

The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us

Rutgers University has made this article freely available. Please share how this access benefits you.
Your story matters. <https://rucore.libraries.rutgers.edu/rutgers-lib/43176/story/>

This work is an **ACCEPTED MANUSCRIPT (AM)**

This is the author's manuscript for a work that has been accepted for publication. Changes resulting from the publishing process, such as copyediting, final layout, and pagination, may not be reflected in this document. The publisher takes permanent responsibility for the work. Content and layout follow publisher's submission requirements.

Citation for this version and the definitive version are shown below.

Citation to Publisher Goldstone, Andrew & Underwood, Ted. (2014). The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us. *New Literary History* 45(3), 359-384. <https://dx.doi.org/10.1353/nlh.2014.0025>.

Citation to this Version: Goldstone, Andrew & Underwood, Ted. (2014). The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us. *New Literary History* 45(3), 359-384. Retrieved from [doi:10.7282/T3222RZT](https://doi.org/10.7282/T3222RZT).

Terms of Use: Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

Article begins on next page

The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us

Andrew Goldstone and Ted Underwood

Forthcoming in *New Literary History*

The history of literary study is primarily remembered as a narrative of conflicting ideas. Critical movements clash, then establish a succession: New Criticism, structuralism, deconstruction, New Historicism. Although scholars have complicated this simple history-of-ideas story in recent decades with an emphasis on social and institutional struggle, generational conflict remains a central framework: instead of struggles among ideas there are struggles among genteel amateurs, professionalized scholars, and so on.¹ In emphasizing conflict, these approaches still leave aside important dimensions of the history of scholarship: assumptions that change quietly, without explicit debate; entrenched patterns that survive the visible conflicts; long-term transformations of the terrain caused by social change. To see more of these kinds of history, another approach is needed—a way of discovering and interpreting patterns on a different historical scale.

Consider the following odd but intriguing example of a century-long change in the discipline that never became an overt debate: literary scholars now mention numbers only about 60% as often as we did in the early twentieth century. Figure 1 indicates the frequency of number words over

1. Key examples include: Gerald Graff, *Professing Literature: An Institutional History*, 20th anniversary ed. (Chicago: University of Chicago Press, 2007); John Guillory, *Cultural Capital: The Problem of Literary Canon Formation* (Chicago: University of Chicago Press, 1993); and Michael Warner, "Professionalization and the Rewards of Literature: 1874–1900," *Criticism* 27 (1985): 1–28. For a recent challenge to this tradition's institutional emphasis, see Rachel Sagner Buurma and Laura Heffernan, "The Common Reader and the Archival Classroom: Disciplinary History for the Twenty-First Century," *New Literary History* 43, no. 1 (Winter 2012): 113–35.

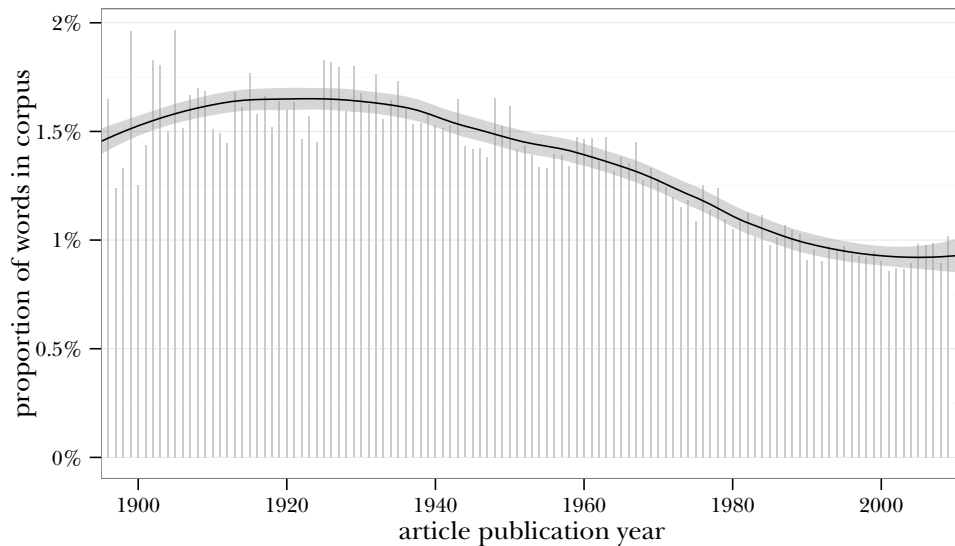


Figure 1: Yearly frequencies of number words from *two* to *hundred* and *first* to *tenth* in a corpus of articles in seven generalist literary-studies journals. The light gray bars are the data; the dark curve is a smoothed version highlighting the trend.

the last century in a collection of generalist literary-studies journals (described in detail below). Counting is a simple rhetorical gesture, and usually an uncontroversial one, but the magnitude of change makes clear that there are tacit customs governing this aspect of scholarly diction as effectively as the explicit premises of any “school.” On the other hand, the century-long trend doesn’t fit neatly into a narrative about, say, the displacement of New Criticism by historicism. What were literary scholars counting in 1930, anyway? We can make a good guess by searching for nouns whose frequency correlates closely with the frequency of numbers: they include *editions*, *dates*, *years*, *lines*, and *verses*. Though this list is suggestive, it still doesn’t explain why the practice of enumeration itself declined.²

James English has argued that contemporary literary studies defines

2. The word counts used for the chart, supplied by JSTOR, don’t include Arabic numerals; however, our own analysis of a sample of scanned articles suggests that all forms of enumeration decline together in late-twentieth-century literary scholarship.

itself through an “antagonism toward counting.”³ While this antagonism is commonly staged as a defense of humanistic tradition, figure 1 suggests that the tradition being defended may actually be relatively recent. Of course, the graph also suggests that this may be an inauspicious time for a pair of literary scholars to make an argument that depends on numbers. In spite of what looks like a tiny upturn in the present decade, we may expect skepticism if we argue that the mere counting of words can redress important blind spots in the history of literary scholarship, highlighting long-term changes—like this one—that were not consciously thematized by scholars. Nonetheless, we believe numbers are useful, not because they’re precise or objective or free from interpretation, but because they can help us grapple with the subtle interpretive problems endemic to cultural history, where a change is often determined by multiple causes. We use a technique known as *topic modeling* to elucidate historical patterns in a corpus of more than 21000 scholarly articles from the last 120 years, and argue that this approach reveals important but hitherto unarticulated trends in literary scholarship. Furthermore, well-known transitions in scholarly method look different in the light of our evidence: concepts taken for granted as the bedrock of literary studies are both more recent and more variable than contemporary debates assume. More broadly, we hope to make clear that quantitative methods like ours promise not definitive simplicity but new ways of analyzing the complexity of literary and cultural history.

TOPIC MODELING

Whether numbers add subtlety or flatten it out depends on how you use them, and a simple graph of word frequency like figure 1 is not necessarily the most nuanced approach. The graph is hard to interpret in part because these words have been wrenched out of context. *Five* might count editions or it might count the length of five long winters. The meanings of words are shifting and context-dependent. For this reason, it’s risky to construct groups of words that we imagine are equivalent to some predetermined concept. A group of numbers may be relatively uncontroversial,

3. James F. English, “Everywhere and Nowhere: The Sociology of Literature After ‘the Sociology of Literature,’” *NLH* 41, no. 2 (Spring 2010): xiii.

but a group of, say, “philological terms” would be pretty dubious. If historicism tells us anything, it’s that the meaning of a term has to emerge from the way it’s used in a specific historical context.

In recent years, researchers in computer science have devised exploratory techniques that can identify groups of words with more sensitivity to the discursive context. In this essay, we use one such technique, a *probabilistic topic modeling* algorithm, which seeks to infer meaningful groupings of words statistically, on the basis of their patterns of usage in a corpus of documents. Of course no interpretive algorithm can pretend to work from a blank slate: using the algorithm requires making careful choices, especially about the conceptual breadth of patterns to be looked for. But an assumption need not be a straightjacket. In fact, one advantage of an algorithmic technique is that it allows researchers to make their assumptions explicit as parameters of the method, which they are then free to vary.

