Mihnea BÂLICI

Faculty of Letters, Babeş-Bolyai University

Cluj-Napoca, Romania

mihnea.balici@yahoo.com

THE EMERGENCE OF QUANTITATIVE STUDIES. ACTUAL FUNCTIONALITIES AND THE ROMANIAN CASE

Abstract: The aim of this essay is to examine the evolution of quantitative studies and the changes that occurred within their usage, portraying the international case as well as debating upon the existence and the relevance of a Romanian case. Given the fact that quantitative studies have become increasingly visible in literary studies during the last decades, due to various influences such as Franco Moretti's theoretical branch or Matthew L. Jockers' macroanalysis and its appropriation in digital humanities, a pertinent study on this methodology is necessary because: 1) it can contribute to a better understanding of the role it has had in this domain and 2) it can bring better insight into its possible use in our country. For these reasons, this essay is seeking to present a history of the quantitative methodological approach as applied on macro scales (therefore, the main domains discussed in this essay are the sociology of literature and digital humanities) and to describe the ways it has influenced or may further influence the study of literature in Romania.

Keywords: quantitative studies, macroanalysis, distant reading, statistical analysis, sociology of literature

Since there has been an important evolution in the general use of quantitative studies in literary research during the last decades, often related to Franco Moretti's theories and methodologies, it is necessary to examine the ways this research instrument has been

used in the past in order to comprehend its actual functionalities. Also, it is equally important to evaluate whether these methodologies have had any impact in other countries outside their countries of origin. Therefore, this paper has two distinct aims. On the one hand, it proposes an analysis of the development of quantitative studies in the literary domain. What it aims at proving is that there has been a certain continuity in the practice of quantitative research of literature, as well as an epistemological break due to the latest technological advances that have logistically facilitated the process of data mining. The next questions are put forward: Has there been a significant broadening of the objectives of research in the *digital humanities* paradigm, in comparison to the first historical uses of quantitative studies? How exactly have the aims of quantitative research broadened? What are the new problems and challenges raised by quantitative studies in humanities and, more precisely, in the study of literary phenomena? On the other hand, this study is trying to track down possible uses of quantitative studies in Romanian criticism. This case is important because the autochthonous critical phenomenon of the last century is characterised by a strong impressionist influence. Since feuilletonism has been the major medium through which literature has been analysed, the aesthetic value of a literary work was the general focus of most of the critics. The main direction in the study of literature has been defined by the axiom of "the autonomy of the aesthetic", so most of the associations of the literary phenomenon with the socio-political domain have been reluctantly viewed. Literary criticism became a strictly axiological undertaking. Therefore, the existence of a positivist tradition can be scarcely traced. This fact might have been caused by the socio-political context of the last century: the dogmatic appropriation of Marxism in the humanities during the communist period might have caused the rise of a hostile attitude towards any so-called "reductionist" approach to literature in later times. However, there have been attempts at quantifying literary phenomena in the Romanian space. In the sociology of literature as well as in the structuralist paradigm, the quantitative method has been, if not practically applied (due to logistical reasons), at least theorized. This essay will also discuss the sociological approach of quantitative studies in the Romanian discourse on literature, given the preoccupation of this domain with large literary corpuses and historical generalisations.

Sociology of literature/Literary sociology

Before any analysis of the actual functionality of quantitative studies in the literary domain, a definition of the involved terms is necessary. What is exactly quantitative research? What are its main objectives? Broadly, it is defined by the quantification of observable and empirical phenomena, in order to systemise the information by using statistical measures and mathematical procedures. Quantitative research is associated with natural sciences, and its appropriation in human sciences took place via sociology, which "has oscillated between a scientific orientation which has led it to ape the natural sciences and a hermeneutic attitude which has shifted the discipline towards the realm of literature" (Lepenies 1), ever since its foundation as a science by Auguste Comte. Commenting upon the conflict and dialogue between sociologists and the literary community during the 19th and 20th century, Wolf Lepenies proves that there had indeed been a relationship of mutual influencing between the two domains. It can be clearly seen that the first scientific approach to human sciences had taken place by the means of sociology. In the 19th century, Hippolyte Taine had been the first to embrace a sociological view of the literary object, viewing it as a product of the socioeconomic environment and not of individual, autonomic authors, following Mme de Staël's thesis that there is a close relationship between art and society. This has been a starting point for the empirical study of literature: Taine's radical point of view about the influence of the social environment in the development of the literary domain led to a new perspective of this cultural phenomenon. Far from being a metaphysical, ineffable and exclusively subjective experience, literature becomes a social product and is perceived as an institution, a commodity and a collective enterprise.

