

Pamphlet **10**  
→ October 2015

# Literary **Lab**

On Paragraphs. Scale,  
Themes, and Narrative Form

**Mark** Algee-Hewitt

**Ryan** Heuser

**Franco** Moretti

**Pamphlets of the Stanford Literary Lab**

ISSN 2164-1757 (online version)

Mark Algee-Hewitt  
Ryan Heuser  
Franco Moretti

## On Paragraphs. Scale, Themes, and Narrative Form

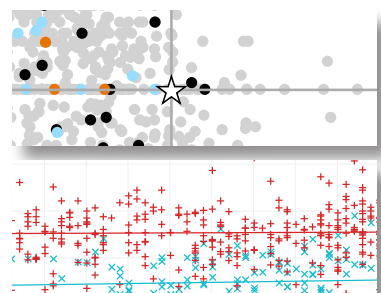
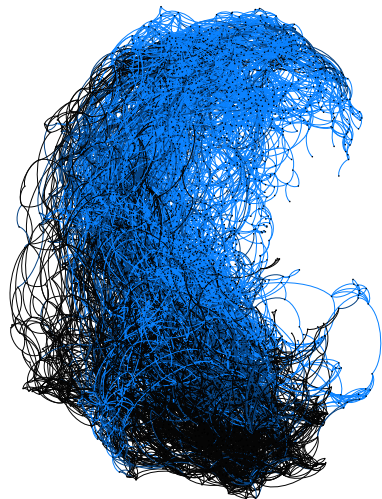
### 1. Micromégas

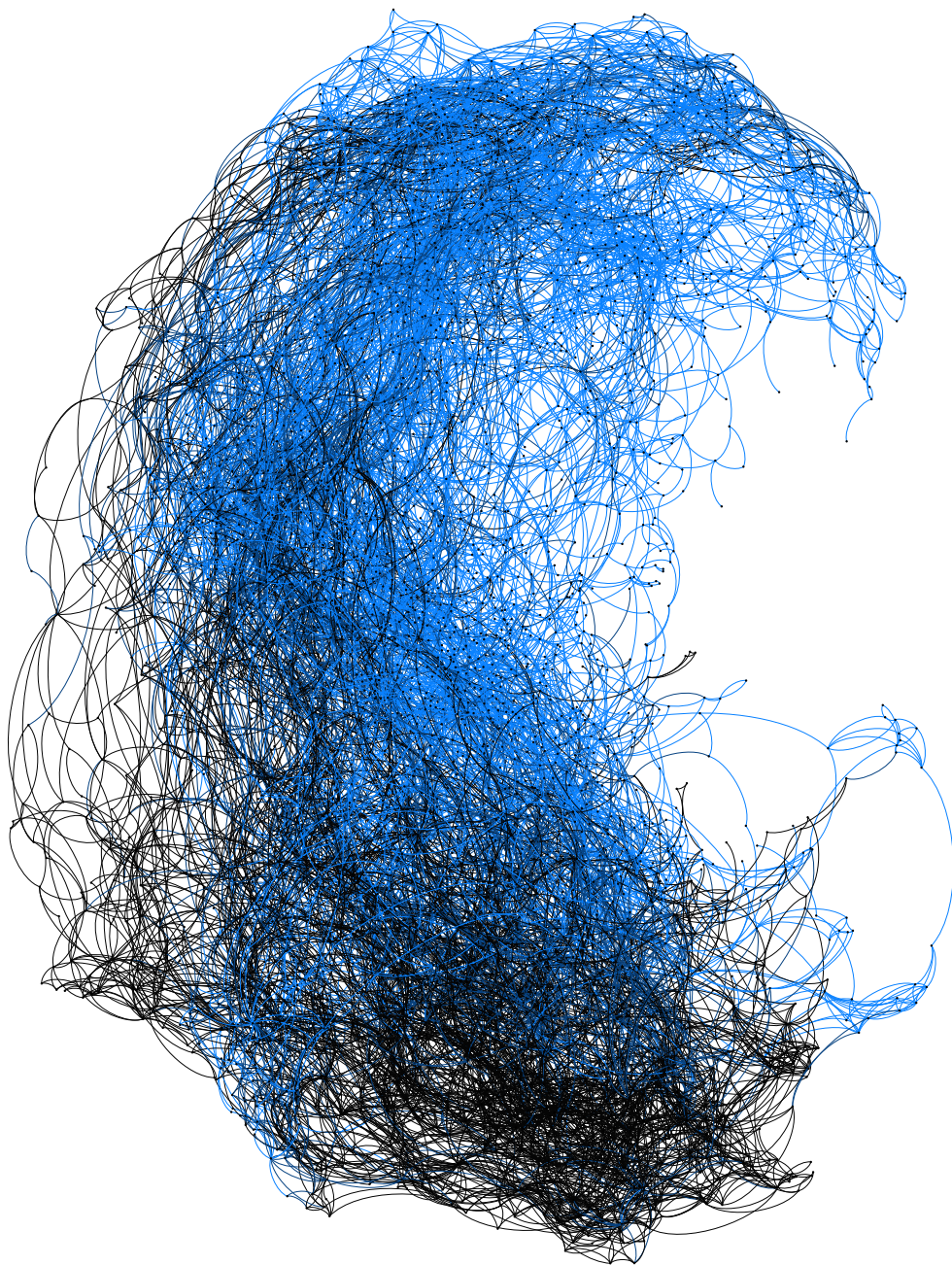
**Figure 1.1** is the final image, and book cover, of Matt Jockers' *Macroanalysis*: a network of over 3,000 nodes, each standing for a nineteenth-century novel, and of 165,000 edges among them, based on the similarity of 104 distinct features. The novels are mere dots, but there are so many of them, and with so many inter-connections, that the result is a big, complicated cloud.

Though the amount of information in this image is unusual, the conjunction of very small units and a very large outcome is typical of the digital humanities. Look at the diagram of "metricality" in **Figure 1.2**: two sets of very simple variables – rising-or-falling rhythm, and binary-or-ternary feet – and centuries of English poetry are mapped onto a single chart. Or take Mark Algee-Hewitt's use of bigrams to investigate nineteenth-century fiction in **Figure 1.3**: millions of minuscule two-word combinations, and a secret divide emerges from a century of novels.

By now, this mix of micro and macro has become the signature of the digital humanities, and of their dramatic impact on the scale at which literature is studied. Dramatic, because what we encounter in these images are

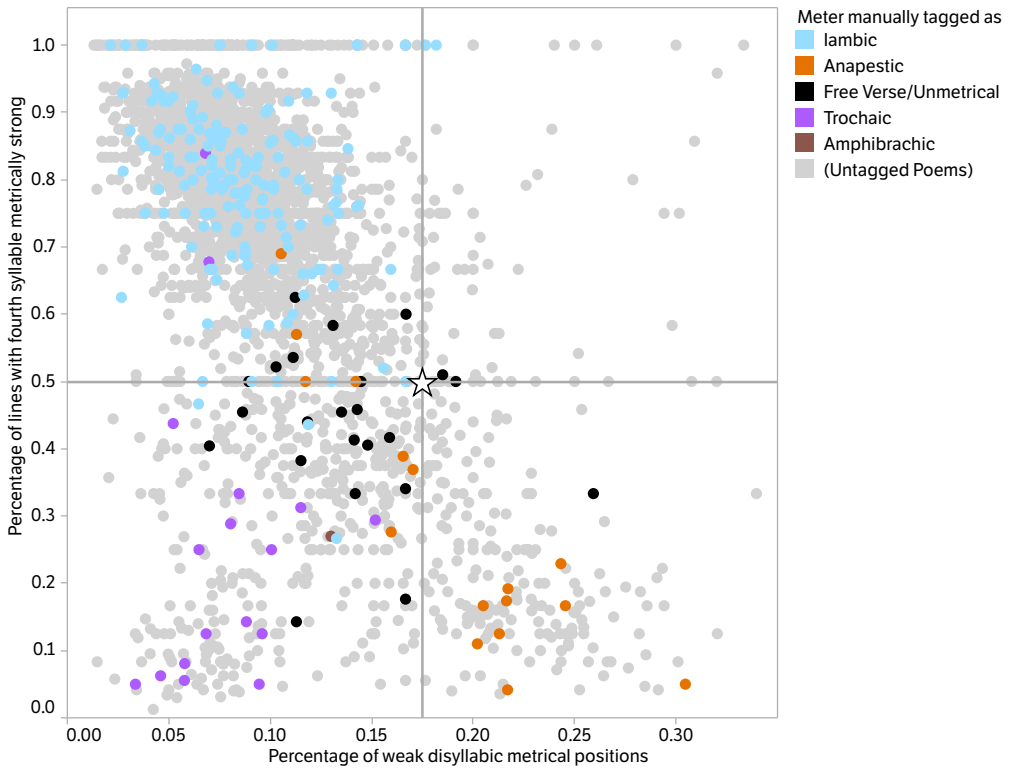
Excerpts of Figures 1.1, 1.2, & 1.3.  
See next two pages for larger images.





**Figure 11: A Network of 3,000 Novels**

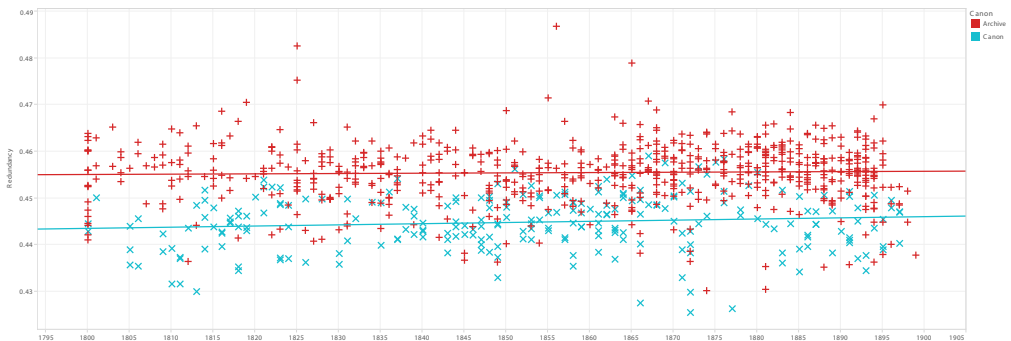
Matthew L. Jockers, *Macroanalysis. Digital Methods & Literary History*, Illinois UP, Urbana, Chicago, and Springfield, 2013.



**Figure 1.2 Mapping English Poetry**

A plot of 6400 poems, sampled evenly, period by period, from the sixteenth through the twentieth century. Colored dots represent a subsample of poems whose meter was manually annotated (see legend). Gray dots indicate poems whose meter is not directly known; their position within a specific quadrant indicates what our program considers their most likely meter (iambic, trochaic, anapestic, or dactylic). Black dots indicate poems in free verse, or with no discernible meter.