The aim of topic modeling is to identify the thematic or rhetorical patterns that inform a collection of documents: for instance, the articles in a group of scholarly journals.⁴ These patterns we refer to as *topics*. If each article were about a single topic, we would only need to sort the articles into categories. But in reality, any article participates in multiple thematic and rhetorical patterns. Similarly, if a given word always expressed a given topic—if *nature*, say, were always part of discussions of pastoral—we could classify topics by sorting individual words into categories. But words have different meanings associated with different contexts: *nature* is also associated, for instance, with science. The algorithm responds to this challenge by modeling a topic as an intersection of vocabulary and context: it identifies groups of words that tend to be associated with each other in a particular subset of documents. In the particular topic model we discuss below,

4. *Topic modeling* is a generic term that can cover a range of specific algorithms. When we refer to topic modeling in the rest of this essay, we are referring to the specific technique we used, Latent Dirichlet Allocation (LDA). We give only a brief explanation of this algorithm here. Matthew L. Jockers gives more detail in his discussion of LDA applied to nineteenth-century novels in “Theme,” chap. 8 in *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013). The LDA algorithm was first described in David Blei, Andrew Ng, and Michael Jordan, “Latent Dirichlet Allocation,” *Journal of Machine Learning Research* 3 (2003): 993–1022. The implementation we use also reflects insights in Hanna M. Wallach, David Mimno, and Andrew McCallum, “Rethinking LDA: Why Priors Matter,” *NIPS* 2009: 1973–81.

for example, *nature* is an important word in multiple topics: in one, it consorts with *garden* and *flowers*; in another, with *biological* and *evolutionary*.

The algorithm completely ignores the syntax of the articles it processes: it looks only for groups of words that tend to occur in the same articles. This pattern-searching should not necessarily be understood as an attempt to reveal deep linguistic, psychological, or cultural structures. It is the task of the interpreter to decide, through further investigation, whether a topic's meaning is overt, covert, or simply illusory. The simplicity of the approach leaves many possibilities open, allowing us to map an archive in a very flexible way. The topics of topic modeling are not simply themes; they might also reflect rhetorical frames, cognitive schemata, or specialized idioms (of the sort that Bakhtin conceived as mixed together in social heteroglossia); if they are capacious enough, topics may even indicate a discourse in Foucault's sense.⁵

Consider a sample topic produced by allowing the algorithm to divide the same set of seven journals considered in figure 1 into 150 topics. This topic, arbitrarily numbered 80, features as its most common words *power*, *violence*, and *fear*. In figure 2, we list the words that are most common in this particular topic and chart the percentage of words in the journals that belonged to the topic in each year. Since topics are defined by context, this topic doesn't necessarily include all the occurrences of any word. For instance *act* occurs here (in a violent context where *deed* might be a good synonym), but *act* also occurs in another topic, in the company of words like *character* and *scene*. To this extent, topic modeling can divide a word into different senses. It doesn't try to guess whether *act* is functioning as a verb or a noun, but it can distinguish contexts where it's associated with violence from those where it's associated with theater.

Topics are interestingly slippery objects that require interpretation. *Violence* might be a reasonable one-word summary of topic 80, but it isn't a complete description. The most common word in the topic, after all, is *power*—a somewhat broader concept. The topic also includes strange details, like what appear to be the names of body parts: *blood*, *head*, *hands*, *face*, and *eyes*. There is a coherence to this list, but it may not be the kind of co-

5. For the connection between topics, frames, and heteroglossia, see Paul DiMaggio, Manish Nag, and David Blei, "Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding," *Poetics* 41, no. 6 (December 2013): 570–606.

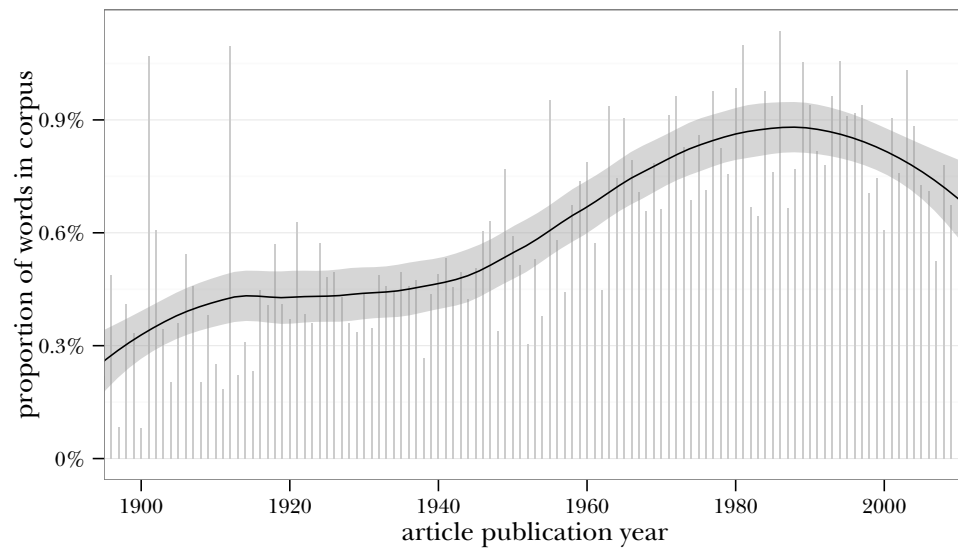


Figure 2: Topic 80 in our 150-topic model. Its most prominent words are: *power, violence, fear, blood, death, murder, act, guilt, face, violent, secret, head, crime, victim, hands, kill, evil, punishment, eyes, shame.*

herence we ordinarily associate with the term *topic*. These words belong together simply because they are in practice used together. For whatever reason—conceptual or psychological—eyes, faces, and hands get mentioned frequently when writers are talking about violence and power.

By itself, this detail is interesting. But the model also reveals a chronological trend: violence became much more prominent in literary-critical discourse over the course of the twentieth century. The frequency of the topic roughly triples between 1890 and 1980. The same is not true of print culture more broadly; in the Google Books Ngram Viewer, for instance, the frequency of these keywords doesn't increase.⁶ We have also contrasted this to a topic model of the *American Historical Review*, where no similar topic appears. This is a trend specific to literary study that eludes our current approach to disciplinary history because this topic was rarely thematized as a subject of debate. Its increasing prominence may have been an indirect consequence of new methodologies, but there was no moment between 1890 and 1980 when scholars explicitly decided to spend more time talking about power, violence, and fear.

Topic modeling is good at revealing quiet changes of this kind. In disciplinary history, changes can slip under the radar for many reasons—because they happened gradually, because they were secondary consequences of a different debate, because they were disavowed, or simply because no one decided to fight about them. Yet these stealthy changes often turn out to be both numerous and intellectually significant. After studying the subtle changes revealed by topic modeling, disciplinary histories organized exclusively around conscious debate start to look unsatisfying.

This is not to say that topic modeling will immediately provide a clear alternative history. There is at least as much room for interpretation and debate with this method as any other. To begin with, scholars have to define the archive they want to model. After an initial exploration focusing on *PMLA* alone,⁷ we chose a group of seven generalist literary-studies journals with long print runs, from Britain and the United States, with articles mostly in English but not exclusively about English-language literature:

6. See http://books.google.com/ngrams/graph?content=power%2Cviolence%2Cfear%2Cblood&year_start=1890&year_end=2000&corpus=15&smoothing=3&share=

7. For the earlier work, see Andrew Goldstone and Ted Underwood, "What Can Topic Models of *PMLA* Teach Us About the History of Literary Scholarship?" *Journal of Digital Humanities* 2, no. 1 (2012).

Critical Inquiry (1974–2013), *ELH* (1934–2013), *Modern Language Review* (1905–2013), *Modern Philology* (1903–2013), *New Literary History* (1969–2012), *PMLA* (1889–2007), and the *Review of English Studies* (1925–2012). Modeling different corpora has convinced us that the patterns we discuss here are durable; adding one more journal to the list would not overturn everything. But our archive is also defined by conscious exclusions: for instance, little magazines are not represented here. We also excluded book reviews, front matter, and anything shorter than 1000 words in order to focus solely on full-length scholarly articles. Other scholars will have valid reasons for studying a different corpus.

