (i.e., "love" or "hate"), but many words track more subtle reactions to the world that would not rise to the level of a clear emotion category. Here, the analogy to Hogan's embodiment no longer holds, since sentiment can often be disembodied or diffuse.

To illustrate, consider how sentiment analysis assigns values to words. Adjectives like "exceptional" or "terrible" offer qualitative judgments without prescribing a specific emotion. Verbs can also have sentiment attached to them, falling across a spectrum. For instance, "crash" (-0.7) implies sudden or violent collision or failure, while "bash" (-0.5) is moderately negative. "Strike" (-0.3) indicates some ambiguity: it can be neutral in some contexts (like labor actions) but often implies conflict. Positive verbs like "fortify," "cultivate," and "amplify" (+0.6), and "radiate," "elevate," "enhance" and "bloom" (+0.7) carry varying degrees of positive sentiment.

Nouns, too, have valence and polarity. "Death" scores -0.8 due to its association with loss, while "blood" scores -0.6, being less intensely negative.

This granularity allows sentiment analysis to capture nuances that might be lost in broader emotional categorizations. Unlike Hogan's interpersonal examples, sentiment analysis is not confined to words signaling particular emotions or characters' experiences. It might surface a peak during a sublime description of nature (Elkins, *Shapes* 156) or during the depiction of an event with no point-of-view or focalization. Rehora (865) points to a lack of distinction between aesthetic and embodied experience as a weakness of sentiment analysis, but it can also be seen as a strength. Sentiment analysis extends to a wide variety of descriptions, perceptions, places, and objects. It can attach to an event or an atmosphere just as easily as a person. By tracking this ebb and flow of sentiment throughout a text, we can reveal patterns of emotional intensity that may not be immediately apparent through a

specialized focus on character, plot or even scene. These arcs can surface in unexpected places, at times even highlighting subtle sentiment cues that could easily go unnoticed.

But this granularity raises our third question: what can sentiment analysis reveal about narrative that traditional approaches might miss? Patrick Colm Hogan's work provides a bridge between traditional narrative theory and computational approaches. In *The Mind and Its Stories*, Hogan argues that narrative patterns are universal across cultures because they stem from common human emotions and shared cognitive processes (2). His more recent work connects our work to his theory that "emotion dynamics shapes stories" (Hogan 2011, 231).

Although his focus is on emotion over sentiment, it is likely the case that shared responses to the world are at play in emotional arc and the power of storytelling. As Hogan notes, for many years, "critics preferred to pretend that feelings did not have much of a role to play in literary experience, at least not strong feelings, at least not in works of merit" (Hogan, "Emotional Impact"). Sentiment analysis confirms this likelihood, suggesting that emotional patterns may be fundamental to how narratives are structured and how they affect readers. However, it's crucial to remember that sentiment analysis remains agnostic about specific emotional experiences or reader responses. It doesn't tell us how a reader will feel, but rather maps the emotional potentials encoded in the text. This is both a limitation and a strength: while it can't predict individual responses, it can reveal patterns that might influence those responses in subtle ways.

As we continue to explore the possibilities of sentiment analysis, we should nonetheless remain aware of its limitations. The term "emotional arc" itself is somewhat misleading, as what we're measuring is sentiment, not emotion. Yet, "sentiment arc" lacks the intuitive appeal needed for broader theoretical discussions. This terminological challenge reflects the broader difficulties in bridging computational methods with traditional literary theory. Looking forward, several questions remain. How do emotional arcs correlate with reader response? How do cultural and historical contexts influence these arcs? And how might understanding these arcs change our approach to literary analysis and theory?

5. What Feature Extraction Can Teach us About Close Reading

We turn next to a critical question: How does this computational approach intersect with one of the most fundamental practices in literary studies—close reading? This intersection not only illuminates the strengths and limitations of both methods but also challenges us to reconsider some of our most deeply held assumptions about literary analysis.

Early forays into emotional arc analysis often focused on the shape itself as significant. Following a tradition that emphasizes a shared structure across stories, approaches like Andrew Reagan’s attempted to operationalize theories that proposed a set number of story shapes, such as the “Rags to Riches” or “Cinderella” tale (Reagan et al. 4). With enough smoothing, it’s true that virtually any story can be made to fit a small number of shapes. This practice of categorizing a large corpus of texts shares much in common with Franco Moretti’s “distant reading” approach (Moretti 48–49). But this approach ignores a crucial step in the typical machine learning pipeline: feature extraction and analysis. In sentiment analysis, one significant feature is the crux points or peaks and valleys. These signal moments when the sentiment changes course, moving from rising to falling or vice versa. These points serve two purposes. First, they can confirm that our model is working well by identifying key moments in the narrative when the language of sentiment changes. Second, they offer a new lens through which to examine the text.

