Distant reading in Russia: Franco Moretti and the formalist tradition

Jessica Merrill
2 May 2018

Franco Moretti’s seminal collection 'Distant Reading' set out his famous quantitative approach to literary criticism and was a key contribution to the emergent field of digital humanities. Moretti’s interest in the ‘big questions’ of literary evolution, literary form and narrative universals was shared and significantly influenced by early twentieth-century Russian Formalism, writes Jessica Merrill.

The essays collected in Franco Moretti’s *Distant Reading* were originally published between 1994 and 2011. At the time, they were key contributions in the emergence of the new field of digital humanities. Today, the digital humanities encompass an increasingly diverse, sophisticated field of scholarship. The translation of Moretti’s seminal collection into Russian [1] provides an occasion to historicize Moretti’s approach and above all to place it in the context of the formalist tradition.

*Distant Reading* is really an intellectual biography. It does not describe how algorithms can be used to study literature; rather, it is a prelude to the more technical stage of Moretti’s work that began with the formation of the Literary Lab at Stanford in 2010 and the statement of ‘Quantitative Formalism’ as the Lab’s working method. *Distant Reading* is thus not a case book for the confirmed digital humanist, or a justification of a method. Instead, it seeks to articulate the motivations which prompted Moretti to take on the challenges of a quantitative formalism.

Moretti’s commentary reveals that his early motivations can be traced to Italian Marxism, particularly the work of Galvano Della Volpe and Lucio Colletti. In the 1970s, these scholars critiqued Gramscian neo-Marxism, advocating a more strictly materialist, empirical and anti-dialectical approach. [2] In his own work, Moretti writes that his preference for an empiricist and materialist approach to literature was first articulated (in his 1983 introduction to *Signs Taken for Wonders*) as a desire for a ‘falsifiable criticism’. [3] It is not just that Moretti wants his arguments to rest solidly on facts, but that he wants to promote a more substantial collaboration in literary studies – one driven by the confirmation and refutation of findings.
In this spirit, *Distant Reading* is structured as a dialog between Moretti and his critics. Moretti repeatedly cites what he takes to be the weightiest objections to his work. He defends his position but also concedes that his critics are ‘right’ and that he ‘should have known better’. [4] As Moretti writes, ‘once you have really been proved wrong, the argument is no longer about you; it’s about a world of facts that everybody agrees to share’. [5]

**Conceptual models: From linguistics to socio-economics**

Moretti’s big questions are reminiscent of those that inspired the first generation of literary theory in Russia at the turn of the 20th century. *Distant Reading* is concerned with world literature, with the ‘laws of literary evolution’, with literary form, with the device as a minimal unit, with narrative universals. It is interesting to compare Moretti’s approach to these questions with the foundational contributions Alexander Veselovsky’s *Historical Poetics*, Russian Formalism, and Czech Structuralism. Although they are not living collaborators, and their work may not – strictly speaking – disprove or confirm Moretti’s assertions, they are speaking the same language.

On a relatively superficial level, they share a certain formal orientation. Moretti’s method begins with defining the basic, quantifiable ‘units of analysis.’ These are often defined in linguistic terms. For example, in the essay ‘Style Inc.: Reflections on 7,000 Titles’, Moretti’s argument hinges on the significance of linguistic categories (adjectives, proper names, nominal sentences, definite versus indefinite articles, space nouns). He defends this approach, writing that ‘for me, formal analysis is the great accomplishment of literary study, and is therefore also what any new approach – quantitative, digital, evolutionary, whatever – must prove itself against’. [6]

In ‘The Slaughterhouse of Literature,’ a study of canon formation on the basis of detective stories, his unit of analysis is the ‘clue’. His hypothesis, which is inspired by Walter Benjamin, is that ‘formal choices’ made by authors can be understood as an effort to ‘eradicate’ … competitors in the literary marketplace. Formal devices such as ‘clues’ are features that allow one work to be differentiated from another, allowing it to be selected by readers and thus (potentially) to end up in the canon rather than the archive (the 99.5% of literature that becomes the ‘great unread’). [7] Moretti’s impulse to begin with formal analysis, with the identification of literary ‘devices’, would have been met with approval by the Russian Formalists. They also understood the literary ‘device’ as a means of differentiating one text from another; e.g. the defamiliarized from the automatized. What is more, the aim of Moretti’s study, ‘to explain the laws of literary history’, was also for them a central goal. [8]

However, even as Moreetti takes on these familiar topics, his approach differs. He is not thinking through language as a model for understanding literature, as these earlier theorists were. His models for literary evolution come not from language, but from economics and biology – specifically Immanuel Wallerstein’s world-systems theory and Darwinian evolutionary biology. This is Moretti’s innovation in tackling a long-standing problem for literary studies: how to connect formal analysis and literary history.