There are also significant technical choices to be made with this method. Having chosen the corpus, the researcher must then decide what constitute the data for the algorithm. In our procedure—a standard one—each article is reduced to a list of how many times each word occurs in each article, ignoring all punctuation and capitalization. We discard a list of very common words, first names, and other terms that in practice interfere with the discrimination of topics, as well as very infrequent words. We also accepted the error associated with using JSTOR's own word counts, provided by the Data for Research service.⁸

Most consequentially in practice, the researcher must also specify the number of topics to produce in advance. The model we discuss here is divided into 150 topics. We also produced models with 50, 100, 200, and 300 topics; although there were many zones of overlap between them, each of these models revealed slightly different patterns. In fact, because the modeling algorithm is probabilistic, the results are slightly different each time you run it. Thus a modeling algorithm cannot be a black box whose pronouncements are taken on faith; researchers should look inside the box and explore different ways of using it. We used a range of different modeling tools, and wrote our own tools, before we settled on a satisfactory method.⁹

8. <http://dfr.jstor.org>.

9. For technical notes on the method, see the appendix, which includes references to our supplementary information online.

THE HAZARDS OF QUANTITATIVE METHOD

No matter how carefully prepared the topic model may be, it is still important to remember that it is a model, a statistical simplification designed to make analyzing thousands of documents tractable.¹⁰ The approximateness of the model is visible throughout our analysis. In addition to topics which appear to join words together meaningfully, there are difficult-to-interpret topics featuring very common words; there are largely coherent topics with “intrusive” words that appear not to belong; and there is the omnipresent low-level froth of randomness, assigning small parts of each document to each topic in an arbitrary way. And behind it all is the flagrant simplification of considering only word *occurrences* and not word *order*.

These are all reasons for caution. Nonetheless, provided we have a sense of how to interpret the model—and how far it can be relied upon—we can use it to draw conclusions. Researchers using topic models of historical and contemporary texts have shown they can be reliable for broad generalizations, though it remains important to evaluate any model’s validity in each new application.¹¹ In our work, the algorithm’s many simplifying assumptions nonetheless yield meaningful insights into the history of literary scholarship. That is not to say that an alternate, more refined technique might not improve upon our conclusions; we hope that others will be able to build on our research, and to that end, we have made our software and our analysis available online (see the appendix for links).

Of course, many humanists are troubled less by the simplifications involved in a particular model than by the whole notion of quantitative method. Scholars worry above all that quantifiers will assume data can speak for themselves, bypassing the intricacies of interpretation that have been central to literary criticism in the last half-century. Thus, Alan Liu

10. For this preliminary study we have chosen to use one of the simplest versions of the Latent Dirichlet Allocation algorithm available. More elaborate versions exist, but they add considerable complexity to the task of interpreting results. Exploring more intricate models with more realistic assumptions is an important task for future work.

11. A non-technical discussion of the validity of Latent Dirichlet Allocation applied to a nineteenth-century newspaper can be found in Robert K. Nelson, introduction to *Mining the “Dispatch,”* <http://dsl.richmond.edu/dispatch/pages/intro>. The validity of a topic model of political speeches is addressed in Kevin M. Quinn et al., “How to Analyze Political Attention with Minimal Assumptions and Costs,” *American Journal of Political Science* 54, no. 1 (January 2010): 209–28.

has recently challenged the use of techniques, including topic modeling, that promise “tabula rasa interpretation,” that seek “to read texts algorithmically and discover word cohorts or clusters leading to themes without acting on an initial concept from an interpreter looking to confirm a particular theme.”¹² This promise is, Liu argues, false; the machine provides only false objectivity, and in the end we must always close the hermeneutic circle with human interpretation.

Liu is certainly right that machine learning does not obviate the need for interpretation. But topic modeling never pretends to dispense with interpretation: a “model” is explicitly a human abstraction, shaped by human assumptions about the kinds of patterns to be looked for.¹³ The point of quantitative modeling, in other words, is not to achieve machine-guaranteed objectivity. Our method does not require us to pretend that our data is “raw” or free from human interference.¹⁴ Instead, our approach tries to extend human interpretation over collections of documents that would otherwise be too large for a single interpreter to survey. We trade refinement for the ability to make comparable interpretations of thousands of documents.

This change of scale made possible by the computer does not free us from the need for an interpretive methodology. Ours is drawn both from literary hermeneutics and from the methodology of the social sciences. As literary scholars, we want to understand how a topic is concretely embedded in discourse; when we interpret a topic, we read specific articles where it is prominent. Going from individual examples to interpreting on a larger scale, however, is also a problem of long standing in the social sciences, especially in sociological content analysis. In content analysis, the language of a large body of texts (most typically, but not only, from the mass media) is “coded” or categorized, and the codings become the basis for inferences about those texts’ meanings in particular contexts. Content

12. Alan Liu, “The Meaning of the Digital Humanities,” *PMLA* 128, no. 2 (March 2013): 414.

13. Our approach is technically described as “unsupervised,” but this means only that the human interpreter doesn’t begin by characterizing specific topics; in a broader sense, the model still begins with human interpretive assumptions. For instance, the human modeler has to specify the number of topics and the degree of blurriness topics are expected to possess.

14. On raw data, see Lisa Gitelman and Virginia Jackson, introduction to “*Raw Data*” *Is an Oxymoron*, edited by Lisa Gitelman (Cambridge: MIT Press, 2013).

analysis has also long made use of computer-assisted techniques, and social scientists' debates in this domain, which have been highly sensitive to the challenges of interpreting texts and the compromises necessary for aggregation, offer a major resource for literary scholarship's attempts to understand large numbers of texts.¹⁵

Quantitative approaches to literary history like ours join in the wider renewal of interest in the sociology of literature.¹⁶ What is best in these new approaches is a shared determination to adapt concepts and techniques from the social sciences—including quantitative techniques—in order to enhance the nuance and precision of our interpretations of literary history. Thus, social-scientific reflections on quantitative content analysis emphasize the need to validate the categories of an analysis, both in terms of their internal coherence and their consistency with known facts about the texts at hand. And they remind us that even large-scale interpretations are contextualized: as Klaus Krippendorff says, what matters is that content analysts “make their chosen contexts explicit.”¹⁷ In our own case, we approach the texts of articles and the categories of the topic model as specialist readers knowledgeable about the terminology of literary study and the kinds of subjects literary scholars have discussed over

15. On the theory and method of content analysis, see Klaus Krippendorff, *Content Analysis: An Introduction to Its Methodology*, 3rd ed. (Los Angeles: Sage, 2013). DiMaggio, Nag, and Blei argue that Latent Dirichlet Allocation is particularly congenial to work in the sociology of culture in “Exploiting Affinities.” Recently, political scientists have been particularly interested in using Bayesian topic-modeling techniques similar to the one we have used; see Quinn et al., “How to Analyze Political Attention”; Justin Grimmer and Brandon M. Stewart, “Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts,” *Political Analysis* 21, no. 3 (Summer 2013): 267–97; and Justin Grimmer, “A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases,” *Political Analysis* 18, no. 1 (2010): 1–35. Computational linguists have used topic models to study academic discourse in particular; for a survey, focused on analyses of recent scholarship, see Daniel A. McFarland et al., “Differentiating Language Usage through Topic Models,” *Poetics* 41, no. 6 (December 2013): 607–25. For a sociologist's perspective on large-scale computer-assisted content analysis of historical texts, see Roberto Franzosi, *From Words to Numbers: Narrative, Data, and Social Science* (Cambridge: Cambridge University Press, 2004).

16. The diverse currents of the new sociologies of literature are surveyed in English, “Everywhere and Nowhere.” In a recent essay, Hoyt Long and Richard So provide a lucid exposition of connections between quantitative digital humanities and sociology, with a focus on network analysis: “Network Analysis and the Sociology of Modernism,” *boundary 2* 40, no. 2 (June 2013): 147–82.

17. Krippendorff, *Content Analysis*, 40.

the last century. And we believe that patterns in vocabulary are of particular significance in the world of scholarship, signaling methodological affinities, known topics of controversy, and underlying assumptions.