As Jockers first remarked and I verified in numerous case studies, these crux points tend to correspond to passages which critics often turn to for close reading (Jockers; Elkins, *Shapes* 37). This correlation raises important questions about the nature of close reading itself and how we select passages for analysis. Close reading, a practice rooted in biblical criticism and popularized by New Critics like I.A. Richards in the 1920s, has a long history (Richards). For a short text like a poem, close reading involves attending to the linguistic choices of every (or almost every) word. But for long-form narrative, the process becomes more complex. How do we choose which passages to read closely in a novel spanning hundreds of pages?

Surprisingly, there’s less theorization of this selection process than one might expect. As I point out in my work on Proust that passage selection can radically alter the interpretation of key themes (Elkins, “Proust’s Consciousness” 217–220). This observation underscores a potential blind spot in our methodology, one that sentiment analysis can help illuminate. Some might argue that the selection process is a craft: it’s developed through apprenticeship and resistant to formalization. They might contend that it’s crucial to preserve this practice against any attempts at quantification. But what if we could write a set of instructions for passage selection that could be applied to any close reading? Sentiment analysis offers just such a possibility, although one that is easier to do with a computer than by a human alone.

It is true that passages identified by sentiment analysis may be slightly longer than those typically chosen for close reading; perhaps “middle reading” is a more apt term (Elkins, *Shapes* 11). Nonetheless, this algorithmic approach to passage selection challenges our assumptions about the most intuitive aspects of literary analysis. This challenge echoes our earlier discussion of how sentiment analysis complicates traditional notions of plot and character.

Just as we found that emotional arcs can reveal structures independent of conventional plot and character elements, here we see that computational methods can identify significant passages whose key features—a change in the general direction of the language of sentiment—might escape our conscious notice.

What this phenomenon demonstrates is that passage selection is undertheorized because we haven't fully understood the features that impact our choices. The language of sentiment creates patterns with clear signals, even if those signals often lie below our conscious perception. It's likely, moreover, that we haven't been paying enough attention to emotion as central to our selection process. This insight also connects back to our exploration of what sentiment can teach us about affect and emotion. Just as we found that sentiment analysis can reveal emotional structures that transcend cultural specificity, here we see it uncovering patterns that transcend individual critical intuition.

Andrew Piper suggests that literary studies often engages in generalization from text to world without offering a clear methodology (Piper 45–46). While his focus is on generalizing from literature to life, this same methodological gap exists when we generalize from a specific passage to the whole text. How can we be sure that the passages we select truly represent the novel as a whole? Could undertheorized close reading actually support confirmation bias, our tendency to focus on details that support our preexisting theories?

Sentiment analysis offers one way to address this concern, providing a method to identify key passages that play crucial roles in the narrative's emotional structure. This is not to say that other passages aren't also important, but it does provide a systematic way to connect our close reading of specific moments to the arc of the entire narrative. Computational methods like sentiment analysis don't replace human interpretation; rather, they provide new entry points and perspectives that can enrich our critical practice, as I hope I've shown here. Still, the intersection of sentiment analysis and close reading challenges us to reconsider some of our most fundamental practices. It invites us to be more explicit about our methodologies and more open to new approaches. As we continue to explore these methods, we may find ourselves rethinking how we analyze individual texts by being more intentional about how we pick passages for close reading.

6. Understanding Stories Better Through Emotional Arc

Having explored the theoretical implications of sentiment analysis for narrative structure and close reading, we next turn to more practical considerations. How can this method deepen our understanding of stories

across? As we'll see, emotional arc analysis not only illuminates individual texts but also offers insights into cultural differences, translation effects, and even real-time social phenomena.