Moretti approaches literary history through the impact of consumer choices on the literary market. In ‘Slaughterhouse’, he describes canon formation as such: ‘readers ...
buy A, inducing its publishers to keep it in print until another generation shows up, and so on'. [9] The ‘market amplifies that initial selection over and over again’, starting a ‘feedback loop of “increasing returns” until Doyle ends up occupying 99.9 per cent of the market for nineteenth century detective fiction.’ [10] There is thus a difficult conceptual jump made: from formal analysis (the presence or absence of ‘clues’) to the behaviour of the literary market (consumer choice and increasing returns). In other words, how do we get from the author’s use of a device to the reader’s decision to buy a book?

The problems here are numerous, from basic causality – the reader ostensibly needs to buy the book before he/she can register the presence/absence of a device – to the complexities of the psychology of perception. In ‘The End of the Beginning’, Moretti admits that he does not have a ‘demonstrable causal relation’ between formal traits and the ‘long term survival of a literary genre’. [11] Moving between the spheres of formal analysis and the socioeconomics of the literary market (between formalism and literary history) is an intriguing research goal, but not an easy one.

Complex Presentiment: Half-Figure in a Yellow Shirt, 1932 by Kazimir Malevich. Source: Wiki Commons

**Literary evolution: centres and peripheries**

How did earlier theorists address this problem? The Russian Formalists treated literature as a ‘verbal mass’ or discursive field. Although they adopted a more sociological approach to literature in the late 1920s, their initial theoretical model for literary evolution came from the study of spoken language, specifically dialectology. This was most obvious in Roman Jakobson’s writings. Throughout his career, Jakobson would claim that ‘there is a close correspondence, much closer than critics believe, between the question of linguistic phenomena expanding in space and time and the spatial and temporal spread of literary models.’ [12] In this model, literary innovations (forms) entered circulation in the same way as linguistic ones. Both the Formalists and Moretti were and are concerned with the diffusion of formal innovations, and specifically with the role of ‘centres’ and ‘peripheries’.

Moretti turns to world-systems theory to explain how places such as Paris or Hollywood end up dominating the world market for literature and film production in the 19th and 20th centuries. He is interested in explaining how forms move, and why it appears that a select few gain global dominance. In his essays ‘Conjectures on World Literature’ and ‘More Conjectures’, Moretti discusses the movement of literary forms from the centre to the periphery, where they ‘exert pressure’ on local literatures. [13] This idea has prompted critique from those who would grant the periphery more agency. Moretti sees this as wishful thinking: ‘the culture industry as a “guest” invited by a “host” who “appropriates” its forms... are these concepts – or daydreams?’ [14]

The idea of selective appropriation by the periphery has, however, been central to Russian philology. When speaking of the transfer of elements of folklore between cultures, Alexander Veselovsky stressed that the receiving sphere was not to be understood as an empty space but as one of ‘counter flows’. This idea was developed in the 1920s and 1930s by the structuralist ethnographer Petr Bogatyrev, whose
ethnographic geography sought to explain why items of urban culture would be adopted by some regions but leave others ‘islands’ untouched. To do this, he studied the ‘structure’ of the receiving cultures.