THE COMPLEXITY OF DIACHRONIC INTERPRETATION

As we have said, topic modeling is congenial to literary history because it divides words across topics, acknowledging that words have multiple senses and are used in different contexts. In this respect, the technique is subtler and more flexible than humanists might initially suppose. But that subtlety also raises problems of historical interpretation. For instance, recall the trend line for topic 80 *power violence fear blood*: it rose steadily to 1980 and then declined. Does this mean that questions of power became less central to literary study in recent decades? Not necessarily, because the word *power* is not restricted to one topic. In our model, it also figures prominently in many other topics—notably 10 *own power text form*, which emphasizes the way power is generated by textual or dramatic representation, and starts to rise just as topic 80 is declining (figure 3). In short, the chronological trajectory of a topic is not the same thing as the trajectory of the individual words that compose it. As Benjamin Schmidt has pointed out, this importantly complicates the interpretation of topic models.¹⁸ It is not, however, an insuperable obstacle. We label topics simply by listing their most common words, but when we interpret the changing frequency of a topic over time, we have also checked the trajectory of individual words to ensure that our interpretation is consistent with this finer level of analysis.

Another way to put this is to say that individual topics always need to be interpreted in the context of the larger model. The same thing is true of all historical evidence: we may try to trace a single theme or trope in an article, but there are always loose ends leading off in the direction of other problems, closely related but not quite the same. Much of the art of composing literary history lies in developing a set of gestures that allow an author to acknowledge these connections in an abbreviated way.

The complexity that makes topic models hard to interpret also makes them particularly good vehicles for revealing the intersections between

18. Benjamin M. Schmidt, “Words Alone: Dismantling Topic Models in the Humanities,” *Journal of Digital Humanities* 2, no. 1 (2012).

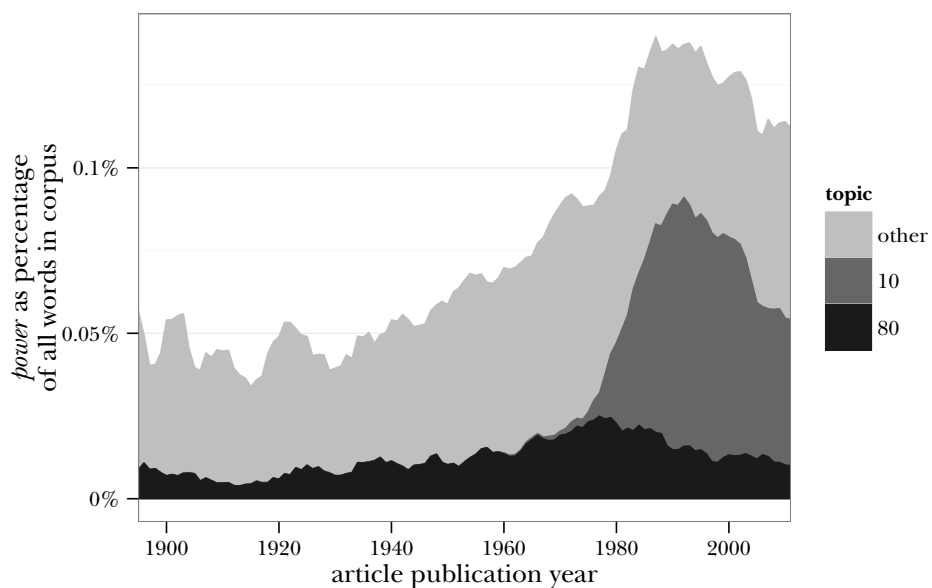


Figure 3: Yearly frequency of the word *power* in our corpus, smoothed as a five-year moving average, and broken down by the topics to which it is assigned in the model. The black region shows occurrences of *power* assigned to topic 80 *power violence fear blood*; the dark gray *power* assigned to topic 10 *own power text form*, and the lightest gray *power* in all other topics.

related social trends. For instance, the rising arc of *power violence fear blood* parallels several other trends in twentieth-century literary study. Topic 74, *death dead father life*, emphasizes mourning and elegy. Although its connection to violence is indirect, it appears in many of the same articles as topic 80, and shares the same trajectory. Is this actually a distinct topic? It depends on the interpretive frame. The model we spend most time discussing in this article divides our collection of literary journals into 150 topics, but a larger or smaller number could be equally valid. In one of our other models, divided into 100 topics, some of the key words in topics 74 and 80 were absorbed into a larger topic, *dream death myth magic*, which focuses on the monstrous, horrific, and uncanny.

The picture provided by the 100-topic model is obviously rather different. It includes all the same evidence: the same articles, the same words. But the evidence has been organized around a different set of centers. Working with multiple topic models of the same corpus, one finds surprisingly concrete confirmation for certain premises of recent critical theory. Discourses are always multiply determined, and lend themselves to multiple valid interpretations. This multiplicity isn't only produced at the margin—because we could change a corpus, for instance, by including or excluding authors. It's equally true at the center of the interpretive act, since the very same corpus can be divided in more than one persuasive way. We're always constituting some figure by excluding some ground, and there is usually more than one interesting pattern that could be produced. This isn't perhaps the lesson that humanists were expecting to learn from quantification. Our expectations have been shaped by more traditional forms of argument where numbers are deployed to prove a predetermined thesis. But topic modeling, and other contemporary applications of data mining, point toward a different hermeneutic, which humanists may find shockingly congenial.

At the same time, it's possible to learn something concrete from these techniques. Different models can slice the same corpus in different ways, but familiar patterns start to recur. For instance, themes of death and violence, broadly construed, clearly play a much larger role in literary study at the end of the twentieth century than they did at the beginning. Moreover, it's possible to rule out certain explanations for this shift. For instance, it might be tempting to suppose that increased critical discussion of these dark themes reflects mainly discussion of modern literature, which has a

reputation for rawness and candor. In fact, there's no evidence of that. In our corpus, the twenty articles where topic 80 *power violence fear blood* is most salient discuss Dante, Jacobean tragedy, Mary Shelley (three times), Charles Dickens (twice), Daniel Defoe, Herman Melville, Dostoevsky, Byron, Shakespeare (four times), and Samuel Richardson (twice), as well as Abu Ghraib.¹⁹ It's likely that twentieth-century critics were drawn to themes of violence for reasons that parallel the development of literary genres, but the shift in scholarship didn't happen primarily because critics were discussing modern authors.

It's reasonable to speculate that the rising importance of violent themes in twentieth-century scholarship reflects an underlying shift in the justification for literary study. Early-twentieth-century scholarship places a fair amount of emphasis on literature's aesthetically uplifting character (as in topic 66 *light like heart day*). As the century proceeds, that emphasis on aesthetic cultivation wanes, and appears to be replaced by a stance that one could characterize as ethical concern. But the transition between these rationales for literary study is not immediate. Instead, midcentury scholarship seems to bridge the gulf by retaining an overtly aesthetic/formal rationale for literary study, while quietly refocusing its formal attention on topics that are ethically provocative: violence, or mourning, or existential individualism (topic 82 *life world own man*). It's easy to miss a change of this kind if your model of disciplinary history is organized by methodological debate, because this change involves a gradual shift of subject rather than a new method.

We have described this explanation as speculative because the underlying causes of this shift are not transparently legible in the model itself. Causes beyond the one we've described are possible. But although topic modeling may not fully explain social causation, it can help us resist simplistic causal explanations that would otherwise seem plausible. For instance, it might have seemed natural to trace recent critical interest in power and violence to Michel Foucault's interest in both themes. (And there may be some truth to that account. Topic 10 in our model, *own power text form*, looks fairly Foucauldian and takes off like a rocket at the end of the 1970s.) But the larger context of our topic model also reveals that critical interest in power and violence was increasing steadily throughout the

19. The "most salient" documents are those in which a topic reaches its highest proportion. Even these documents always include other topics.

twentieth century. By resisting our tendency to attribute change to well known writers and “schools,” topic modeling can help us develop a nuanced, multifaceted account of the history of scholarship, where gradual social change overlaps with the influence of particular ideas.

BEYOND THE HISTORY OF “CRITICISM”

In describing our work as an investigation of the history of literary *scholarship*, we are already implying an argument: it would be more traditional to say that we are considering the history of literary *criticism*, in the manner of a standard reference like *The Cambridge History of Literary Criticism*. But, as Gerald Graff argues in his history, the centrality of criticism to the academic discipline of literature was a comparatively late development, produced by a series of conflicts: “Scholars versus Critics,” as several of his chapter titles have it (and then: “History versus Criticism”). Between 1915 and 1930, he argues, “scholar and critic emerge as antithetical terms”; the critics fight for a place in the academy in the subsequent decades; criticism achieves dominance after the Second World War; and then it finally fades into the background as conflicts over “theory” come to the fore.²⁰ Taking a long view of the institution, *criticism* is not a comprehensive label.