Let's begin with an example that demonstrates how sentiment analysis can highlight, rather than obfuscate, different emotional responses across time and culture. In my reading of Defoe's *Robinson Crusoe*, feature extraction highlighted cruxes that center on religious passages (Elkins, *Shapes* 71). Many contemporary secular Western readers might not find these highly-charged religious moments particularly salient, potentially overlooking them when selecting key passages for analysis. Yet by surfacing these cruxes, sentiment analysis alerts us to how readers might have experienced the narrative differently when it was first published, as well as giving us clues about how Defoe intentionally shaped his narrative as a spiritual arc. It is also helpful in illuminating how contemporary readers from other religious traditions may find shared emotional readerly response (Alkodimi), or how scholars argue for the centrality of religion to the colonialist depiction (McInelly). Sentiment analysis, in other words, helps surface aspects of a differential reading experience across historical periods and cultural contexts.

This differential reading becomes particularly evident in translation studies, an area where sentiment analysis has had a special impact. Consider the case of English translations of the *Odyssey* touched upon earlier. While translations into the same language can be relatively similar, minor differences surfaced by sentiment analysis can point to cultural variations in interpreting key events. One crucial distinction, for instance, is how different translations portray the story's ending after Odysseus' return. Does the narrative conclude on a happy note, or is the reunion between Odysseus and Penelope forever haunted by past events? Different translations suggest slightly different outcomes through careful selection of sentiment-laden language (Elkins, *Shapes* 94–98).

Multilingual methods for sentiment analysis further enhance our ability to compare translations with the original across different languages and cultures. Inspired by Patrick O'Neill's *Transforming Kafka: Translation Effects*, Strain et al. compared Kafka's *The Trial* in its original German against translations into French, Spanish, and English. Their findings supported O'Neill's suggestion that translation choices have significantly altered Kafka's narratives, revealing vastly different emotional arcs across translations (Strain). While this idea isn't entirely new, examining the arcs helps us visualize and quantify just how dramatically the same narrative can change in translation.

Moving beyond traditional literary texts, sentiment analysis proves equally illuminating when applied to contemporary digital discourse. Consider the case study of the Sri Lankan financial crisis, where De Silva et al. applied

diachronic sentiment analysis to tweets.⁵ Just as emotional cruxes often correlate to plot events in fiction, crux extraction aligned with key historical events. The analysis surfaced moments of deepening crisis, including the storming of the president’s house, a subsequent “pool party,” and the president’s furtive departure. Notably, following these events, the sentiment arc flattened into a relatively negative line, suggesting that although the popular uprising resulted in a momentary celebration, the end result was uniformly negative for everyday Sri Lankans (De Silva).

Another intriguing case study comes from Gimbel et al., who analyzed tweets before and after five highly-contested U.S. midterm elections were called. Even with smoothing, they observed extreme oscillations following several of the elections, making simple feature detection of cruxes challenging. This “ringing effect,” akin to snapping a taut rubber band, suggests an emotional vibration within a community, with voices both very positive and very negative reacting strongly to the official announcements. Larger oscillations seemed to indicate less confidence in the election results (Gimbel).

These examples from social media analysis connect back to our earlier discussion of how sentiment analysis challenges traditional notions of narrative. They show that short, individual commentaries can, in aggregate, create emotional arcs similar to those in fictional narratives. This broader sense of “story” invites literary scholars to reconsider the connections between life and art, suggesting that while elements of fiction are undoubtedly constructed, they may also mirror organic aspects of lived experience.

This mirroring of life and art is perhaps most poignantly illustrated in Song’s study of end-of-life narratives. Examining five non-fiction memoirs, Song found that key cruxes identified through sentiment analysis corresponded to emotional turning points in the narratives. Human-in-the-loop analysis confirmed that these cruxes aligned with Kübler-Ross’s five stages of grief (Kübler-Ross), though not necessarily in the order originally proposed. This aligns with Kübler-Ross’s later acknowledgment that these stages aren’t always experienced linearly (Kübler-Ross and Kessler). These results not only validate the effectiveness of sentiment analysis in identifying emotionally significant moments but also suggest how personal narratives might naturally structure themselves around commonly-shared emotional patterns.

⁵ Tess McNulty rightly calls for more attention devoted to theorizing and working with recent cultural media (McNulty). This is certainly the case for tweets, for which keyword selection is crucial. As just one example, a selection of tweets concerning “reproductive rights” will turn up a very different social network and “story” than “abortion.” The kinds of ethical concerns Wendy Hui Kyong Chun raises are also apt. The practices used in the case studies mentioned in this essay do not profile individuals but instead focus on collective aggregates. Studies that emotionally profile individuals are far more questionable. Since these studies, changes to X have made it difficult to scrape tweets in ways that were once available to researchers—a real loss to our ability to document contemporary events as they unfold. While a focus on individual researchers and ethical practices is always warranted, it’s important to remember that large tech companies continue to emotionally profile individuals on a scale unavailable to the typical researcher. See, for example Joseph’s (July 30, 2024) discussion of Meta’s facial recognition profiling.