On the question of the agency of the periphery, then, the use of economic and linguistic models leads in different directions. Moretti comes to the problematic conclusion that the ‘(literary) periphery is not necessary for the existence of the (literary) core’ (108). He concedes that it is not reasonable to think of formal innovation at the centre as occurring ‘autonomously’, but remains committed to an economically inspired understanding of world literature as a system of ‘inequality’ in which a ‘strong, systematic constraint’ is exerted by some literatures over the others. [16] While Moretti’s model is largely unidirectional, the Russian Formalists and Czech Structuralists tended to speak about the centre and periphery as a cyclical system. In his Aesthetic Function, Norm, and Value as Social Facts (1936), Jan Mukařovský wrote that ‘it frequently happens that a canon which has sunk to the lowest periphery is suddenly elevated to the very centre of aesthetic activity and becomes – in an altered form, of course – once again a new and vital norm.’ [17] In this earlier approach, scholars likewise relied on the concept of the ‘centre’ to disseminate innovations, but saw that creative transformation of norms occurs on the periphery.

This linguistically inspired approach is consistent with the current sociolinguistic position on language change which, for example, ‘universally points to the working class and lower middle class as the originators of sound change in contemporary American English’. [18] It is, however, in contrast to work in other fields, such as that of Everett Rogers, which finds that the diffusion of technical advances comes from the top. [19] Moretti’s turn from the linguistic model raises important questions: does a work of literature travel like a commodity, evolve like an organism, spread like a linguistic innovation, or disseminate itself like a technical advance? Can we even generalize about ‘literature’ as a category or will different genres, for example, need different models?

**World literature: formal theories of ‘plot’**

Another point of comparison between Moretti’s Distant Reading and earlier Russian work on world literature is the treatment of plot as a narrative stratum that is formally static and geographically mobile. The study of traveling plots goes back at least to the influential work of the German philologist Theodore Benfey, who sought to trace a common stock of tales found in Europe to their Indian origins, developing what became known in folklore studies as the ‘theory of borrowing’. Veselovsky built on Benfey’s work in his own studies of the poetics of plots, which in turn inspired Vladimir Propp, Mikhail Bakhtin, Olga Freidenberg and the Russian Formalists, among others.

A premise Moretti shares with these Russian theorists is that to study the history of plots, one must posit that literary narrative can be thought of as being stratified. Moretti, with his gift for simplification, puts it as follows: ‘foreign plot; local characters; and then, local narrative voice’ [20]; or even more simply: ‘the novel is a composite form, made of the two distinct layers of … plot and style’. [21] This allows for the non-synchronicity of literary history; for older cultural phenomena to persist or reappear alongside more recent developments. In ‘The Novel: History and Theory’, Moretti concludes that the pre-modern adventure plot was taken up by the modern novel because it allowed for the ‘the
functional articulation of ideological needs’. [22] This is a formulation that could have come from Veselovsky’s Historical Poetics. As the stratum of the text that persists over time formally and moves geographically, it would seem that plot is a key unit of analysis for the study of world literature. But how are we to approach it in quantitative terms?

Moretti’s last chapter in *Distant Reading*, ‘Network Theory, Plot Analysis’ proposes that plot can be quantified using network theory. In his example, he creates a network for Shakespeare’s *Hamlet* by tracking who speaks to whom (‘two characters are linked if some words have passed between them’). [23] This is a departure from the Russian tradition, which approaches plot in terms of the actions of the dramatis personae. Following Veselovsky, Propp approached the morphology of plot by focusing on the actions (verbs) in a story: who *does* what to whom, and at what point in the chain of events? Although Moretti doesn’t take this approach, there is an abundance of Proppian work in quantitative narrative studies. Mark Finlayson, for example, has created an algorithm designed to learn Vladimir Propp’s theory of narrative structure from an annotated corpus of Russian folktales. [24] In *Distant Reading*, however, it is Viktor Shklovsky’s thinking on plot that is consistently evoked.

Moretti repeatedly cites Shklovsky, calling him ‘the critical genius of the age’. [25] When discussing narrative, Moretti refers to Shklovsky’s central concepts, but Moretti divides and separates concepts that for Shklovsky are conjoined in a productive tension. These conceptual pairs include: curiosity vs. suspense, forward momentum vs. complication, plot vs. style, and ‘high’ vs. ‘low’ genres. In his *Theory of Prose* and other writings from the 1920s, Shklovsky argued that narrative plots are generally formed from a combination of two principles: ‘ring structure’, which provides a beginning and ending; and ‘stepped construction’, that decelerates the progress from beginning to end. The first principle, one could say, evokes curiosity, while the second creates suspense by slowing down the progress towards the promised ending. Shklovsky derives both principles from the simplest of ‘low’ genres – riddles and fairytales – and proceeds to show how, together, they continue to inform even much more complex, ‘high’ forms of literature. All of these concepts feature in Moretti’s writing on prose structure, but Shklovsky’s synthetic theory is broken down into disparate kinds of narrative.