This established history provides a good test of our model’s validity. We can follow the historical rise and fall in the frequencies of topics which include *critic*, *critics*, *critical*, and *criticism* as frequent words: there are two of these, topic 16 *criticism work critical theory* and topic 94 *literary literature new work*. There are also several more in which only some *critic*-words are prominent. Topic 16 appears to correspond to explicit discussions of literary criticism; varieties of literary criticism, including New Criticism, are the subjects of its most prominent articles.²¹ The topic reaches maximum frequency between 1940 and 1960 before falling off through the 1980s (figure 4, top). Thus it appears true that the rising mentions of *critic*-words and the frequency of topic 16 do correspond to the period of criticism’s rise to hegemony, when it is a subject of active debate. We could stop there, having more or less validated our broadest historical premise. But if we compare the changing shape of topics to the historical trend of a single word, a more nuanced story emerges: the word *criticism* alone rises

20. Graff, *Professing Literature*, 122.

21. Citations are in the online supplement (see appendix).

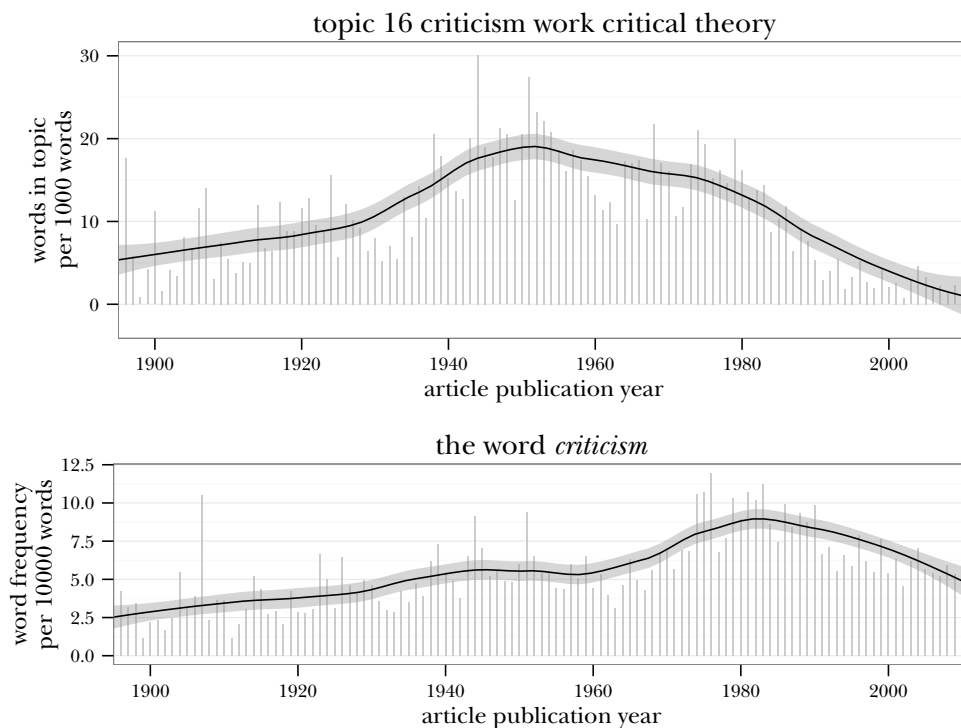


Figure 4: The rise and fall of criticism. *Above*: the proportion of the corpus occupied by topic 16 *criticism work critical theory*. *Below*: the proportion of the corpus occupied by the single word *criticism*.

again to a peak around 1980 (figure 4, bottom), but the model allocates these uses to a group of newer topics, none of which become prominent until the 1980s. These topics correspond to a variety of debates during the emergence of poststructuralist literary theory: predominantly topic 94 *literary literature new work*, but also topics 20 *reading text reader read*, 39 *interpretation meaning text theory*, and 143 *new cultural culture theory*.²² This transformation, too, is compatible with Graff’s description of a transition from a period in which criticism is a frequent focus of discussion to one in which it is part of the background to other methodological discussions.

Although this part of the story is still familiar, our model adds nuance

²² Some uses of most words, include *critic* and derivatives, are allocated to hard-to-interpret, semantically broad topics (such as 68 *see both modern early*), but not enough to affect the changes over time we are discussing.

to accounts that emphasize a few individual actors in conflict. It shows the emergence and subsequent naturalization of the discourse of criticism over the whole course of the twentieth century, reminding us that the very idea of the discipline of literary study as criticism is the product of a historical development. Of course, because our corpus includes only scholarly journals, not little magazines or general-interest periodicals, it is unsurprising that the word *criticism* does not feature prominently in the early part of the century in our model. What the model does emphasize, though, is the legitimization of criticism as it became a topic of discourse in field-encompassing scholarly journals like *PMLA* and *Modern Philology*.²³ The model helps to see why *Critical Inquiry* was a likelier name for a journal in 1974 than it would have been in 1924.

Indeed, our study of the corpus of journals helps to historicize a whole family of basic concepts (figure 5). As Sheldon Pollock has written, “The humbling force of genealogy must be part and parcel of every disciplinary practice.”²⁴ It is striking to see the late emergence of *reading* as a key word in scholarship: topic 20 *reading text reader read* takes off only in 1980 (accompanied by another latter-day topic in which *reading* is important, topic 10 *own power text form*). Though *reading* is not absent from our corpus in earlier years, our model suggests that reading did not emerge as a thematic concern of literary scholarship until quite recently. The same holds true of *interpretation*, sometimes taken to be the indispensable core skill of literary studies or even of all the humanities. Topic 39 *interpretation meaning text theory* is narrowly confined in time, peaking sharply in the decade between 1975 and 1985 (with much of the topic belonging to debates over deconstruction in *Critical Inquiry*). Again, in our model, interpretation, interpreters, and meanings become key words only late in the history of literary scholarship. In a sense, the era of “high” literary theory transformed the language of scholarship even more than we might assume: from the point of view of the topic model, we see not so much the introduction of a new interpretive methodology as the introduction of interpretation,

23. The isolated sharp peaks for topic 16 in the early part of the century are due to a small number of articles discussing criticism of the past: the most prominent pre-1930 articles for this topic are two lengthy philological dissertations in *PMLA* on “Friedrich Schlegel and Goethe, 1790–1802” (1906) and “The Development of John Dryden’s Literary Criticism” (1907).

24. Sheldon Pollock, “Future Philology? The Fate of a Soft Science in a Hard World,” *Critical Inquiry* 35, no. 4 (Summer 2009): 948.

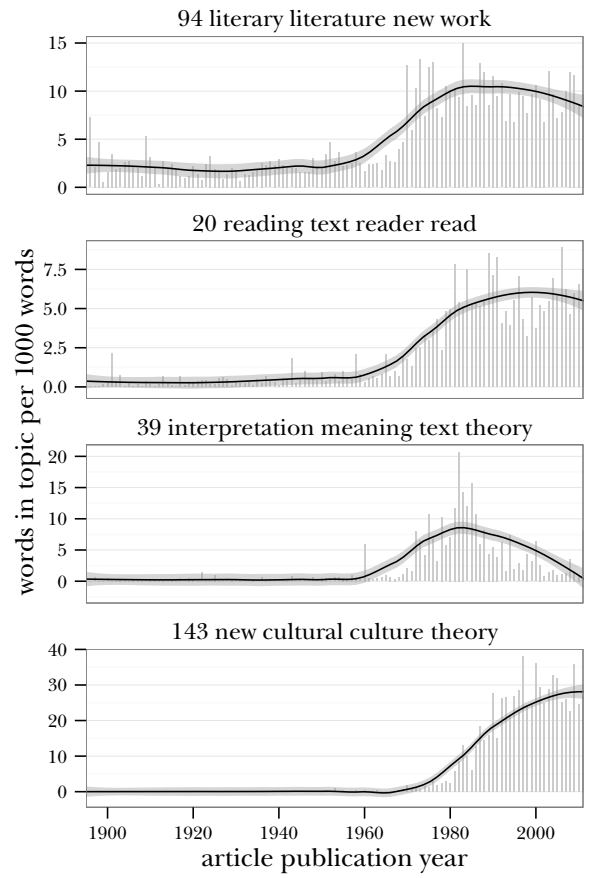


Figure 5: The emergence of “high” theory in the topic model.