7. Sentiment Analysis in the Age of Large Language Models

In this final section, we turn to the cutting edge of sentiment analysis to ask: how might large language models (LLMs) transform the landscape of sentiment and emotion analysis? LLMs may soon address some of the challenges we've encountered while opening up new avenues for exploration.

While it's commonplace to use ensemble models in computer science, the need to compare and evaluate many models to pick the one best tailored to a particular text is a high bar for the average humanist with modest coding and data analytic skills. One can easily arrive at a combinatorial explosion that makes it difficult to evaluate many different models, each with its own set of crux points. Recent large language models (LLMs) offer an easier and more streamlined approach. We found that making a function call to GPT-4 via OpenAI's API worked fairly well (Chun and Elkins 15). This is likely because dynamic word embeddings, which capture the complexity of semantic variations in hundreds of dimensions, are able to take into account the multivariate nature of language better than simpler lexicon models.

The power of LLMs lies in their contextual understanding, a feature that addresses one of the key limitations discussed earlier in traditional sentiment analysis. Unlike simpler models that might struggle with irony, sarcasm, or complex emotional states, LLMs have a much deeper grasp of context. This could lead to more nuanced interpretations of sentiment, potentially bridging the gap between computational analysis and the kind of subtle reading literary critics engage in. Moreover, LLMs offer the possibility of far more fine-grained and performant emotion analysis. Instead of just classifying text using sentiment, these models might be able to simultaneously categorize emotions into multiple categories. This aligns with Alan Cowen's recent research documenting at least 27 distinct categories of emotion that can be mapped onto a continuous three-dimensional space (A. S. Cowen et al. E7900). Narratologist Jim Phelan has long argued that our emotional vocabulary is inadequate for the kinds of nuanced analysis literary critics perform (Phelan 73). Cowen's latest approach, mapping emotions onto a complex, three-dimensional continuous semantic space, may bring affective science and literary criticism closer together.

The cross-lingual capabilities of LLMs also promise to revolutionize comparative literature studies. In general and until recently, it was standard practice to translate texts into English before applying sentiment analysis. English-language models were so performant that this was often as good if not better than using models targeted for under-resourced languages. While that practice may be fine for commercial needs, it's definitely a problem for the kind of linguistic nuance that surfacing emotional arcs in literary texts demands. The proliferation of LLMs pre-trained or fine-tuned on other languages will change all this and offer new avenues for literary scholars, especially when it comes to working on non-Anglophone texts.

I found uneven performance in French using BERT, a smaller language model based on the Transformer architecture. Here more recent experiments using the open-source French-engineered Mistral fared much better (Elkins, “In Search of a Translator”). There are now LLMs specifically designed for high performance in a wider variety of languages. New models, for example JAIS for Arabic, have been benchmarked for a range of NLP tasks including sentiment analysis with promising results. More work is needed to confirm how well they work in edge cases like highly subtle literary texts. In my recent work comparing the original Proust with a variety of English translations, Mistral was able to surface very minor linguistic choices that enabled fine-grained comparisons.

This ability to work across languages could address some of the concerns we raised earlier about how translation choices can significantly alter emotional arcs. It might allow for more direct comparisons between original texts and their translations, potentially revealing subtle shifts in emotional tenor that occur in the translation process. The historical language understanding of LLMs also offers exciting possibilities. Trained on vast corpora of historical texts, these models might be better equipped to analyze sentiment in older works, potentially addressing any bias of presentism in sentiment analysis discussed earlier. This could revolutionize how we approach texts from different historical periods, allowing us to more accurately capture the emotional nuances as they might have been understood by contemporaneous readers.

We might also imagine LLMs that can track character emotions separately from narrative tone, or models that can analyze visual and textual sentiment simultaneously in illustrated texts or films. Multimodal sentiment analysis, which takes into account sound, image and text, is advancing quite rapidly (Chun, “AI Multimodal Sentiment Arcs”). Such developments could offer new insights into the complex interplay between different narrative elements that we touched upon in our discussion of Woolf’s “distributed heroine” structure.