From Mark Algee-Hewitt, Ryan Heuser, J.D. Porter, Jonathan Sensenbaugh, Justin Tackett, “Trans-Historical Poetry Project”, Literary Lab.



**Figure 1.3. Canon and Archive in nineteenth-century Britain**

This image charts the frequency with which combinations of any two consecutive words repeat themselves in a sample of nineteenth-century novels: the greater the repetition, the less informative the given text. As can be seen, this very simple measure of linguistic redundancy reveals a striking difference between canonical texts (in blue) and non-canonical ones (red).

From Mark Algee-Hewitt, Sarah Allison, Marissa Gemma, Ryan Heuser, Franco Moretti, Hannah Walser, “Canon/Archive. Large-Scale Dynamics in the Literary Field”, Literary Lab.

*the extremes* of the literary scale, whereas criticism has traditionally worked with *the middle* of the scale: a text, a scene, a stanza, an episode, an excerpt ... An *anthropocentric* scale, where readers are truly “the measure of things”. But the digital humanities, Alan Liu has written, have changed these coordinates, by “focusing on microlevel linguistic features [...] that map directly over macrolevel phenomena”.<sup>1</sup> Exactly. And how does one study literature, in this new situation?

One option would be to focus exclusively on the very small and the very large. This is very much what happened with Leo Spitzer’s *Stilkritik* – one of the great theoretical precedents of computational criticism. For Spitzer, all that mattered were the “detail” and the “whole”; no middle scale, but only a long series of “back-and-forth movements (first the detail, then the whole, then another detail, etc.) [...] until the characteristic ‘click’ occurs, which is the indication that detail and whole have found a common denominator”.<sup>2</sup> From a certain type of conjunctions, to French symbolist poetry; from neologisms, to Rabelais’ *oeuvre* in its entirety; from the detail of noun deformation, to the worldview of *Don Quixote*. Detail and whole – very small and very large: Micromégas – and *only* detail and whole:

At its most perfect, the solution attained by means of the circular operation is a negation of steps: once attained, it tends to obliterate the steps leading up to it (one may remember the lion of medieval bestiaries – concludes the critic named Leo – who, at every step forward, wiped out his footprints with his tail, in order to elude his pursuers!) [37]

The present project is the opposite of Spitzer’s lovely simile: instead of “obliterating” intermediate steps, we want to make them totally explicit, by proposing a new middle scale for the study of literature. This scale, is that of the paragraph.

## 2. Auerbach, Watt, and the paragraph

Unlike sentences and (to a lesser extent) chapters, paragraphs remain an understudied scale of prose writing.<sup>3</sup> We all know they exist – we all *write* paragraphs all the time – but we

---

1 Alan Liu, “Where Is Cultural Criticism in the Digital Humanities?” <http://dhdebates.gc.cuny.edu/debates/text/20>

2 “Linguistics and Literary History”, 1948, in Leo Spitzer, *Representative Essays*, Stanford UP, 1988, pp. 36, 38.

3 The first systematic study of the paragraph in English appears to have been Edwin Herbert Lewis’ excellent Ph.D. dissertation *The History of the English Paragraph*, Chicago UP, 1894. For more recent studies, see Paul C. Rodgers, Jr., “Alexander Bain and the Rise of the Organic Paragraph”, *Quarterly Journal of Speech*, 1965, and “A Discourse-centered Rhetoric of the Paragraph”, *College Composition and Communication*, 1966; Francis Christensen, “A Generative Rhetoric of the Paragraph”, *College Composition and Communication*, 1965; and R. E. Longacre, “The Paragraph as a Grammatical Unit”, *Syntax and Semantics*, vol. 12, Academic Press, New York 1979.

Most of these scholars consider the structure and function of the paragraph to be fundamentally the same as that of the sentence, and they usually mention Alexander Bain’s *Manual of English Composition and Rhetoric* (1866) as the basis for this type of approach. “[Bain’s] paragraph”, writes for instance Rodgers, “is simply a sentence writ large [...] sentence and paragraph alike display an organic structure and employ the same means to secure it”; and then, summarizing the developments of paragraph theory in the late nineteenth century: “the paragraph now is an expanded sentence not only structurally but logically and semantically as well [...] devoted to the amplification and enforcement of the single idea announced in its topic sentence.” (“Alexander Bain”, pp. 406, 408). An even more radical continuity between distinct textual scales is asserted by Longacre, for whom “a paragraph resembles a long

don't really know how they work.<sup>4</sup> So, as a first step, we turned back to a previous pamphlet on literary scale – “Style at the Scale of the Sentence” – with the idea of studying how style changes when one shifts from the scale of the sentence to that of the paragraph. After all, wasn't the most programmatic essay of American stylistics entitled “On the first *paragraph* of Henry James' *The Ambassadors*”? And think of the opening words of the pilot essay for Auerbach's *Mimesis*: “Dieser Absatz steht im neunten Kapitel des ersten Teils...”<sup>5</sup> *Dieser Absatz*: this paragraph. *Mimesis* is beginning to take shape – the title of the 1937 essay is the first occurrence of that “serious imitation of the everyday” that will provide the *Leitmotiv* of the book – and “paragraph” is the first word that crosses Auerbach's mind.

Style at the scale of the paragraph. But there was something odd about the status of the paragraph in Watt and in Auerbach. They both analyzed style, and drew their evidence from paragraphs that were explicitly reproduced as such; the *form* of the paragraph, though, never entered their analysis. Paragraphs contained style, but they did not *shape* it. Tellingly, Watt referred to the opening of *The Ambassadors* three times as a “paragraph”, and *thirteen* times as a “passage”: although he was analyzing a paragraph, he didn't see it as one. Something similar happened in Auerbach's chapters on French realism: he referred to paragraphs obliquely, by turning to cognitive metaphors of a pictorial (the paragraph as “a ‘scene’”, “a portrait”, “a picture”, “the coarse realism [of] Dutch painting”) or musical (“a leading motif, which is several times repeated”, “the first words of the paragraph state the theme, and all that follows is but a development of it [...] a resumption, a variation”) nature.<sup>6</sup> Scene, portrait, picture, motif, theme ... For Auerbach, paragraphs were clearly not stylistic units, but *thematic* ones. And in the world of digital humanities, thematic study means: topic modeling.<sup>7</sup>

### 3. Thematic focus

Though thematic concepts may no longer be the “chaos” evoked by Propp in the opening pages of the *Morphology*,<sup>8</sup> most theorists agree that they remain disturbingly opaque, especially when it comes to the articulation of “theme” and “motif”.<sup>9</sup> Here, we will follow the

---

sentence on the one hand and a short discourse on the other hand.” (“The Paragraph as a Grammatical Unit”, p. 116).

4 Or at least: how they work *within narrative fiction*. Most studies of the paragraph have focused on non-literary discourse, and composition courses – where American students encounter the paragraph as form – are also structured around critical exposition, not narrative form.

5 See “Über die ernste Nachahmung des alltäglichen”, *Travaux du séminaire de philologie romane*, Istanbul 1937.

6 Erich Auerbach, *Mimesis*, 1946, Princeton U.P., 1974. pp. 455, 470, 483, 509, 510, 470, 484.

7 In what follows, we assume the existence of a relationship between thematics and topic modeling – so much so that we will use the adjectives “thematic” and “topical” almost interchangeably – though we are aware that our observations on this point have hardly a systematic character.

8 “If a division into categories is unsuccessful, the division according to theme leads to total chaos. We shall not even speak about the fact that such a complex, indefinite concept as ‘theme’ is either left completely undefined or is defined by every author in his own way.” Vladimir Propp, *Morphology of the Folktale*, 1927, Texas UP 1968, p. 7.

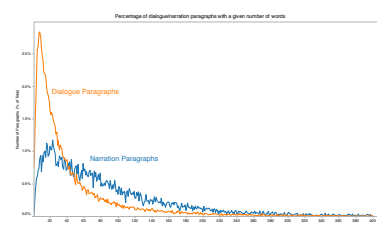
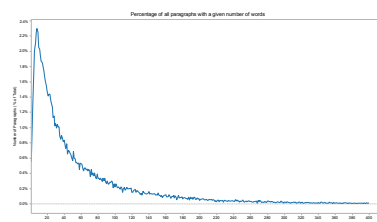
9 “The various distinctions that have been drawn between *theme* and *motif* remain vaguer still, so much so that their

widely shared assumption that themes tend to be large, abstract, synthetic,<sup>10</sup> and are usually not stated openly in the text, but produced by an act of interpretation;<sup>11</sup> whereas motifs tend to be explicit, delimited,<sup>12</sup> and concrete.<sup>13</sup>

Our initial hypothesis was simple: if paragraphs were indeed thematic units, then they would have a higher “thematic focus” than the abstract textual segments that were routinely used by topic modeling researchers.<sup>14</sup> Testing this hypothesis required however a series of preliminary steps, the first of which consisted in establishing the length of the paragraphs in our corpus. Initially, we measured *all* paragraphs, mixing narrative instances, like those analyzed by Auerbach and Watt, with exchanges of dialogue among characters (**Figure 3.1**). Though typographically marked in the same way, however, the two types of paragraphs played quite different roles in the architecture of the novel – especially in the mid-Victorian corpus we had selected – and we decided to disaggregate them, obtaining the results that are presented in **Figure 3.2**.

From this point onwards, we restricted our investigation to narrative paragraphs, along the lines of Auerbach and Watt, proceeding to a three-way

Excerpts of Figures 3.1 and 3.2. See next page for larger images.



definitions are often interchangeable.” Cesare Segre, “From Motif to Function and back again”, in Claude Bremond, Joshua Landy, Thomas Pavel, eds., *Thematics. New Approaches*, SUNY Press 1995, p. 22.

10 “The theme (what is being said in a work) unites the separate elements of a work [...] The idea expressed by the theme is the idea that *summarizes* and unifies the verbal material in the work”. Boris Tomashevsky, “Thematics”, 1925, in Lee T. Lemon and Marion J. Reis, eds, *Russian Formalist Criticism. Four Essays*, Nebraska UP, 1965, pp. 63, 67.