meaning, and reading as themselves the main themes of scholarly discussion, especially in new journals like *Critical Inquiry* and *New Literary History*.²⁵

Indeed, the word *literature* itself, though always present in the corpus, is used more frequently after 1970. The late jump in prominence of topic 94 *literary literature new work* (see figure 5) gives another clue to the surprising rise in the scholarly fortunes of the concept. The model indicates that the conceptual building-blocks of contemporary literary study become prominent as scholarly key terms only in the decades after the war—and some not until the 1980s.²⁶ We suggest, speculatively, that this pattern testifies not to the rejection but to the *naturalization* of literary criticism in scholarship. It becomes part of the shared atmosphere of literary study, a taken-for-granted part of the *doxa* of literary scholarship.²⁷ Whereas in the pre-war decades, other, more descriptive modes of scholarship were important, the post-1970 discourses of the literary, interpretation, and reading all suggest a shared agreement that these are the true objects and aims of literary study—as the critics believed. If *criticism* itself was no longer the most prominent idea under discussion, this was likely due to the tacit acceptance of its premises, not their supersession. From this perspective we might return to the puzzle posed in figure 1, and suggest that the decline of enumeration in literary scholarship reflects the gradual naturalization of an assumption that the discipline's aims are fundamentally critical and interpretive rather than descriptive. The resurgence of historicism in recent decades may not yet have done much to alter this deeper disciplinary premise.

Topic modeling thus challenges presentist assumptions in method-

25. Needless to say, this does not mean earlier scholarship lacked theory, only that the key words of “Theory” really were a novel grouping. Even this claim must be qualified on the basis of our model: religious allegory and biblical interpretation are topics of long standing in literary scholarship, and topic 75 *god christ divine christian*, in which *allegory*, *allegorical*, and *interpretation* are all frequent words, constitutes between 0.5% and 1.5% of the corpus fairly regularly between 1895 and 1985. Topic 39 overtakes it only after 1980.

26. Readers may wonder whether the late prominence of *literary* is due only to its presence in the title of *New Literary History*. However, topic 94's striking post-1980 increase is visible even when all articles from *NLH* are eliminated. The trend is also validated by the models we constructed with different numbers of topics.

27. On literary doxa, see Pierre Bourdieu, *The Rules of Art: Genesis and Structure of the Literary Field*, trans. Susan Emanuel (Stanford: Stanford University Press, 1996), pt. 2, chap. 1.

ological debate by revealing the recent emergence of concepts we take as central to the discipline. It also challenges our assumptions in another sense, by breaking apart concepts we tend to treat as historical unities. Much discussion of twentieth-century literary theory is organized around a tension between historicism and formalism. We know that different approaches to form are possible; we tend to associate these different approaches with successive critical movements called “philology,” “New Criticism,” “structuralism,” and “post-structuralism.” But there are alternate ways of dividing this story. Instead of revealing a succession of critical movements (all discussing a shared object called *form*), our topic model tends to divide the concept of form itself into different facets. A full list of these would be too large to discuss here, but a few of them are visualized in figure 6. To make this picture simpler to interpret, we’ve chosen to graph the frequencies of word groups here instead of using topic frequencies, but these word groups are drawn from topics in our model and follow roughly the same chronological trajectory as those topics.²⁸

“Style” and “manner” are very durable concepts; in the early twentieth century they coexisted with a strong emphasis on versification, which has since tended to diminish. A critical vocabulary stressing imagery and symbolism begins to dominate discussion of literary form between 1940 and 1980; this correlates strongly with interest in the “pattern” organizing each work as a whole. At the same time, the contrast between literal and figurative language (especially metaphor) assumes the centrality we now associate with it.

Some aspects of this story are not surprising: we might have expected, for instance, that interest in all of these formal concepts would decline in the 1980s. Yet the overall picture here is rather different from our received account of a conflict between successive theoretical systems. Among other things, figure 6 suggests that this story was more cumulative than agonistic: while interest in different aspects of form has been volatile, none of the discourses represented here are finally displaced. Instead they build up like layers, and eventually coexist in roughly equal proportions.

But although some themes stretch across the century, pre-1945 schol-

28. We created these four word groups by selecting words that were relatively prominent in topics 17, 29, 53, and 118, and whose trajectories over the timeline correlated closely with the trajectories of those topics.

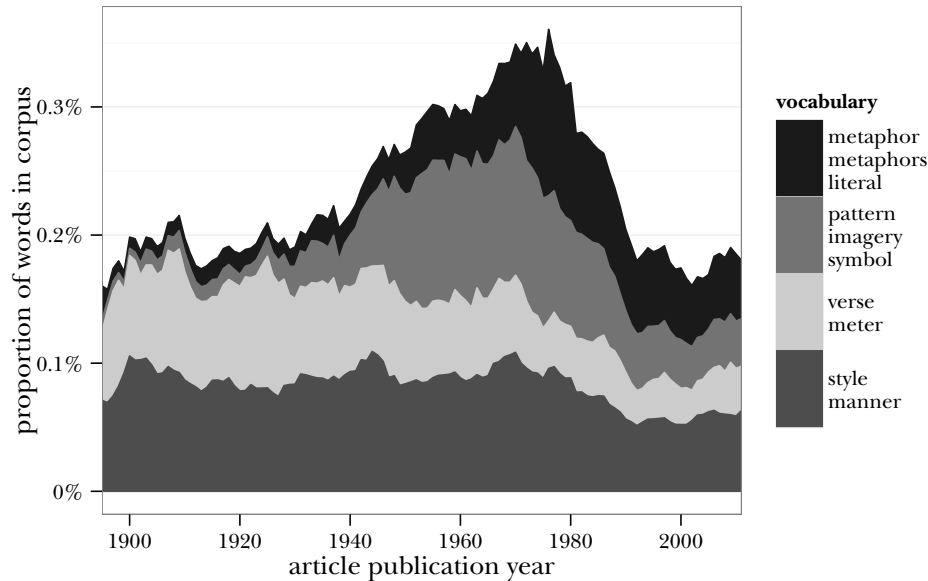


Figure 6: Overlapping waves of interest in different aspects of form.

arship was in other ways quite remote from us. Discussion of grammar was once far more central than it is now: topic 73 *verb examples use other*, whose exemplary articles have titles like “The Position of the Genitive in German” and “Impersonal *haber* in Old Spanish,” is very prominent in *PMLA* up to 1910, but captures little of the corpus after 1925. Similarly, topics that indicate textual scholarship, like topics 117 *text ms line reading*, 133 *ms manuscript fol manuscripts*, and 142 *edition first text printed*, are much more prominent before 1950 than after.

Our model also reveals another distinction between the scholarship published in the first half of our period and that published after: its use of material in foreign languages. The modeling algorithm regularly identifies “topics” of common words and names from particular languages: for example topics 34 *piu tasso canto perche* and 95 *nur hat man noch*. We counted 22 of the 150 topics focused on languages other than English (counting Old English, but not Middle English); of these, some 15 have downward trends after 1940.²⁹ The explicit discussion of grammar and

29. The exceptions are a French-language topic, 124 *plus cette aux meme*, and a Spanish-language topic, 45 *spanish mas spain entre*, which have some presence through about 1980 before again fading away. Foreign-language topics contribute much less to articles in our

the incorporation of foreign language text can contrast with topic 17 *language words word meaning*, a structuralist/poststructuralist language-and-metaphor topic. In fact, the words *words* and *word* themselves are, in the early part of the century, allocated to the philological topics described above, but, after 1960, they are predominantly part of that same topic 17. Our model gives indications of a turn from the polyglot philology of the early century to an increasingly monoglot theoretical discourse about language; non-English languages—especially medieval languages—simply take up less space over time.

This survey of major methodological trends has not uncovered any radically new movements; if it had, it would cast more doubt on the topic model than on our accepted history. But the model highlights the very broad scope of the major historical changes in the objects of literary study, and furnishes us with a significant negative conclusion: neither interpretation, nor criticism, nor form, nor texts, nor language itself can be thought of as the invariant core of the discipline of literary studies. Even key concepts of long standing, like those of language, form, and text, mutate in ways that their changing topic-affinities show. The existence of this marked historical variation should complicate any attempt to redirect literary studies *back* to a putative disciplinary core. Debates about “returns to philology” or “returns to form” tend to imply a more unified tradition for literary study than our topic models reveal. Thus, when Marjorie Levinson describes “New Formalism,” she speaks of attempts “to recover for teaching and scholarship in English some version of their traditional address to aesthetic form”; such attempts share, she says, a desire to reinstate “reading, understood in traditional terms as multilayered and integrative responsiveness to every element of the textual dimension.”³⁰ But our analysis recasts form as a multilayered, historically changing concept, and it gives reasons to wonder whether supposedly traditionalist “readers” are participating instead in the comparatively recent discourse of topic 20 *reading text reader read*.³¹

corpus in the last two decades.