It should also be possible in the very near future for those without coding skills to use LLMs to classify and visualize sentiment much more easily. This democratization of tools will be a welcome advancement, allowing more traditionally-trained literary scholars to explore sentiment analysis and answer many of the questions that still remain. Indeed, LLMs offer the possibility that quantitative analysis with a range of methods, not just sentiment analysis, might soon be within reach for a much wider group of practitioners.

As with any powerful tool, the use of LLMs in sentiment analysis raises important ethical considerations. There is certainly the risk of potential biases in training data and of over-relying on AI interpretations. As stressed already, computational methods should complement rather than replace human interpretation. Still, as these tools become more accessible, they have the

potential to transform not just how we analyze individual texts, but how we conceptualize the very nature of emotion in narrative. The challenge for literary scholars will be to engage critically with these new methods, leveraging their power while remaining grounded in our tradition of literary practice.

Submitted: August 24, 2024 EDT. Accepted: January 02, 2025 EDT. Published: September 10, 2025 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

- Abbott, H. Porter. *The Cambridge Introduction to Narrative*. 2nd ed., Cambridge University Press, 2008, <https://doi.org/10.1017/CBO9780511816932>.
- Alkodimi, Khaled Abkar. "Allegorical Depictions of the Spiritual Life of Man in *Robinson Crusoe* (1719): An Islamic Perspective." *Cogent Arts & Humanities*, vol. 10, no. 1, 2023, <https://doi.org/10.1080/23311983.2023.2259659>.
- Archer, Jodie, and Matthew Lee Jockers. *The Bestseller Code: Anatomy of the Blockbuster Novel*. St. Martin's Press, 2016.
- Aristotle. *Poetics*. Translated by S. H. Butcher, The Internet Classics Archive, <http://classics.mit.edu/Aristotle/poetics.html>. Accessed 27 July 2024.
- Barrett, L. F., et al. "Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements." *Psychological Science in the Public Interest*, vol. 20, no. 1, 2019, pp. 1–68, <https://doi.org/10.1177/1529100619832930>.
- Barrett, Lisa Feldman. *How Emotions Are Made: The Secret Life of the Brain*. Houghton Mifflin Harcourt, 2017.
- Campbell, Joseph. *The Hero with a Thousand Faces*. 3rd ed., New World Library, 2008.
- Chun, Jon. *A Novel Method to Visualize Multimodal AI Sentiment Arcs in Long-Form Narratives*. <https://github.com/jon-chun/multisentimentarcs>. Accessed 27 July 2024.
- Chun, Jon, and Katherine Elkins. "eXplainable AI with GPT4 for Story Analysis and Generation: A Novel Framework for Diachronic Sentiment Analysis." *International Journal of Digital Humanities*, vol. 5, 2023, pp. 1–26, <https://doi.org/10.1007/s42803-023-00069-8>.
- Chun, Wendy Hui Kyong, et al. "'Understanding' Asians: Anti-Asian Racism, Sentimentality, Sentiment Analysis, and Digital Surveillance." *Critical Inquiry*, vol. 50, no. 4, 2021, pp. 425–51, <https://doi.org/10.1086/728946>.
- Cordaro, Daniel T., et al. "The Recognition of 18 Facial-Bodily Expressions across Nine Cultures." *Emotion (Washington, D.C.)*, vol. 20, no. 7, 2020, pp. 1292–300, <https://doi.org/10.1037/emo0000576>.
- Cowen, A. S., et al. "Sixteen Facial Expressions Occur in Similar Contexts Worldwide." *Nature*, vol. 589, 2021, pp. 251–57, <https://doi.org/10.1038/s41586-020-3037-7>.
- Cowen, Alan S., et al. "Self-Report Captures 27 Distinct Categories of Emotion Bridged by Continuous Gradients." *Proceedings of the National Academy of Sciences*, vol. 114, no. 38, 2017, pp. E7900–09, <https://doi.org/10.1073/pnas.1702247114>.
- Crawford, Kate. "Artificial Intelligence Is Misreading Human Emotion." *The Atlantic*, Atlantic Media Company, 13 Nov. 2023, <http://www.theatlantic.com/technology/archive/2021/04/artificial-intelligence-misreading-human-emotion/618696/>.
- De Silva, Cherantha. *How Did Sri Lankan Protestors End Up in the President's Pool? Understanding the Evolution of an Occupy-Style Protest: A Story of Economic Turmoil, Declining Social Sentiment and Resulting Political Change*. 2022, https://digital.kenyon.edu/dh_iphs_prog/59.
- Ekman, Paul. "An Argument for Basic Emotions." *Cognition and Emotion*, vol. 6, no. 3–4, 1992, pp. 169–200, <https://doi.org/10.1080/02699939208411068>.
- . "Cultural Differences in Emotional Expressions." *Paul Ekman Group*, <http://www.paulekman.com/blog/cultural-differences-in-emotional-expressions/>. Accessed 23 Jan. 2024.
- Elkins, Katherine. "In Search of a Translator: Using AI to Surface What's Lost in Translation." *Frontiers of Computer Science*, Aug. 2024, <https://doi.org/10.3389/fcomp.2024.1444021>.