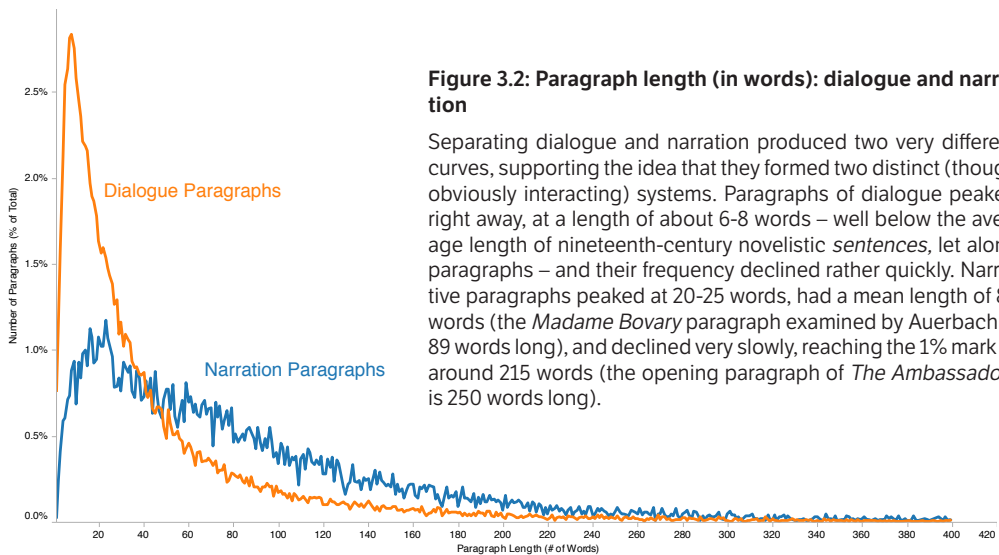
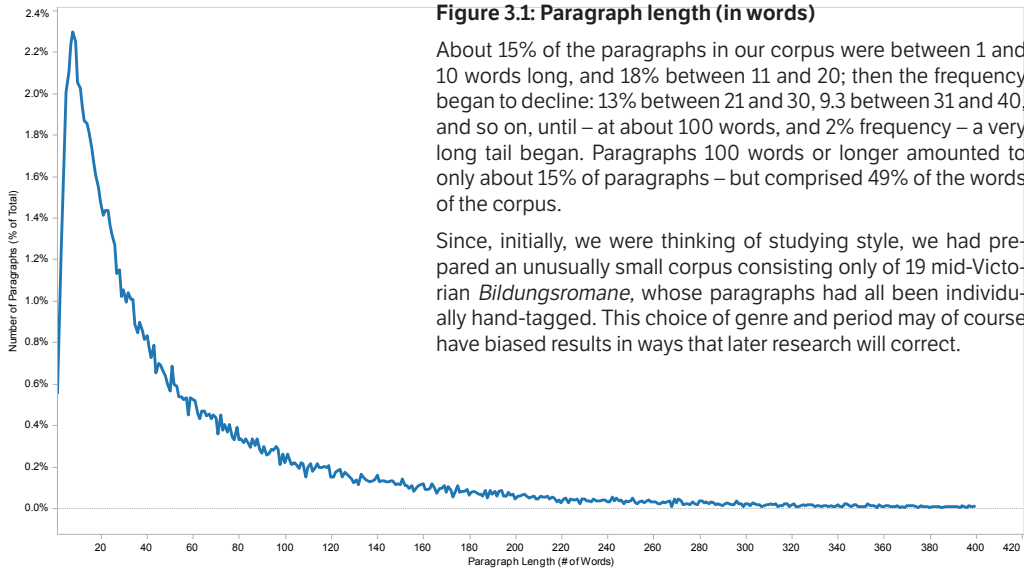
11 “A theme is not a component: there is no element in a literary work that can be called its theme [...] A theme is not an expression: although the theme is sometimes formulated explicitly, more usually it emerges implicitly, without corresponding to any specific expression in the text [...] A theme is not a segment within the text-continuum, but a construct put together from discontinuous elements in the text [...] a theme is a construct (a conceptual construct, to be precise), put together from discontinuous elements in the text.” Shlomith Rimmon-Kenan, “What is Theme and how do we get at it?”, in *Thematics. New Approaches*, cit., pp. 10-11, 14.

12 “After reducing a work to its thematic elements, we come to parts that are irreducible, the smallest particles of thematic material: “evening comes”, “Raskolnikov kills the old woman”, “the hero dies”, “the letter is received”, and so on. The theme of an irreducible part of a work is called the *motif*: each sentence, in fact, has its own motif.” Vladimir Propp, *Morphology of the Folktale*, cit., p. 67.

13 “Undoubtedly motif is more concrete, and theme more abstract”, concludes Bremond; while Segre, taking his cue from musicology, points out that “according to [various musicological] definitions, *theme* and *motif* are in a relationship of complex to simple, composite to unitary; [...] themes are elements that span an entire text or a considerable part thereof, while motifs – of which there might be many – are more localized elements.” See Claude Bremond, “Concept and Theme”, in Werner Sollors, ed., *The Return of Thematic Criticism*, Harvard UP 1993, pp. 49-50fn 2; and Cesare Segre, “From Motif to Function and back again”, cit., pp. 24, 25.

14 “Abstract”, in the sense that most researchers select as their units segments of equal length (usually 1,000 words), regardless of the actual subdivisions of the text; “mechanical” would be another way to describe the relationship between segment and text. In principle we have nothing against abstract and mechanical approaches to literature; it’s just that – in this specific case – we think that a different choice might produce better results.

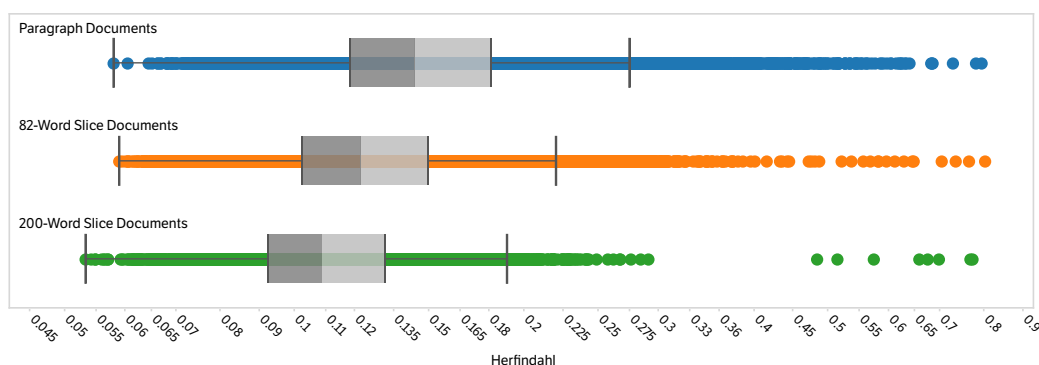




The chart suggested that nineteenth-century novels required readers to shift between long (and even very long) narrative stretches, and pointed dialogue among characters: a counterpoint of “written” and “oral” that seems to have crystallized in mid-century, and was later radicalized in the novels of Henry James. Flaubert offered his own striking version of the short/long alternation: the shock of his most legendary paragraph – “Il voyagea”, near the end of *Sentimental Education* – lies in seeing a paragraph *reduced* to the simplest of sentences (pronoun + intransitive verb). Implicitly, this radical compression reveals how different the functions of paragraphs and sentences usually are.

comparison between paragraphs, textual segments 82 words in length (82 being the mean length of narrative paragraphs), and a second group of segments 200 words in length (to approximate the 1,000-word size usual in topic modeling). In order to measure these three groups' thematic focus we borrowed from economics two statistical indicators – Gini's index of wealth inequality, and Herfindahl's measure of market concentration – that aim at establishing how finite resources (in our case, the number of words in a given paragraph) are distributed among different actors (in our case, the different topics present in the corpus). Combining the two measures, we determined how much of a paragraph's semantic space was concentrated into the “hands” of a single (or a few) topics. And, as **Figure 3.3** shows, thematic concentration turned out to be indeed significantly higher in paragraphs than in segments of equivalent length, and *much* higher than in 200-words segments.

Let us be clear about the meaning of these findings. First of all, we did *not* “discover” that paragraphs were thematic units; scholars who had studied the paragraph had long established this “fact”, which we had all learned in elementary school, and had “known” ever since. But we proved that this “well-known fact” was actually true, and could be “recognized” by a topic-modeling program, thus proving its reliability; two instances of corroboration which, though hardly exciting in themselves, have their modest role to play in the process of research. More significantly, our results suggest that – if one wants to use topic modeling to analyze literature – then *paragraphs are a better unit than “mechanical” segments*, and should replace them in future research. And the same for thematics: if, as we have seen, no one really knows “where” to look for themes in a text, our findings suggest that paragraphs are probably the best starting point: by concentrating thematic material within their limited space, they act *as the textual habitat of themes*. What this concretely means, is the object of the next two sections.



**Figure 3.3. Thematic Focus**

In this image, the line bisecting the three “boxes” indicates the median value for the group; the dark grey and light grey sections indicate the two central quartiles; and the “whiskers” represent the upper and lower quartile, with outliers indicated by individual dots.

We topic-modeled the three groups separately, on the basis of 50 topics, with hyperparameters on; for each segment, the model would consider at least 10 “Mallet” words, excluding names and very frequent words, and overlay the topics to the given segment.

## 4. Mono-topical paragraphs

Paragraphs specialize in the “concentration” of themes. But how, exactly? **Figures 4.1, 4.2, and 4.3**, offer the beginning of an answer, exemplifying what we ended up calling “mono-topical” paragraphs: that is to say, paragraphs in which a single topic occupies at least half of the available semantic space. On the basis of its top 50 words, such dominant topic could be described as “marriage and expectations” (**Figure 4.1**, from *Middlemarch*), “entering a house” (**Figure 4.2**, from *Villette*), and “direct, emotional communication” (**Figure 4.3**, from *Adam Bede*). Let us say right away that the semantic consistency of these topics is often questionable,<sup>15</sup> and that, whereas the paragraph from *Adam Bede* is semantically very close to the core of its dominant topic – many of whose 50 top words appear in the paragraph itself – the opposite is true of the *Middlemarch* case, where it’s hard not to feel a strong disconnect between the meaning of the paragraph and that of its supposedly dominant topic. A possible way out from this confusing state of affairs is offered by Claude Bremond, with his distinction between concept and theme:

There is in Rousseau’s work a concept of the social contract, but a theme of reverie [...] The one [i.e. the concept] takes off from the varied concrete and goes towards abstract unity. The other tends to exemplify a supposedly defined notion by immersing it in the context of various situations; it takes an abstract entity and makes it a point of departure for a series of concrete variations [...] theme overflows and incessantly calls into question the concepts forged in order to apprehend it. This is the consequence of *exemplification*: to the characteristics judged

15 In the *Middlemarch* passage, for instance, the terms “liberality”, “personal independence”, and “expectations” suggest the idea of “personal identity”, whereas “reason”, “imagine”, “suppose”, and “prove” point towards some kind of interpretive activity. The same in *Villette* – where “garden”, “door”, “cabinet”, “steps”, and “round” refer to space, and “loud”, “cautious”, “tinkle”, and “parley” to sound – and in *Adam Bede*, where “felt”, “deep”, and “sign” indicate emotions, and “stood” and “move” location. These contradictions are ubiquitous in topic modeling, and clearly weaken its analytical power; in “Words Alone: Dismantling Topic Models in the Humanities”, *Journal of Digital Humanities*, Winter 2012, Benjamin M. Schmidt has convincingly explained this Janus-faced behavior with the excessively long historical arcs of the corpora from which topics are extracted. It’s a very plausible hypothesis, which however doesn’t apply to our corpus, which is entirely drawn from a single, rather homogenous generation.