30. Marjorie Levinson, “What Is New Formalism?” *PMLA* 122, no. 2 (March 2007): 559–60. It is impossible to resist noting that the most prominent topics in this notably ambivalent essay about formalisms are, according to our model, topics 143 *new cultural culture theory* and 58 *social work form own*.

31. For returns to philology, see Edward Said, “The Return to Philology,” chap. 3 in *Humanism and Democratic Criticism* (New York: Columbia University Press, 2004), and, most recently, Jerome McGann, “Philology in a New Key,” *Critical Inquiry* 39, no. 2 (Winter

PRESENT-DAY DISSENSUS

Having insisted so much on our model's power to reveal the variability of disciplinary watchwords, it is natural to ask what themes have risen to prominence in the last two decades. Here the topic model seems to us to comport with a widely shared understanding about the directions literary scholarship has taken. The historical and cultural turn of literary scholarship is attested to in a series of topics that begin their ascent in the late 1970s or 1980s and rise into the present (figure 7): topics 15 *history historical new modern*; 143 *new cultural culture theory*; 138 *social society public class*; and 58 *social work form own*, which labels Marxist terms of discourse.³² In the last three decades literary studies has, in our model, taken a striking turn towards discussing—or theorizing, or criticizing; the topic model does not tell us which—the subject matter of the social sciences. The remaining topics on the rise in the last three decades (figure 7, right) concern media and technology; politics; science; the postcolonial and transnational; economics; and the law. Only two recent topics, concerned with philosophy and sentiment, lack this social emphasis.

Though this historical and cultural turn is well known, our approach allows us to show the comprehensiveness of the shift in the discourses of literary study. One might have expected a sharp “New Historical” rise and fall to match the earlier critical or deconstructive spikes we discussed above, perhaps centering on discussions about historicist method; instead, we see a broad movement into a whole family of new themes that grow in parallel, all of them couched in at least partly novel vocabularies. The familiarity of the recent topics in figure 7 belies the extraordinary change in literary scholarship's orientation in the space of the last generation. In this case, unlike in the earlier transition from the interests of philological scholarship to those of criticism in the mid-century, the changing

2013): 327–46. Pollock's “Future Philology?” is a soberer account of the long past and doubtful future of philology.

32. The specifically Marxist character of topic 58 is made clear by a longer listing of its most prominent words (*own ideology society material production*) and documents (including several titles by Ernesto Laclau, Tony Bennett, Slavoj Žižek, and Fredric Jameson among the top twenty). There is one more late-rising topic, rather opaque because it includes some very common and polysemous words: topic 10 *own power text form*. But the topic includes Foucauldian key-terms like *order*, *discourse*, and *authority*, and the most prominent articles engage topics like Orientalism, the “author-function,” and submission.

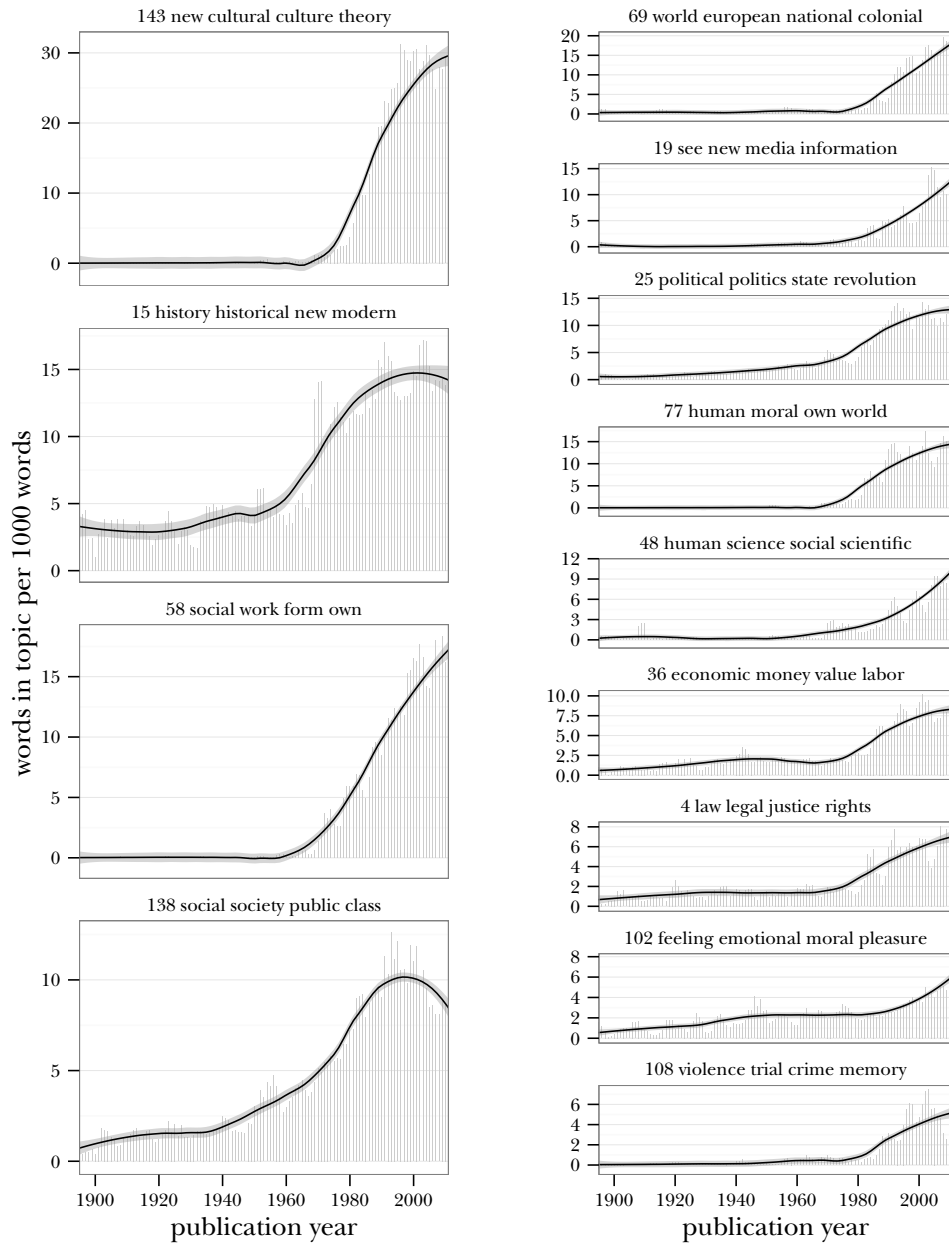


Figure 7: Topics on the rise in the last three decades overlap conspicuously with the subject matter of the social sciences. Note that the vertical scales of the small graphs are not all the same.

scholarly language gives fewer clues to changing method; it is alteration of subject matter that appears most consequential from the perspective of our model.

Such changes must play a role in continuing conflicts over the disciplinary autonomy and prestige of literary study: the earlier bases of claims to distinction—in philological scholarship, in “criticism,” in form, or in interpretive theory—are no longer as central to the discourse of the field. The emergent discourses are palpably interdisciplinary. This interdisciplinary shift is likely specific to literature scholarship; when we constructed a topic model of the *American Historical Review*, we did not see such marked thematic shifts into other disciplinary vocabularies in recent years. Our topic model does not allow us to explain the causes of the shift in literary scholarship. But it does, we believe, lend weight to accounts that emphasize literary studies’s move away from notions of the specifically literary and onto terrain shared with other disciplines. In her ethnographic study of academic judgment in interdisciplinary committees, Michèle Lamont contrasts the “lack of consensus” about objects of study and criteria of evaluation in literary study with other disciplines’ more settled standards of legitimacy. She suggests that in broadening the field, “English scholars may have indirectly lowered the value of purely literary analytical tools within their broader analytical toolkits.”³³ This broadening of analytical toolkits is confirmed by the changing topic-composition of articles in the 1990s and 2000s. Such changes have affected broad swathes of scholarly writing.