Following Tzvetan Todorov, Moretti separates curiosity and suspense into ‘two entirely different forms’ of narrative, [26] positing that ‘the acceleration of narrativity’ and ‘the construction of complexity ... are totally at odds with each other’. [27] He suggests that the first – forward-driving, easy prose – is typical of ‘dime novels’, while complex ‘high’ prose, such as Henry James’s *The Ambassadors*, is to be understood as an exception to the rule. This opposition is pursued in Moretti’s comparison of the European and Chinese novel form – with the latter marked by extreme complexity, described as vertically oriented ‘parallel prose’, in contrast to the horizontality and relative simplicity of the European novel. [28] It is worth noting that Shklovsky’s argument in *Theory of Prose* is that all narrative, from the folktale to Tolstoy’s *War and Peace*, can be understood in terms of his two principles. For Shklovsky, all narrative relies on tension between forward movement and complicating structures.

Nevertheless, it is remarkable how much Shklovsky and Moretti have in common. Like Moretti, Shklovsky’s approach is reductionist: his *Theory of Prose* begins with the riddle as his simplest example of narrative and then builds up to fairytales, novellas, collections
of stories (e.g. *The Arabian Nights*), concluding with the novel and modernism. Shklovsky’s goal was also a theory of world literature: he wanted to identify universal principles of plot construction. So why does Moretti break apart Shklovsky’s more synthetic theory? My sense is that this reflects a need to identify seemingly self-sufficient units of analysis for quantification. If Shklovsky were to respond to *Distant Reading*, he might ask: Can we talk about these tendencies as separate categories (e.g. forward propulsion and deceleration), or shouldn’t keep them together? How will we make sure that we haven’t over-dissected the object of study – broken apart its minimal units?

**Beyond parochialisms**

In outlining some points of comparison between this (largely) Russian tradition of literary theory and Moretti’s book, I have argued that they appear to have a common language. I mean this primarily in terms of shared concepts and research goals. But it is also true on the level of discourse. Moretti and Shklovsky both cultivate an irreverent, jocular tone that speaks of their position as iconoclasts seeking to shake up their field. They even make the same joke about the Roman goddess of childbirth. [29] More generally, both pose big questions in an effort to counter a dominant tendency towards particularizing, contextualizing approaches. This ethos spurred the Russian Formalists in their highly influential, highly collaborative work – and it would seem to be an essential ingredient for the success of quantitative formalism in the 21st century.

The rich legacy of Russian literary theory is, however, only one argument for translating *Distant Reading* into Russian. Some of Moretti’s critics work have focused on the Anglocentric aspect of his project. [30] *Distant Reading* makes it clear that, for Moretti, the motivation for digital humanities is the promise of the idea of world literature. For this, it is essential to have collaborators from outside the academic ‘centres’ of western Europe and the United States.

*This research was supported by the Russian Science Foundation, project number 16-18-10250.*

**Footnotes**


2. Franco Moretti, *Distant Reading*, 2013, 180, 155.

3. Ibid. 64.

4. Ibid. 108.

5. Ibid.

6. Ibid. 204.
7. Ibid. 70–71; 66-67.

8. Ibid. 77.

9. Ibid. 68.

10. Ibid. 146, 69, 70.

11. Ibid. 142–144.


14. Ibid.


20. Franco Moretti, *Distant Reading*, 2013, 57; italics in original.

21. Ibid. 132; italics in original.

22. Ibid. 178.

23. Ibid. 214.

24. Mark Finlayson, ‘*Inferring Propp’s Functions from Semantically Annotated Text*’, *Journal of American Folklore*, 2016. 129 (511), 53–75.


26. Ibid. 83.

27. Ibid. 163.

28. Ibid. 170.


**Published 2 May 2018**

Original in **English**
Downloaded from eurozine.com (https://www.eurozine.com/distant-reading-russia-franco-moretti-formalist-tradition/)
© Jessica Merrill / New Literary Observer / Eurozine