- . "Proust's Consciousness." *Proust's In Search of Lost Time: Philosophical Perspectives*, edited by Katherine Elkins, Oxford University Press, 2022, <https://doi.org/10.1093/oso/9780190921576.001.0001>.
- . *The Shapes of Stories*. Cambridge University Press, 2022, <https://doi.org/10.1017/9781009270403>.
- Elkins, Katherine, and Jon Chun. "Can Sentiment Analysis Reveal Structure in a Plotless Novel?" *ArXiv*, 31 Aug. 2019, <http://arxiv.org/abs/1910.01441>.
- Franta, Andrew, and Sean Silver. "Sentiment Analysis and the Sentimental Novel." *Critical Inquiry*, vol. 50, no. 3, pp. 375–584, <https://doi.org/10.1086/728948>.
- Freytag, Gustav. *Technique of the Drama: An Exposition of Dramatic Composition and Art*. Translated by Elias J. MacEwan, Scott, Foresman and Company, 1900.
- Gimbel, Ben. *Quantifying Polarization around Election Denial: Measuring Public Sentiment Changes in the 2022 Midterms*. 2022, https://digital.kenyon.edu/dh_iphs_prog/66.
- Gow, Alexander. *Blood in the Water: Storytelling and Sentiment Analysis in ABC's Shark Tank*. 2021, https://digital.kenyon.edu/dh_iphs_ss/4.
- Harris, Davida, et al. *Sentiment Analysis Amidst the COVID-19 Pandemic: Comparing the Emotional Valence of President Trump and Governor Cuomo's Coronavirus Briefings*. 2020, https://digital.kenyon.edu/dh_iphs_prog/16.
- Hogan, Patrick Colm. "The Emotional Impact of Literature." *The Routledge Companion to Literature and Cognitive Studies*, edited by Jan Alber and Ralf Schneider.
- . *The Mind and Its Stories: Narrative Universals and Human Emotion*. Cambridge University Press, 2003, <https://doi.org/10.1017/CBO9780511499951>.
- Holben, Mari, et al. "The Rise of Anti-Asian Sentiment With COVID-19: A Sentiment Analysis of Twitter Data." *IPHS 200: Programming Humanity. Paper 15*, 2020, https://digital.kenyon.edu/dh_iphs_prog/15.
- Homer. *Odyssey*. Translated by Alexander Pope, Project Gutenberg, 2021.
- Jameson, Fredric. *The Political Unconscious*. Cornell University Press, 1981.
- Jockers, Matthew L. *Revealing Sentiment and Plot Arcs with the Syuzhet Package*. 2015, <http://www.matthewjockers.net/2015/02/02/syuzhet/>.
- Joseph, Jibin. "Meta to Pay \$1.4 Billion to Settle Texas Facial-Recognition Lawsuit." *PCMag*, 30 July 2024, <http://www.pcmag.com/news/meta-to-pay-14-billion-to-settle-texas-facial-recognition-lawsuit>.
- Keen, Suzanne. "Narrative Empathy." *The Living Handbook of Narratology*, edited by Peter Hühn et al., Hamburg University Press, 2013, <https://doi.org/10.1515/9783110316469.521>.
- Kübler-Ross, Elisabeth. *On Death and Dying*. Macmillan, 1969.
- Kübler-Ross, Elisabeth, and David Kessler. *On Grief and Grieving: Finding the Meaning of Grief Through the Five Stages of Loss*. Scribner, 2005.
- Liu, Bing. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, 2015, <https://doi.org/10.1017/CBO9781139084789>.
- Liotard, Jean-François. *The Postmodern Condition: A Report on Knowledge*. Translated by Geoff Bennington and Brian Massumi, University of Minnesota Press, 1984, <https://doi.org/10.2307/1772278>.
- McInelly, Brett. "Expanding Empires, Expanding Selves: Colonialism, the Novel, and Robinson Crusoe." *Studies in The Novel*, vol. 35, 2003, pp. 1–21.
- McNulty, Tess. "Content's Forms." *New Literary History*, vol. 54, no. 1, 2022, pp. 795–851. *Project MUSE*, <https://doi.org/10.1353/nlh.2022.a898330>.