Excerpts of Figures 4.1, 4.2, and 4.3; larger images next two pages.

“Then, why don't you extend your li  
independence is as important to me  
have personal expectations from Br  
expectations from Bulstrode. Motiv  
But as to money and place in the wc  
clear that I am not determined by cc

– *Middlemarch: A Study of Provincial Life*

Topic #	Mallet Topic Score	Assigned Topic Score	
0	60.47%	21 / 30 [70.0%]	mind, marriage, husband, feeling, knew, point, makir determined, aware, sense, g opportunity, view, persona n't, dear, sir, suppose, give,

All at once, quick rang the bell-- qui  
whisper. Rosine darted from her cal  
two minutes in parley: there seeme  
hand; she stood on the steps, lifting

– *Villette*, by Charlotte Bronte (1853)

Topic #	Mallet Topic Score	Assigned Topic Score	
14	56.9%	20 / 30 [66.7%]	door, room, opened, open, passed, side, doors, place, s library, end, key, large, ser
19	12.21%	4 / 30 [13.3%]	looked, eyes, voice, turned, silent, silence, began, gave, room, sat, knew, passed, us speech, questions hand, eyes, face, head, han

Slowly, while Dinah was speaking, H  
arms. They stood so a long while, fo  
without any distinct thought of it, h  
she was sinking helpless in a dark g  
welcomed by the wretched lost one.  
down on the straw pallet together, t

– *Adam Bede*, by George Eliot (1859)

Topic #	Mallet Topic Score	Assigned Topic Score	
48	46.68%	16 / 30 [53.3%]	looked, eyes, face, turned, h silent, minutes, speaking, lo quick, answer, met, half, slo
25	23.3%	8 / 30 [26.7%]	heart, poor, tears, face, death hard, broken, strange, grief,

“Then, why don't you extend your liberality to others?” said Will, still nettled. “My personal independence is as important to me as yours is to you. You have no more reason to imagine that I have personal expectations from Brooke, than I have to imagine that you have personal expectations from Bulstrode. Motives are points of honor, I suppose-- nobody can prove them. But as to money and place in the world,” Will ended, tossing back his head, “I think it is pretty clear that I am not determined by considerations of that sort.”

— *Middlemarch: A Study of Provincial Life*, by George Eliot (1872)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
0	60.47%	21 / 30 [70.0%]	mind, marriage, husband, fact, present, wife, time, subject, question, kind, opinion, family, wished, felt, sort, give, feeling, knew, point, making, position, interest, grandcourt, idea, expected, reason, case, regard, uncle, means, determined, aware, sense, giving, side, circumstances, reasons, person, affair, held, occasion, future, general, object, opportunity, view, personal, find, speak
22	12.57%	4 / 30 [13.3%]	n't, dear, sir, suppose, give, hear, mind, things, wo, care, woman, father, call, remember, speak, ah, talk, married, friend, poor, feel, glad, continued, wife, bad, fellow, tone, boy, mine, understand, husband, people, leave, word, live, aunt, girl, stay, trouble, papa, place, bear, sort, head, pretty, exclaimed, kind, matter, fool
39	6.3%	2 / 30 [6.7%]	things, people, world, men, find, time, knew, life, feel, women, matter, set, wrong, times, end, sort, deal, hard, bad, person, work, suppose, kind, poor, mind, worse, half, lose, felt, fancy, wanted, making, friends, true, ways, doubt, care, show, story, easy, clever, thoughts, reason, give, fault, stupid, place, began, spite
42	6.21%	2 / 30 [6.7%]	hand, head, looked, chair, round, hands, side, turned, hat, eyes, table, face, sat, forward, arm, sit, seated, standing, room, stood, seat, sitting, turning, window, feet, walk, end, corner, distance, rising, held, began, leaning, close, book, laid, holding, threw, suddenly, time, fingers, walked, open, ground, turn, fire, entered, ring, sofa
27	2.99%	1 / 30 [3.3%]	dress, white, hair, black, wear, bonnet, silk, pink, wore, gold, dressed, round, small, red, gown, blue, lace, handkerchief, diamonds, necklace, shawl, cap, large, silver, clothes, frock, plain, pretty, curls, ring, short, green, jewels, tied, brown, bright, satin, suit, handsome, neck, mama, gray, muslin, yellow, robe, gloves, fingers, bracelet, colored

Figure 4.1

### Figures 4.1, 4.2, 4.3. Mono-topical paragraphs.

In these figures, each of the 50 topics created by the topic-modeling program is indicated by a number and color, making it possible to recognize its presence in the body of the paragraph: topic 0, blue/purple, dominates figure 4.1, topic 14, teal, figure 4.2, and topic 48, violet/purple, figure 4.3.

All the paragraphs we discuss from now on consist of 30 “Mallet” words: a length that correlates with the highest combined Gini/Herfindahl scores, and is therefore ideal to investigate thematic focus. Since Mallet does not consider proper names, function words, and other very frequent entities, a paragraph with 30 Mallet words has on average about 100 actual words.

Of the two scores on the left of the image, “Mallet Topic Score” indicates the amount of paragraph space that Mallet considers occupied by a given topic. As most words participate in several topics (though, usually, with a different “rank” within them), all 50 topics have a non-zero Mallet score in every paragraph in the corpus. “Assigned Topic Score” indicates for its part how many of the 30 Mallet words in the given paragraph are being “assigned” by the program to a specific topic. In this case, the only topics with non-zero values are those which have been assigned at least one of the 30 Mallet words.

Finally, we have run the model five times over the entire corpus (with very similar results) and are drawing our examples from all of these runs. As a consequence, there may be slight inconsistencies in the numbering and coloring of specific topics.

All at once, [quick](#) rang the bell-- [quick](#), but not [loud](#)-- a [cautious tinkle](#)-- a [sort of warning](#), [metal whisper](#). Rosine [darted](#) from her [cabinet](#) and ran to [open](#). The [person](#) she [admitted stood](#) with her two [minutes](#) in [parley](#): there seemed a [demur](#), a [delay](#). Rosine came to the [garden door](#), [lamp](#) in [hand](#); she [stood](#) on the [steps](#), [lifting](#) her [lamp](#), looking [round vaguely](#).

— *Villette*, by Charlotte Bronte (1853)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
<a href="#">14</a>	56.9%	20 / 30 [66.7%]	door, room, opened, open, house, stood, entered, window, heard, hall, looked, stairs, fire, light, table, round, closed, shut, passed, side, doors, place, step, led, parlor, small, dark, chamber, staircase, front, steps, bed, floor, corner, furniture, glass, library, end, key, large, servant, hand, candle, servants, chair, sat, rooms, brought, empty
<a href="#">19</a>	12.21%	4 / 30 [13.3%]	looked, eyes, voice, turned, tone, face, spoke, felt, heard, speak, smile, answer, speaking, glance, manner, time, eye, silent, silence, began, gave, met, pause, question, stranger, word, round, surprise, hand, expression, strange, turning, room, sat, knew, passed, usual, low, reply, feeling, quick, quickly, smiled, quiet, conversation, slight, countenance, speech, questions
<a href="#">43</a>	12.05%	4 / 30 [13.3%]	hand, eyes, face, head, hands, tears, looked, lips, arms, heart, voice, turned, arm, shook, father, held, round, kissed, mother, pale, stood, pressed, kiss, felt, spoke, fell, laid, cheek, child, cry, touch, bent, gave, neck, touched, door, speak, sat, answer, whispered, trembling, burst, smiled, lifted, whisper, silence, paused, shoulder, suddenly
<a href="#">25</a>	3.57%	1 / 30 [3.3%]	felt, mind, feeling, sense, life, strong, consciousness, mother, presence, husband, brought, painful, feel, dread, effect, state, conscious, future, sort, nature, reason, present, longer, power, feelings, pain, effort, experience, trouble, thoughts, fear, turned, speech, creature, moments, change, wanted, stronger, brother, began, strength, pride, care, pity, turn, ready, imagination, hard, shock
<a href="#">33</a>	3.12%	1 / 30 [3.3%]	eyes, face, hair, looked, eye, head, beauty, figure, woman, features, dark, light, white, black, mouth, large, smile, brow, tall, fine, expression, girl, full, blue, delicate, pale, nose, handsome, glance, color, soft, red, dress, fair, countenance, brown, thin, air, bright, pretty, sort, lips, appearance, set, forehead, shape, contrast, fire, age

Figure 4.2

[Slowly](#), while Dinah was [speaking](#), Hetty rose, took a [step forward](#), and was [clasped](#) in Dinah's [arms](#). They [stood](#) so a long while, for neither of them [felt](#) the [impulse](#) to [move](#) apart again. Hetty, without any [distinct](#) thought of it, hung on this something that was come to [clasp](#) her now, while she was [sinking helpless](#) in a [dark gulf](#); and Dinah [felt](#) a [deep](#) joy in the first [sign](#) that her love was [welcomed](#) by the [wretched](#) [lost](#) one. The [light](#) got [fainter](#) as they [stood](#), and when at last they [sat](#) down on the [straw pallet](#) together, their [faces](#) had become [indistinct](#).