That does not mean, however, that our model proves the existence of a “crisis” in literary studies; such a conclusion lies beyond the scope of our analysis. And as James English has argued for the case of English studies, the crisis narrative ignores the implications of the continuing spread of English literature study in university curricula across the world. Still, our exploration of the corpus of scholarship diverges from the literary curricula he describes, which, according to his investigation, remain focused on canonical literature, having “succeeded all too well in containing the ‘threat’ posed by cultural studies.”³⁴ The status of the canon is not easy

33. Michèle Lamont, *How Professors Think: Inside the Curious World of Academic Judgment* (Cambridge: Harvard University Press, 2009), 76, 72–73.

34. James F. English, *The Global Future of English Studies* (Chichester, UK: Wiley-Blackwell, 2012), 150.

to track in our model—proper names present a challenge for this topic-modeling algorithm, and a scholarly article is in any case not the same thing as a syllabus. But in our model, the social, political, historical, and broadly cultural vocabularies that emerged in the period of cultural studies do not show a drop-off analogous to the retrenchment English observes in the curriculum. This suggests a divergence between research and teaching that would be highly consequential.

CONCLUSION

David Armitage and Jo Guldi have recently proposed that computational methods are well suited to support a return to the *longue durée*.³⁵ The long twentieth century described here is not quite Braudelian in scope, but we have similarly argued that topic modeling offers a new way to explore long timelines, allowing literary historians to dramatize changes that may be too gradual, too distributed, or too unconscious to condense into a case study. In saying this, however, we mean to offer a methodological resource, not make a claim about the true rhythms of literary history. The algorithm we used doesn't presuppose continuity, and nothing in our argument would forbid scholars from providing a Foucauldian or Kuhnian account of rupture. Although we have emphasized long-term trends that are otherwise difficult to trace, our model also reveals rapid transformations of scholarly language, like the sharp peak in linguistic and rhetorical terminology associated with deconstruction circa 1980. In fact, an overview of our model, with its many topics that rise sharply from around 1980, suggests that the diction of literary scholarship changed somewhat more rapidly in the last three decades of the twentieth century than it had before—an acceleration we have also confirmed using other evidence.³⁶

35. David Armitage and Jo Guldi, "The Return of the *Longue Durée*: An Anglo-American Perspective." Forthcoming in *Annales. Histoire, Science sociales* 69 (2014).

36. Setting aside the topic model, we examined decade-to-decade changes in word frequencies in *PMLA* alone (leaving other journals out of the picture, to avoid confounding the addition of new titles with overall changes in diction). We randomly selected an equal number of articles from each decade of *PMLA*, calculated the average word frequencies for that decade, and compared each decade to a decade 20 years later, using cosine similarity as a metric. The overall pace of change is more rapid after the 1970s than it had been in the earlier portion of the print run—which is not, of course, to diminish the importance of the gradual changes still underway.

Quantitative methods may be especially useful for characterizing long, gradual changes, because change of that sort is otherwise difficult to grasp. But the methods we used in this article don't prescribe a particular scale of historical analysis; on the contrary, one of their advantages is an ability to reveal overlapping phenomena on different scales, or even transformations of the pace of change itself.

The transformations of literary study, quiet and otherwise, are not over, and our readers may wonder where the methods applied in this study would themselves fit in the history of the discipline. It is tempting to frame computational methods as an intervention in familiar conflicts. Jerome McGann has described digital methods as “philology in a new key”; Franco Moretti and the members of the Stanford Literary Lab have spoken of “quantitative formalism”; and Stephen Best and Sharon Marcus include some computational methods within a tendency to turn away from hermeneutic suspicion and back to the “surfaces” of literature.³⁷ Although our discussion is based on an analysis of language, we do not believe that a topic-modeling exploration like ours must necessarily lead to formalism, philology, or surface reading. Our topic model certainly reveals trends on the linguistic surface of the scholarly journals we studied: for instance, a decline in the relative prominence of languages other than English. But it also reveals changes in critical practice and in the underlying social rationales for literary study. Unlike a surface reading—but like a sociological content analysis—it reveals patterns of representation that may not be visible to the individuals participating in them. In these respects, the quantitative approach we adopt fits rather well into what we've described as a recent tendency for literary studies to develop stronger connections to social science.

On the other hand, the fact that numbers are involved doesn't give computational methods a purely objective character. Like more familiar kinds of evidence, topic modeling requires interpretation. With thoughtful interpretation, it can support critical interventions we might not ordinarily associate with numbers. Topic modeling can help us critique presen-

37. McGann, “Philology in a New Key”; Franco Moretti, *Distant Reading* (London: Verso, 2013); Sarah Allison et al, “Quantitative Formalism: An Experiment” (Stanford: Stanford Literary Lab, 2011), <http://litlab.stanford.edu/LiteraryLabPamphlet1.pdf>; Stephen Best and Sharon Marcus, “Surface Reading: An Introduction,” *Representations* 108, no. 1 (November 2009): 1–21.

tism, by uncovering the recent emergence of supposedly traditional concepts. It can challenge reification, by revealing social changes that don't align with the slogans we have used to organize history. Finally, it can help us grapple with the multiply-determined character of historical change, by showing how trends overlap, and can be lumped or split to make more than one persuasive pattern. Quantitative methods are unfamiliar in literary study, and many observers have inferred that they must encode values alien to the discipline. But we have found topic modeling to be a remarkably flexible technique that shows the past of literary scholarship in a new light and supports its present aspirations in unexpected ways.

Rutgers University, New Brunswick
University of Illinois at Urbana–Champaign

APPENDIX: TECHNICAL DETAILS

The models we use in this article were generated by MALLET, which is becoming a standard tool for topic modeling in the humanities: Andrew Kachites McCallum, “MALLET: A Machine Learning for Language Toolkit” (2002), <http://mallet.cs.umass.edu>. We also made use of David Mimno’s “mallet: A Wrapper Around the Java Machine Learning Tool MALLET” (2013), <http://cran.r-project.org/web/packages/mallet/>.

The topic model described here can be explored interactively at <http://www.rci.rutgers.edu/~ag978/quiet/>. Our R language scripts for creating the model are available at <http://github.com/agoldst/tmhl>. We will supply the raw topic model data on request.

The corpus we modeled included 21,367 articles and 13,221 distinct author names (the latter only approximates the number of different authors). The articles are listed at <http://www.rci.rutgers.edu/~ag978/quiet/#/bib>.

We divided the corpus into 150 topics. We tried different settings of this parameter before settling on 150 topics as leading to a model whose results were most easily interpreted for our purposes. We label each topic with the four words most commonly assigned to it. In creating the model, we enabled the MALLET feature that optimizes “hyperparameters,” which can be understood as assumptions about the degree of blurriness to be expected in topic distributions over documents and words.

Our list of 6,970 “stop words” (words excluded from the modeling process) is available at <http://hdl.handle.net/2142/45709>. A list of stop words embodies an assumption that certain words are too common (for instance: determiners, prepositions, modal verbs) or too arbitrarily distributed (given names, abbreviations) to provide interpretable evidence. These lists can powerfully shape a model; we sought to be conservative about what we excluded, but other approaches could be equally valid. We also standardized to U.S. spelling and removed low-frequency words (words that were not among the top hundred thousand for this corpus).

JSTOR’s optical character recognition process does produce errors. We addressed this by using the stop list to exclude a few common categories of error (e.g., “ofthe”). We also noticed and corrected some obvious errors in the metadata. For more detail and justification on the rationale for these choices in topic modeling, see David Mimno, “Computational Historiography: Data Mining in a Century of Classics Journals,” *ACM Journal on Computing and Cultural Heritage* 5, no. 1 (April 2012): Article 3.

In our plots, the vertical axis shows the proportion of words in a given year after stop words have been discarded. The smoothing curve seen in most figures is produced by local regression (loess) smoothing; the lighter-colored band is an estimate of the smoother’s standard error. The plots with filled-in areas of color have been smoothed instead using a five-year moving average. Our plots show the years between 1895 and 2010. Though we included the full runs of journals in our model, the earliest years have too few data points to show clear trends on the chart, and the most recent years are unevenly covered by JSTOR.