- Moretti, Franco. *Distant Reading*. Verso Books, 2013.
- Noorily, Jill. *AI Reads Playboy (but Not for the Articles): Revealing Cover Trends with Deep Neural Networks*. 2022, https://digital.kenyon.edu/dh_iphs_ss/17.
- Oatley, Keith. "Emotions and the Story Worlds of Fiction." *Narrative Impact: Social and Cognitive Foundations*, edited by Melanie C. Green et al., Lawrence Erlbaum Associates, 2002, pp. 39–69.
- O'Neill, Patrick. *Transforming Kafka: Translation Effects*. University of Toronto Press, 2014, <https://doi.org/10.3138/9781442623798>.
- Phelan, James. *Experiencing Fiction: Judgments, Progressions, and the Rhetorical Theory of Narrative*. Ohio State University Press, 2007.
- . "Rhetorical Theory, Cognitive Theory, and Morrison's 'Recitatif': From Parallel Play to Productive Collaboration." *The Oxford Handbook of Cognitive Literary Studies*, edited by Lisa Zunshine, Oxford University Press, 2015, pp. 120–35.
- Piper, Andrew. *Can We Be Wrong? The Problem of Textual Evidence in a Time of Data*. Cambridge University Press, 2020, <https://doi.org/10.1017/9781108922036>.
- Piper, Andrew, and Sunyam Bagga. "Towards a Data-Driven Theory of Narrative." *New Literary History*, vol. 53, no. 4, 2022.
- Pope, Alexander, translator. *The Odyssey of Homer*. Project Gutenberg, 2021.
- Propp, Vladimir. *Morphology of the Folktale*. Translated by Laurence Scott, 2nd ed., University of Texas Press, 1968.
- Reagan, Andrew J., et al. "The Emotional Arcs of Stories Are Dominated by Six Basic Shapes." *EPJ Data Science*, vol. 5, no. 1, 2016, pp. 1–12, <https://doi.org/10.1140/epjds/s13688-016-0093-1>.
- Rebora, Simone. "A Survey of Computational Sentiment Analysis." *Digital Scholarship in the Humanities*, vol. 35, no. 4, 2020, pp. 857–72.
- Richards, I. A. *Practical Criticism: A Study of Literary Judgment*. Routledge & Kegan Paul, 1929.
- Shaheen, Erin. *Adapted Arcs: Sentiment Analysis and The Sorcerer's Stone*. 2021, https://digital.kenyon.edu/dh_iphs_ss/1.
- Sinclair, Betsy, and Billy Acree. *About: Stories That Win*. <http://www.storiesthatwin.org/about>.
- Song, Hemmi. *On Death and Emotion: Evaluating the Five Stages of Grief in End-of-Life Memoirs Using AI Deep Learning Models*. 2022, https://digital.kenyon.edu/dh_iphs_ss/20.
- Strain, Flannery. *The Trials of Translation: A Cross-Linguistic Survey of Sentiment Analysis on Franz Kafka's Trial*. 2022, https://digital.kenyon.edu/dh_iphs_prog/55.
- Vishnubhotla, et al. "The Emotion Dynamics of Literary Novels." *ACL*, 2024, <https://doi.org/10.18653/v1/2024.findings-acl.150>.
- Vonnegut, Kurt. "Kurt Vonnegut on the Shapes of Stories." *YouTube*, 30 Oct. 2010, <http://www.youtube.com/watch?v=oP3c1h8v2ZQ>.
- . "The Sexual Revolution." *Palm Sunday*, Penguin Random House, 2009, pp. 277–90.
- "What Is Narrative Theory?" *Project Narrative*, Ohio State University, <http://projectnarrative.osu.edu/about/what-is-narrative-theory>.
- Woolf, Virginia. *To the Lighthouse*. Harcourt, Brace & Co., 1927.