— *Adam Bede*, by George Eliot (1859)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
<a href="#">48</a>	46.68%	16 / 30 [53.3%]	looked, eyes, face, turned, hand, voice, sat, stood, spoke, round, heard, silence, tone, time, speak, felt, room, smile, door, silent, minutes, speaking, low, glance, word, eye, turning, began, passed, paused, suddenly, head, expression, manner, quick, answer, met, half, slowly, standing, immediately, arm, walked, lips, pale, usual, moved, gave, strange
<a href="#">25</a>	23.3%	8 / 30 [26.7%]	heart, poor, tears, face, death, dead, cry, felt, eyes, fear, pain, die, soul, child, knew, fell, bear, time, woman, voice, cold, hard, broken, strange, grief, life, bitter, weak, burst, hands, body, terror, deep, strength, hand, lips, sight, pale, cruel, misery, trembling, comfort, broke, suffering, despair, sorrow, lost, terrible, breath
<a href="#">21</a>	8.94%	3 / 30 [10.0%]	mother, father, child, horn, poor, children, daughter, girl, looked, baby, wc, mamma, years, boy, dear, woman, wife, darling, husband, sister, knew, world, brother, time, brought, face, beautiful, sweet, nurse, mid, dead, house, heart, happy, boys, papa, pretty, loved, arms, parents, grave, remember, called, save, fond, quiet, day, whispered, talked
<a href="#">31</a>	3.06%	1 / 30 [3.3%]	white, black, dress, hair, looked, bonnet, round, wear, clothes, wore, large, coat, dressed, gold, silk, hat, cap, gray, gown, blue, pretty, red, small, gloves, shawl, hands, lace, waistcoat, pink, pair, shoes, head, clean, yellow, green, tied, boots, box, neck, silver, figure, handkerchief, pocket, eyes, worn, brown, gentleman, light, hanging
<a href="#">28</a>	3.13%	1 / 30 [3.3%]	door, room, house, opened, open, window, table, stood, looked, entered, hall, large, light, small, chair, side, place, stairs, dark, windows, parlor, rooms, floor, staircase, closed, key, corner, furniture, round, walls, shut, glass, fire, hand, led, passage, chamber, locked, end, opposite, appeared, books, doors, wall, standing, seated, library, sitting, gallery
<a href="#">5</a>	3.45%	1 / 30 [3.3%]	family, gentleman, person, knew, acquaintance, called, fact, society, conversation, friends, present, pleasure, manner, talk, time, house, gentlemen, kind, friend, held, appeared, character, ladies, subject, general, fine, air, business, visit, appearance, pleased, interest, received, place, company, observed, excellent, agreeable, daughter, mentioned, occasion, attention, private, brought, talked, neighborhood, high, generally, persons

Figure 4.3

as pertinent for the definition of concept, theme adds a network of associated ideas [...] Thematization therefore consists of an indefinite series of variations on a theme whose conceptualization [...] still remains to be completed.<sup>16</sup>

“Variations” that generate “networks of associated ideas”: this is the key. The *Middlemarch* paragraph exemplifies the theme of “Personal Independence” by placing it in the context of comparable choices (Lydgate, to whom the words are addressed, faces very similar problems to Will’s), thus providing a variation that “calls into question” the initial idea. Furthermore, the idea of themes as an “indefinite series of variations” explains why the internal contradictions of topics exist in the first place: they arise from that open-ended activity of association that – far from being an unfortunate accident – is for Bremond *the very point of* thematization.

Bremond’s thesis accounted for the semantic contradictions of topics; *removing* those contradictions was however a different story – and, finding ourselves unable to do so, we concluded that topics remained too slippery for an in-depth semantic analysis of our corpus. So, we turned once again to “Style at the Scale of the Sentence”, but this time to completely reverse its approach: instead of taking the scale of the sentence as the means to investigate the literary phenomenon of style, the other way around: we would take the literary phenomenon of themes, and use it to study the scale of the paragraph. This reversal of ends and means had a memorable precedent in Tomashevsky’s 1925 essay, where the distinction between “bound” and “free” motifs – which was supposed to throw light on their thematic content – ended up having nothing to do with thematics, and everything to do with the theory of plot.<sup>17</sup> Something similar here: mono-topical paragraphs interested us less for what their central topic “meant”, than for what it “did” in the narrative structure: defining a major character (*Middlemarch*), introducing an enigma (*Villette*), or announcing a major turning point in the plot (*Adam Bede*). All functions, we realized, that made mono-topical paragraphs quite similar to Tomashevsky’s bound motifs, thus opening an unexpected and promising path from thematics to narratology.

But there was a problem. Unlike fairy tales, modern novels are known to have very few bound motives; and indeed, once we calculated how many mono-topical paragraphs were to be found in our corpus, the total oscillated between 1 and 4% (**Figure 4.4**). And the other paragraphs?

---

16 Claude Bremond, “Concept and Theme”, in Werner Sollors, ed., *The Return of Thematic Criticism*, Harvard UP 1993, pp. 47, 48, 49. Bremond’s notion of “exemplification” is not to be confused with that of “illustration”: whereas the former places an abstract notion in the world of concrete phenomena, emphasizing its potential for metamorphosis, the latter uses the world of concrete phenomena to prove the inalterable validity of a general notion (as, say, in Harriet Martineau’s *Illustrations of Political Economy*). Illustration is interested in the fable because it allows to reassert the moral; exemplification is interested in the moral because it allows to multiply fables.

17 “Mutually related motifs form the thematic bonds of the work. From this point of view, the story is the aggregate of motifs in their logical, causal-chronological order [...] The motifs which cannot be omitted are *bound motifs*; those which may be omitted without disturbing the whole causal-chronological course of events are *free motifs*.” Boris Tomashevsky, “Thematics”, cit., p. 68. Notice the silent slippage from “thematic bonds” to “causal-chronological order”.

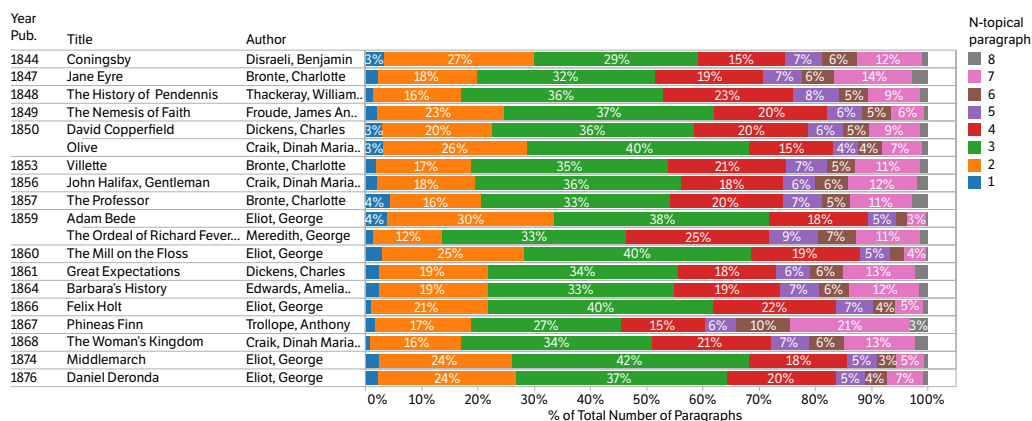


Figure 4.4. Thematic concentration in mid-Victorian *Bildungsromane*

We will return again to the contents of this chart. For now, let's simply point out that the highest frequency of mono-topical paragraphs is a mere 4% – one paragraph every 25 – and in just two of our 19 texts: *The Professor* and *Adam Bede*.

## 5. Physiology of the Paragraph

Faced with the evidence of Figure 4.4, our first impulse was to run to the opposite end of the spectrum, turning our attention to those “poly-topical” passages where – in a Harlequin-like texture of multiple colors – five or more topics were needed to occupy 50% of the available thematic space (Figures 5.1 and 5.2). In lieu of the intensity of turning points and central characters, the rapid criss-crossing of themes produced the light-hearted “surprise” of Meredith's paragraph (Figure 5.1), and opened the door to the inflow of the *Middlemarch* minor characters; or more exactly, squeezing so many characters in such a small space – nine in a hundred words, in Eliot's case – made them all appear “minor”, enlarging by contrast the space for the narrator's ironical comments (Figure 5.2).

These paragraphs offered a neat counterpoint to mono-topical paragraphs, and were also more frequent than them: between 9 and 40% as opposed to 1-4%. But, once again, they were relatively atypical of the corpus as a whole: in Figure 4.4, a single group of paragraphs (“three-topical” ones, in which three topics occupy 50% of semantic space) had a larger presence than the mono- and poly-topical cases

Excerpts of Figures 5.1, 5.2, 5.3, and 5.4; larger images next two pages.

Richard **fac**ed about to make a **querulous** **retort**, **disarmed** him. The **lad's** **unhappy** **nose**, though i becoming **discolored**. To **upbraid** him would be **position**, and **exclaiming**, “Here!” **dropped** down **contemplate** him as a **puzzle** whose every new **m**

– *The Ordeal of Richard Feverel. A History of Father and*

Topic Mallet Assigned  
# Topic Topic  
Score Score

It came very **lightly** indeed. When Dorothea **quit** that Mrs Cadwallader had **stepped** across the **pa** with Celia in a **matronly** way about the **baby**. An **coming** back, it was to be **hoped**, quite **cured** of ] the ‘Pioneer’-- somebody had **prophe**sied that it **colors** for want of **knowing** how to help itself, be Ladislaw, was gone or going. Had Sir James **hea**

– *Middlemarch: A Study of Provincial Life*, by George Eli

Topic Mallet Assigned

But there was such an **influence** in Mi new **school-books** under my **arm**, I **be** my **airy** old **room**, the **grave** **shadow** o to make the **past** more **indistinct**. I **sa** were out of **school** for good at three); yet.

– *David Copperfield*, by Charles Dickens (18

Guy's **instinct** of **flight** was, his **mothe** desire; you shall go.” I had not **expect** **bound** up in him as she was, **accuston** with her, and “**petted**” her more than **reluctance**, some **grieved** **entreaty**-- b up as if her **allowing** him to **quit** her w

– *John Halifax, Gentleman*, by Dinah Maria

Richard **fac**ed about to make a **querulous** **retort**. The **injured** and **hapless** **visage** that **met** his **eye** **disarmed** him. The **lad's** **unhappy** **nose**, though not exactly of the **dreaded** hue, was really becoming **discolored**. To **upbraid** him would be **cruel**. Richard **lifted** his **head**, **surveyed** the **position**, and **exclaiming**, "Here!" **dropped** down on a **withered** **bank**, **leaving** Ripton to **contemplate** him as a **puzzle** whose every new **move** was a **worse** **perplexity**.

— *The Ordeal of Richard Feverel. A History of Father and Son*, by George Meredith (1859)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
19	12.21%	4 / 30 [13.3%]	looked, eyes, voice, turned, tone, face, spoke, felt, heard, speak, smile, answer, speaking, glance, manner, time, eye, silent, silence, began, gave, met, pause, question, stranger, word, round, surprise, hand, expression, strange, turning, room, sat, knew, passed, usual, low, reply, feeling, quick, quickly, smiled, quiet, conversation, slight, countenance, speech, questions character, time, circumstances, life, friends, degree, spirit, grandfather, present, influence, mind, friend, means, feelings, society, considerable, received, position, purpose, intelligence, scarcely, period, interest, deeply, returned, respect, impossible, visit, knowledge, days, family, youth, devoted, conduct, secret, result, met, existence, quitted, confidence, years, difficult, views, future, occasion, success, general, manner, appeared
5	11.76%	4 / 30 [13.3%]	felt, mind, feeling, sense, life, strong, consciousness, mother, presence, husband, brought, painful, feel, dread, effect, state, conscious, future, sort, nature, reason, present, longer, power, feelings, pain, effort, experience, trouble, thoughts, fear, turned, speech, creature, moments, change, wanted, stronger, brother, began, strength, pride, care, pity, turn, ready, imagination, hard, shock
25	9.23%	3 / 30 [10.0%]	n't, dear, sir, suppose, give, hear, mind, things, wo, care, woman, father, call, remember, speak, ah, talk, married, friend, poor, feel, glad, continued, wife, bad, fellow, tone, boy, mine, understand, husband, people, leave, word, live, aunt, girl, stay, trouble, papa, place, bear, sort, head, pretty, exclaimed, kind, matter, fool
22	9.75%	3 / 30 [10.0%]	hand, eyes, face, head, hands, tears, looked, lips, arms, heart, voice, turned, arm, shook, father, held, round, kissed, mother, pale, stooped, pressed, kiss, felt, spoke, fell, laid, cheek, child, cry, touch, bent, gave, neck, suddenly, touched, door, speak, sat, answer, whiskered, trembling, burst, smiled, lifted, whisper, silence, paused, shoulder, suddenly
43	9.22%	3 / 30 [10.0%]	things, people, world, men, find, time, knew, life, feel, women, matter, set, wrong, times, end, sort, deal, hard, bad, person, work, suppose, kind, poor, mind, worse, half, lose, felt, fancy, wanted, making, friends, true, ways, doubt, care, show, story, easy, clever, thoughts, reason, give, fault, stupid, place, began, spite
39	6.3%	2 / 30 [6.7%]	

Figure 5.1

It came very **lightly** indeed. When Dorothea **quitted** Caleb and **turned** to **meet** them, it **appeared** that Mrs Cadwallader had **stepped** across the **park** by the **merest** chance in the **world**, just to **chat** with Celia in a **matronly** way about the **baby**. And so Mr Brooke was **coming** back? **Delightful!** -- **coming** back, it was to be **hoped**, quite **cured** of **Parliamentary fever** and **pioneering**. **Apropos** of the 'Pioneer'-- somebody had **prophesied** that it would soon be like a **dying** **dolphin**, and **turn** all **colors** for want of **knowing** how to help itself, because Mr Brooke's **protégé**, the **brilliant** young Ladislaw, was gone or going. Had Sir James **heard** that?

— *Middlemarch: A Study of Provincial Life*, by George Eliot (1872)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
42	14.7%	5 / 30 [16.7%]	hand, head, looked, chair, round, hands, side, turned, hat, eyes, table, face, sat, forward, arm, sit, seated, standing, room, stood, seat, sitting, turning, window, feet, walk, end, corner, distance, rising, held, began, leaning, close, book, laid, holding, threw, suddenly, time, fingers, walked, open, ground, turn, fire, entered, ring, sofa
35	14.48%	5 / 30 [16.7%]	mother, horn, child, daughter, wc, woman, father, heart, wife, girl, dear, loved, friend, poor, mid, beautiful, sister, happy, looked, world, nurse, years, mc, care, sweet, face, quiet, mamma, talk, husband, grave, creature, papa, scarcely, darling, loving, strange, baby, tender, times, pride, learned, thinking, memory, beloved, true, kiss, tenderness, meet
2	11.67%	4 / 30 [13.3%]	ladies, gentlemen, fine, air, pretty, society, gentleman, party, guests, pleasure, company, dinner, french, conversation, dance, agreeable, gave, women, charming, style, handsome, taste, beautiful, pleasant, perfect, evening, fair, admiration, called, fashion, talked, ball, english, amusement, general, full, high, grand, delightful, beauty, appeared, visitors, distinguished, girl, country, lively, girls, present, world
28	11.86%	4 / 30 [13.3%]	night, bed, light, time, sleep, lay, cold, strange, sat, fell, day, heard, hour, felt, morning, fire, dark, half, knew, evening, sound, silence, silent, heart, hours, room, stood, passed, past, darkness, dream, wind, slept, hear, asleep, dead, grew, lie, warm, rest, remembered, fast, thoughts, face, weary, voice, breath, rain, broke
18	8.92%	3 / 30 [10.0%]	life, heart, world, happy, sweet, happiness, beautiful, day, loved, youth, beauty, years, nature, live, delight, felt, pure, full, dream, soul, real, days, affection, thoughts, kind, secret, eyes, read, friendship, light, calm, future, fair, lived, fancy, sad, hearts, bright, strange, tenderness, earth, peace, fresh, spirit, presence, lives, passion, charm, passed
20	6.69%	2 / 30	day, time, morning, house, evening, hour, place, days, night, week, return, leave, coming, late, hours, usual, stay, brought, half, called, visit, walk, passed, longer, early, afternoon, change, dinner, work, returned, room, knew, order, o'clock, sit,

Figure 5.2

### Figures 5.1-5.2. Poly-topical Paragraphs

The paragraph in 5.1 occurs in the middle of an agitated episode in Richard Feverel's childhood, in which his mind oscillates between several courses of action. The *Middlemarch* passage in 5.2 is part of a larger scene which originates in anonymous gossip, and acquires momentum thanks to the intervention of sir James Chetham and Mrs Cadwallader; it is one of the many moments when Eliot's minor characters occupy the center of the stage.



But there was such an **influence** in Mr. Wickfield's old **house**, that when I **knocked** at it, with my new **school-books** under my **arm**, I **began** to **feel** my **uneasiness** **softening** away. As I went up to my **airy** old **room**, the **grave shadow** of the **staircase** seemed to **fall** upon my **doubts** and **fears**, and to make the **past** more **indistinct**. I **sat** there, **sturdily conning** my **books**, until **dinner time** (we were out of **school** for good at three); and went down, **hopeful** of becoming a **passable sort** of **boy** yet.

— *David Copperfield*, by Charles Dickens (1850)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
8	26.13%	9 / 30 [30.0%]	felt, mind, feeling, life, sense, strong, dread, presence, consciousness, thoughts, feel, change, painful, pain, husband, future, state, nature, feelings, anxiety, effect, fear, things, time, making, vision, experience, conscious, trouble, turn, effort, possibility, present, moments, past, wanted, imagination, silence, begun, full, vague, force, mother, power, sudden, difficult, hard, meeting, longer
11	17.45%	6 / 30 [20.0%]	life, years, time, day, days, ago, year, remember, lived, knew, place, happy, past, things, world, strange, heart, mind, change, passed, months, memory, thoughts, times, weeks, lost, began, dream, end, longer, remembered, changed, present, weary, grown, sad, death, rest, living, forgotten, night, read, daily, live, school, hours, brought, future, lives
28	14.5%	5 / 30 [16.7%]	door, room, house, opened, open, window, table, stood, looked, entered, hall, large, light, small, chair, side, place, stairs, dark, windows, parlor, rooms, floor, staircase, closed, key, corner, furniture, round, walls, shut, glass, fire, hand, led, passage, chamber, locked, end, opposite, appeared, books, doors, wall, standing, seated, library, sitting, gallery
13	9.1%	3 / 30 [10.0%]	hand, head, hands, eyes, arm, round, looked, face, shook, chair, sat, arms, turned, laid, side, shoulder, held, table, stood, putting, fingers, kissed, holding, lips, neck, leaning, forward, gave, shaking, feet, finger, knees, hair, kiss, close, mine, hat, tears, leaned, pressed, fire, bent, corner, lay, cheek, laughed, mouth, foot, touch
9	9.55%	3 / 30 [10.0%]	n't, fellow, suppose, sir, things, mind, give, wo, dear, people, poor, bad, care, pretty, sort, boy, deal, woman, hear, call, talk, head, money, find, fine, bit, afraid, girl, live, speak, remember, time, half, day, wanted, ah, hard, understand, turn, place, world, laugh, talking, fool, feel, wife, word, thinking, glad
48	4.06%	1 / 30 [3.3%]	looked, eyes, face, turned, hand, voice, sat, stood, spoke, round, heard, silence, tone, time, speak, felt, room, smile, door, silent, minutes, speaking, low, glance, word, eye, turning, began, passed, paused, suddenly, head, expression, manner, quick, answer, met, half, slowly, standing, immediately, arm, walked, lips, pale, usual, moved, gave, strange

Figure 5.3

Guy's **instinct** of **flight** was, his **mother felt**, **wisest**, **safest**, best. “My **boy**, you shall have your desire; you shall go.” I had not **expected** it of her-- at least, not so **immediately**. I had thought, **bound** up in him as she was, **accustomed** to his **daily sight**, his **daily fondness**-- for he was more with her, and “**petted**” her more than any other of the **children**-- I had thought to have seen some **reluctance**, some **grieved entreaty**-- but no! Not even when, **gaining** her **consent**, the **boy looked** up as if her **allowing** him to **quit** her was the **greatest kindness** she had ever in his **life bestowed**.

— *John Halifax, Gentleman*, by Dinah Maria Craik (1856)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
43	26.69%	9 / 30 [30.0%]	knew, question, mind, time, give, subject, matter, friend, speak, doubt, word, truth, answer, idea, present, point, felt, case, understand, place, means, feeling, reason, heard, position, friends, duty, fact, husband, offer, find, spoken, making, things, wife, impossible, opinion, circumstances, spoke, aware, purpose, wished, true, simply, called, secret, understood, accept, thinking
21	14.62%	5 / 30 [16.7%]	mother, father, child, horn, poor, children, daughter, girl, looked, baby, wc, mamma, years, boy, dear, woman, wife, darling, husband, sister, knew, world, brother, time, brought, face, beautiful, sweet, nurse, mid, dead, house, heart, happy, boys, papa, pretty, loved, arms, parents, grave, remember, called, save, fond, quiet, day, whispered, talked
20	11.72%	4 / 30 [13.3%]	men, people, women, things, life, world, work, sort, true, children, woman, hard, find, learn, set, wrong, care, poor, end, live, side, small, understand, friends, deal, common, girls, generally, boys, rest, feel, easy, working, easily, place, suppose, living, means, worth, clever, age, pleasant, teach, bad, wise, turn, knowing, give, reason
8	11.93%	4 / 30 [13.3%]	felt, mind, feeling, life, sense, strong, dread, presence, consciousness, thoughts, feel, change, painful, pain, husband, future, state, nature, feelings, anxiety, effect, fear, things, time, making, vision, experience, conscious, trouble, turn, effort, possibility, present, moments, past, wanted, imagination, silence, begun, full, vague, force, mother, power, sudden, difficult, hard, meeting, longer
11	8.93%	3 / 30 [10.0%]	life, years, time, day, days, ago, year, remember, lived, knew, place, happy, past, things, world, strange, heart, mind, change, passed, months, memory, thoughts, times, weeks, lost, began, dream, end, longer, remembered, changed, present, weary, grown, sad, death, rest, living, forgotten, night, read, daily, live, school, hours, brought, future, lives
33	9.06%	3 / 30 [10.0%]	heart, life, woman, loved, happy, world, happiness, sweet, felt, affection, nature, knew, loving, beautiful, soul, passion, feeling, true, tender, care, feel, youth, wife, strong, tenderness, day, full, friend, sympathy, give, beauty, creature, lover,

Figure 5.4

### Figures 5.3-5.4. Mid-topical paragraphs

The two paragraphs describe the immediate aftermath of an important (though not decisive) episode: in *David Copperfield*, David is back from his first day in a new school, and is worried about his future there; while the paragraph from *John Halifax* presents the consequences of a fight between John's two sons, Guy and Edwin.

combined. **Figures 5.3 and 5.4** offer two typical instances of these “middle” paragraphs, which make up over one third (on average, 37%) of our corpus.

Here, finally, we were looking at the “typical” paragraph in our corpus, and at the thematic configuration that characterized it. In the case of *David Copperfield*, one of the three topics appeared to be centered around feelings and expectations (influence, feel, uneasiness, doubts, fears, hopeful); a second one around house and furniture (house, knocked, room, staircase); while the third, more uncertain, had something to do with the passing of time (began, softening, past: not much, but among the 50 top words of the topic the temporal dimension was more visible). *John Halifax* presented a similar tri-partition: a topic clustering around rational decisions (wisest, safest, expected, consent, kindness); one around family relations (mother, boy, petted, boy); and a third, more opaque, related to a general idea of human nature (instinct, accustomed, children; in this case, too, a look at the entire topic strengthened these otherwise tenuous hints). In both novels, the first two topics recalled the polarity of “soul” and “world” in Lukàcs’ *Theory of the Novel*: one was concerned the inner sphere (emotional or rational), and the other with a part of the external context (house, family); while the third topic, less clearly defined, was more open to the narrator’s comments, and narrative framing in general.

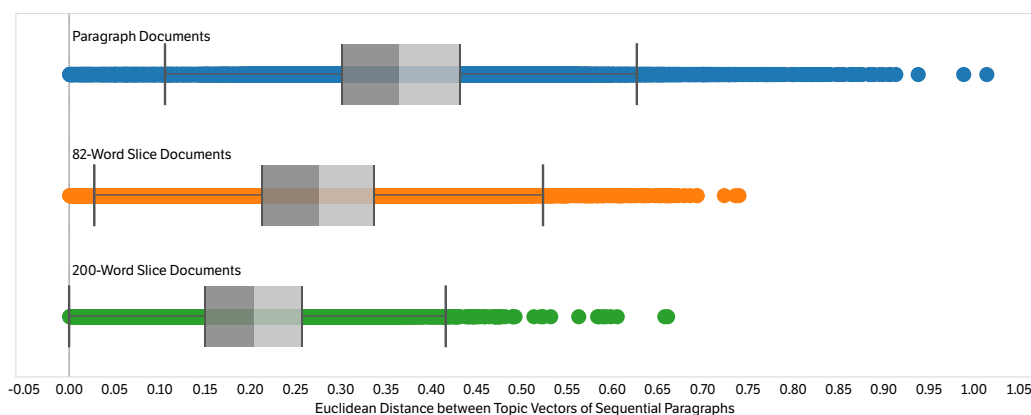
Paragraphs are the textual *habitat* of themes, we wrote earlier; and within this *habitat*, the most typical combination is this mix of three topics (or thereabouts). “Three” is here a strictly empirical result, without any Hegelian grandeur; two topics, or four, wouldn’t change anything (and in fact, given the different ways of measuring thematic focus, a topic’s “score” may easily vary). All that matters is the clear prevalence of these “middle” paragraphs: taken together, 2-, 3-, and 4-topical paragraphs add up to 75% of **Figure 4.4**, suggesting that, most of the time, novelistic thematics will avoid both the intensity of mono-topical paragraphs, and the ironic orchestration of poly-topical ones, settling for a simpler, slightly unremarkable modulation between a small group of elements; a quiet, solid “fit” between the soul, the world – and the narrator.

If the initial comparison between paragraphs and “mechanical” segments had established the greater thematic focus of paragraphs, then, these later findings specify that focus as a kind of *thematic combination*: neither the large notions routinely associated with the idea of the “theme” (War; Nature; Travel), nor those “indivisible units” often labeled as “motifs” (“a bomb explodes”; “falling leaves”; “the train leaves the station”), but *the interaction of a few topics within the frame of an everyday event*. Event; because, let’s not forget it, these are paragraphs in a story; paragraphs that *make* the story. David’s state of mind changes after returning to Mr Wickfield’s house; Guy’s mother accepts the fact that it’s best for him to go away. An action has occurred; the initial situation has been transformed (“I had not expected it of her”) into a different one. And it has been transformed, *by the encounter of distinct topics*: David’s feelings don’t simply change straightforwardly from “fears” to “hopefulness”: they do so, by taking a detour through “house” and “room” and “staircase” (“there was such

influence in Mr. Wickfield’s old house”). The paragraph is not a pawn that makes its orderly one-way move towards the end of the story; it’s a knight that advances by combining two axes in a single move. For now, it’s just a metaphor for how paragraphs contribute to the plot. But more can be expected, in the future, from the encounter of thematics and narratology.

## 6. Lateral explorations

So far, we have been focusing on the internal structure of paragraphs. But what about their “external” borders? Was the typographical break between one paragraph and the next also the sign of a *thematic* discontinuity? Since we had already established which topics were present in any given paragraph (and in the 82- and 200-word segments as well), measuring their discontinuity was relatively simple, and revealed an even greater difference among the three groups than had been the case for thematic focus (Figure 6.1; for thematic focus, see Figure 3.3 above).



**Figure 6.1 Thematic Discontinuity**

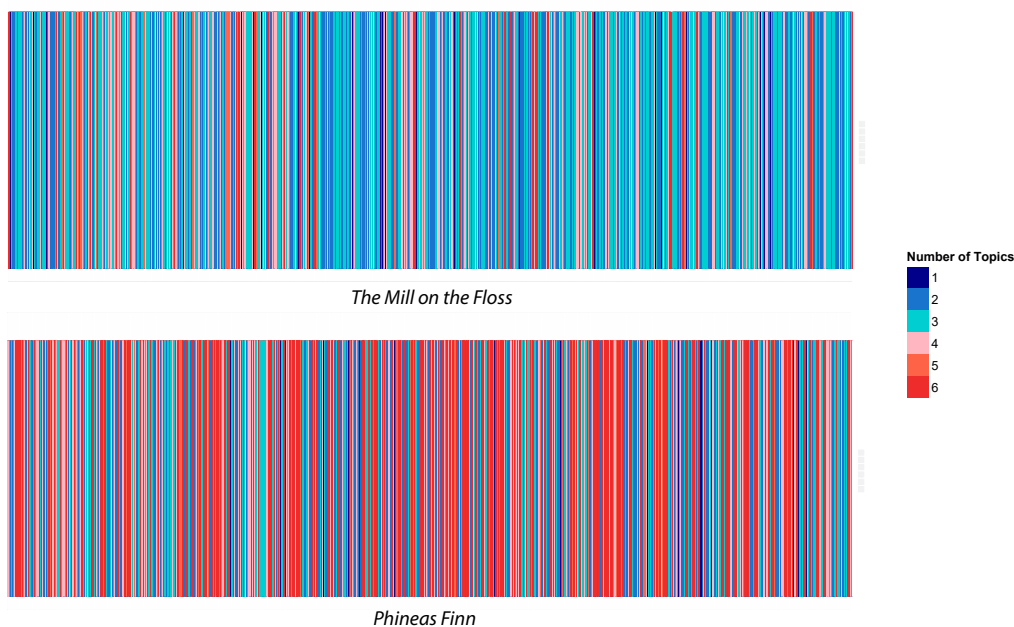
While there are several possible reasons for two paragraphs to be discontinuous with each other, meaningful discontinuities tend to fall into three main categories: shifts in narrative perspective (from one character or setting to a different one), movement from cause to effect (or vice-versa), and a sort of thematic “unfolding” in which a highly focused paragraph is followed by one which introduces associated notions, or places the primary topic in the wider world of the novel.

Though the present pamphlet is an attempt at a quantitative morphology of the paragraph, and not a historical study, that discontinuity *between* paragraphs should be stronger than focus *within* paragraphs is bound to be the result of evolution: of the fact that a “paragraph” originally indicated “a symbol placed in the margin to indicate a noteworthy break in the flow of discourse”, and only later did it “come to signify the stretch of language between breaks” (Rodgers, “A Discourse-centered Rhetoric”, 4). Or in other words: first came the impulse to segment the flow of discourse, and only later did writers strengthen the internal consistency of each segment. First came discontinuity, and then focus.

In this chart, discontinuity is measured by comparing the posterior probabilities for all topics in two sequential paragraphs. By measuring the Euclidean distance between the probability vectors, we are able to compare the similarity of paragraphs based upon which topics they share and with what probability each of them is present in the paragraph.

Since discontinuity is based on pairs of consecutive paragraphs, it was inevitable that, at some point, we would wonder: Why not examine *longer* paragraph series – in principle, as long as the text itself? Did thematic discontinuity follow a hidden rhythm that would allow us to “sequence” the novels we were studying? **Figure 6.2** attempts to do exactly that: each bar represents a paragraph in *The Mill on the Floss* and *Phineas Finn*, color-coded for thematic focus, from the deep blue of mono-topical paragraphs, to the bright red of six or more topics.

Suggestive as these first findings were, no clear pattern emerged from the way paragraphs followed each other; the only constant seemed to be the frequent shift between different types, none of which novels would pause too long on. In a follow-up study, we plan to examine paragraph sequences within chapters, to see if “local” patterns – here dwarfed by the text as a whole – become discernible at that intermediate scale. Meanwhile, our final exploration took us (half) outside of literature, as we decided to compare the thematic organization of our novelistic corpus with two different forms of discourse: a corpus of biographical texts (non-fictional, but with a strong narrative component), and a miscellaneous sample of non-narrative non-fiction, ranging from *Kansas, its Interior and Exterior Life* to *Pessimism: a History and a Criticism*. (Both corpora were drawn from the same period as the *Bildungsroman*, and had the same total number of words.) We wanted to know whether the affinity between topics and paragraphs was specific to novels, or appeared in other forms of writing as well.



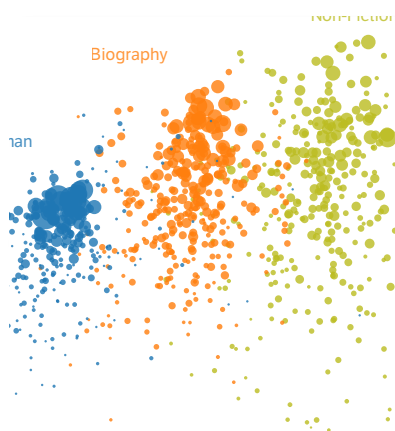
**Figure 6.2. Paragraph sequences in *The Mill on the Floss* and *Phineas Finn***

The color schemes of these images convey an immediate sense of the thematic “temperature” of the two novels, contrasting Eliot’s intense concentration with Trollope’s sardonic tone. The affinity between poly-topical paragraphs and irony in *Phineas Finn* is the same we have briefly discussed in connection with **Figure 5.2** above.

And the results were perfectly clear: both discontinuity and – more dramatically – focus scores were *much* higher in biography, and especially in non-fiction, than in the novelistic corpus. **Figure 6.3**, which combines the two measurements, makes the difference between the three groups impossible to miss.

Why such a dramatic gap? The paragraph in **Figure 6.4**, from the appropriately named *Description of the Skeleton of an Extinct Gigantic Sloth*, is a good instance of the thematic focus that can be achieved by a non-fictional text: all 30 Mallet words in the paragraph, no one excluded, belong to the same “skeletal” topic. In a novel, such single-mindedness is unimaginable: not only are the most frequent paragraphs dominated by two-three distinct topics, but, even in the most intense of mono-topical paragraphs, 10, 12, 15 of the 30 Mallet words fall routinely outside of the dominant topic. It’s not that themes are less relevant, in fiction; rather, they become relevant – they contribute to the advancement of plot – *by associating with other themes*, rather than by fully unfolding in isolation. If their “focus score” is lower than that of non-fiction, then, that is not a flaw, but the very condition of their presence in narrative texts.

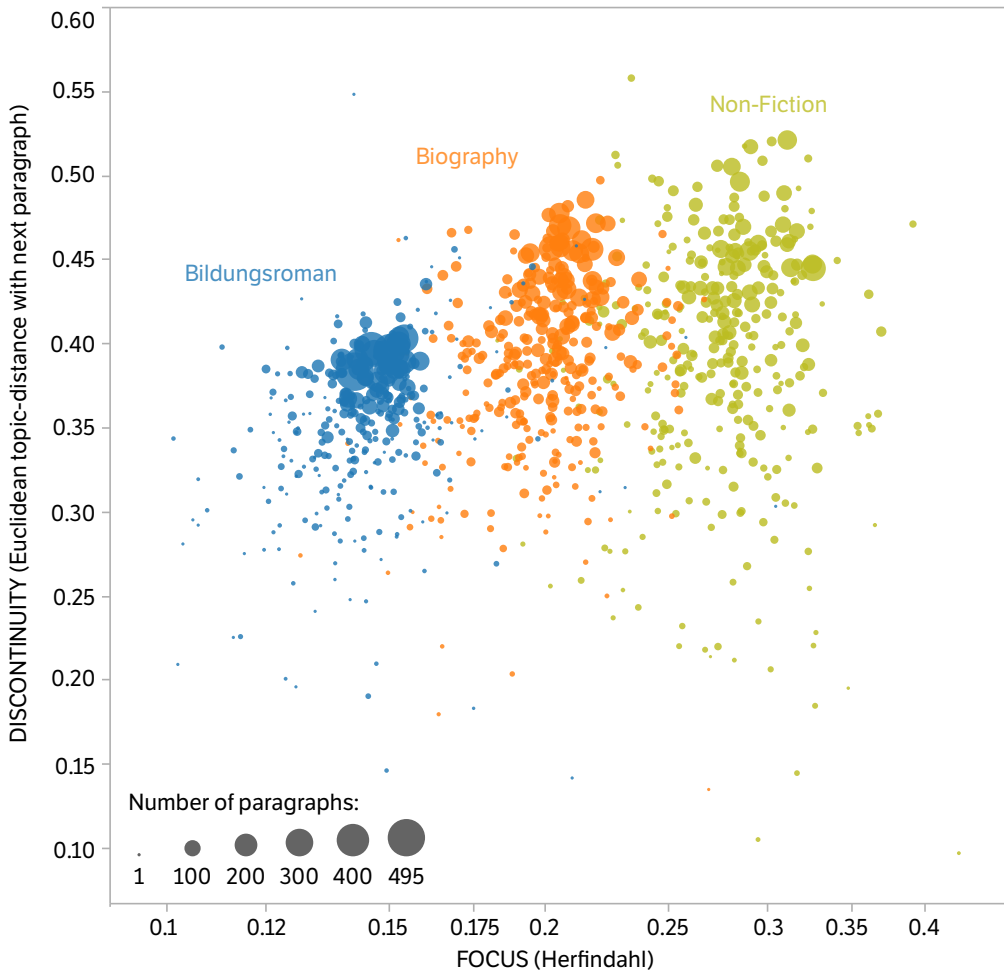
Excerpts of Figures 6.3 and 6.4; larger images next page.



The bones forming the body of the posterior plate of a quadrate form, and the plate; and they each present not sternal bones, and the remaining ei

— *Description of the skeleton of an extinct natural affinities, and probable habits of t*

Topic #	Mallet Topic Score	Assigned Topic Score	
36	92.77%	30 / 30 [100.0%]	surface, bone, anterior, articular, metacarpal, concave, size, transverse, fig, articulation processes, skeleton, plate



**Figure 6.3 Focus and Discontinuity in three Different Registers**

In this chart, the x-axis measures the focus of paragraphs, and the y-axis their discontinuity. The separation between the three discourses – and especially between fiction and non-narrative non-fiction – is unmistakable, and persists even if we topic model each text separately: that is to say, the greater focus of non-fictional texts is not due to the fact that the topics of two books dealing with Kansas and pessimism tend to be more unlike each other than those of two novels (although this is also true); rather, non-fictional paragraphs are both more focused and more discontinuous *even within each individual text*.

The bones forming the body of the sternum may be divided into two parts, a broad and flat posterior plate of a quadrate form, and an anterior rhomb or cube projecting from the middle of the plate; and they each present not fewer than ten articular surfaces, two for the contiguous sternal bones, and the remaining eight for portions of two pairs of sternal ribs.

— *Description of the skeleton of an extinct gigantic sloth, Mylodon robustus, Owen : with observations on the osteology, natural affinities, and probable habits of the megatherioid quadrupeds in general*, by Richard Owen (1842)

Topic #	Mallet Topic Score	Assigned Topic Score	Top 50 Words for Topic
36	92.77%	30 / 30 [100.0%]	surface, bone, anterior, articular, side, posterior, bones, outer, process, form, upper, convex, middle, lower, base, metacarpal, concave, size, ridge, end, vertebrae, distal, slightly, large, rough, species, margin, length, proximal, transverse, fig, articulation, teeth, small, half, broad, presents, phalanx, surfaces, dorsal, skull, vertical, breadth, cavity, processes, skeleton, plate, extinct, diameter

**Figure 6.4 Thematic focus in a non-fictional paragraph**

## 7. Literary Scale

The first pamphlet of the Literary Lab, “Quantitative Formalism”, had been mostly concerned with the frequency of individual words: units which proved to be surprisingly effective at distinguishing literary genres, but not complex enough for an analysis of their internal mechanisms. Two years later, in “Style at the Scale of the Sentence”, the shift in scale from words to sentences allowed us to investigate a literary phenomenon – the emergence of style – which had remained inaccessible at the level of individual words. After two more years, another change of scale – this time, from sentences to paragraphs – has brought to light thematic structures which had not been visible at the level of sentences. And the first glimpses of the scale of the chapter suggest the possibility of a further shift, from themes to the narrative unit of the “episode”. It’s impossible not to venture a general hypothesis: *in literature, different scales activate different structural features.*

Different scales, different features. It’s the main difference between the thesis we have presented here, and the one that has so far dominated the study of the paragraph. By defining it as “a sentence writ large”, or, symmetrically, as “a short discourse”,<sup>18</sup> previous research was implicitly asserting the irrelevance of scale: sentence, paragraph, and discourse were all equally involved in the “development of one topic”. We have found the exact opposite: *scale is directly correlated to the differentiation of textual functions.* By this, we don’t simply mean that the scale of sentences or paragraphs allows us to “see” style or themes more clearly. This is true, but secondary. Paragraphs allows us to “see” themes, *because themes fully “exist” only at the scale of the paragraph.* Ours is not just an epistemological claim, but an ontological one: if style and themes and episodes exist in the form they do, *it’s because writers work at different scales – and do different things according to the level at which they are operating.*<sup>19</sup>

Different scales, different features. But if style “emerges” at the scale of the sentences, and themes at that of the paragraph, this does not mean that they “disappear” at a different scale: since chapters are made of paragraphs, and paragraphs of sentences, the larger unit cannot but retain (some of) the features that had emerged at the lower scale. So, for instance, although we had abandoned the idea of studying style at the scale of the paragraph, we couldn’t help noticing that we *had* encountered style in the course of our research – most clearly in the “irony” evoked in relation to the poly-topical paragraphs of figures 5.2 and 6.2.<sup>20</sup> At the opposite end of the spectrum, mono-topical paragraphs with a strong dominant topic

---

18 Here is a typical passage, from Lewis’ *History of the English Paragraph*, p. 22: “devoted, like the sentence, to the development of one topic, a good paragraph is also, like a good essay, a complete treatment in itself.”

19 In fairness to previous work on the paragraph, the role of scale in multiplying textual functions may be greater in literature than in other types of discourse. The researchers mentioned in footnote 3 have frequently pointed out that Bain’s foundational thinking had a strong “logical” component, and much of their own work has been conducted on a type of text – philosophical, legal, and in general non-fictional – in which a continuity of scale, linking sentences and paragraphs (and perhaps the text in its entirety) is more plausible than in literary discourse.

20 Significantly, this persistence of style was mostly associated with those cases in which *the paragraph’s internal unity was weakest, and the autonomy of sentences correspondingly strongest*; in other words, style was more visible when its elective scale – the sentence – retained a certain independence *vis-à-vis* higher forms of integration.

(figures 4.1-4.3) had a narrative force that could easily “carry over” into the higher scale of the chapter, influencing in depth its narrative composition.<sup>21</sup>

Different scales, different features; and with each new scale, a significant increase in textual complexity. This is not because “higher” scales are *in themselves* more complex: themes are not “more complex” than style, nor episodes more complex than either; they are just completely different – incommensurable, really – in the task they perform. Complexity does not reside “in” any feature or scale of the text: it arises from the fact that multiple scales are embedded into each other, and interact in multiple ways. In the paragraph, we have uncovered the specific “play” of a mid-level structure, which both builds upon smaller components, and acts as a building-block of a much larger object. In this looking both “below” and “above” itself, paragraphs enjoy a uniquely central position in the economy of texts.

---

<sup>21</sup> The narrative potential of paragraphs is clear in Watt’s essay on *The Ambassadors*, with its frequent evocation of James’ “progress towards the foreordained illumination”, “progressive and yet artfully delayed clarification”, and so on (Albert E. Stone, Jr., ed., *Twentieth-Century Interpretations of “The Ambassadors”*, Prentice-Hall, N.J. 1969). By contrast, narrativity is silently downplayed in Auerbach’s pages on *Madame Bovary*: “nothing particular happens in the scene, nothing particular has happened just before it. It is a random moment from the regularly recurring hours...” (*Mimesis*, cit, p. 488).