However, trying to define or set methodological limits to the sociology of literature as an independent and homogenous branch in literary studies becomes increasingly difficult, the reason being that "there is virtually no established corpus of knowledge called the sociology of literature" (Swingewood 12). Still, Alan Swingewood succeeds in underlining two principal directions in the sociological study of literature at the end of the 19th century and the beginning of the 20th century. Firstly, one that "became known as positivism", that was searching for "material, scientifically ascertainable connections" between literature and "'facts' such as climate, geography, and race". Secondly, a direction with a "strong critical conception", defined by the

"critical tendency through Marxists such as Lukács and Goldmann" (Swingewood 30-31), therefore by the explanation of the individual text in relationship to the social context and the tracing of ideological factors that appear in the literary work. What Swingewood points out is the methodological problem of both of these directions: the first one, represented by Robert Escarpit and the study of the social factor inside the literary institution, is prone to reductionism, due to its constant reference to the external aspects of literary phenomena, while the second one, represented by genetic structuralism, risks falling into speculative debates about the importance of the social factor in the constitution of the literary object without basing these assumptions on any empirical data. However, it is important to mention that the distinction Swingewood makes is to prove his affinity to the second direction of the sociology of literature, explaining that the first one is way too "dogmatic".

However, the basic distinction would be that between a sociology of the institutionalisation of literature and an analysis of the sociologic factors within a literary work. The second direction (the Goldmannian one) is closer to the structuralist tradition, while the first one is mostly positivist and empirically based. During the 20th century, quantitative studies have been used especially in this first branch of the sociology of literature. Robert Escarpit is one of the most significant figures in this field. His perspective on literature has its origins in the Marxist perspective on art: literature as a merchandise, as a product, as a marketing enterprise, economically and socially determined, thus explainable and quantifiable. He proposes a threefold affiliation of literature: "to the world of individual spirits, of abstract forms and of collective structures"[1] (Escarpit 4), insisting on the collective status of the literary act while criticising the lack of a "veritable sociological perspective". Given the fact that until Escarpit the main methodologies used in the study of literature have been either the documentation on the author's "spiritual" biography, either the sole interpretation of the literary texts' meaning (what we would call nowadays "close reading"), his approach to literature had been pioneering in the French tradition of literary criticism. The inventorying and systemising of book production, based on the analysis of the social structures (the authors' and readers' social classes, their jobs, their economic status, the

---

[1] "[...] la lumea spiritelor individuale, a formelor abstracte și a structurilor colective" (My translation).

grade of illiteracy in France, the capital-periphery relationship, the role of the editor and of the librarian in the distribution of books), is vital for the understanding of the way in which literature works and is constituted. "Statistical data emphasize the basic features of the literary act"[2] (Escarpit 24), and quantitative research has the role of analysing the external factors of the French book market and the social position of its agents: the author, the editor and especially the reader. Book reception is a central point in Escarpit's theory. By studying the literary public and the general opinion of the popular masses, the evolutionary and dynamical aspect of literature is highlighted, demonstrating that there is a close connection between socioeconomic factors and the literary strategies employed by the authors. This comes from Escarpit's critical view on the selective, cannon-based perspective of literary criticism (which is very similar to the Morettian perspective): "Just as scholar maps show us in an erroneous way a giant Alaska smashing a very small Mexico, twelve or fifteen years of Versailles are smashing, in the 17th century, sixty years of French literary life"[3] (Escarpit 5). Escarpit also insists on the importance of so-called "infraliterary" (or "paraliterary") phenomena, which were regarded as irrelevant in the past but had had an important function in the evolutionary process of literature: "the belonging of a literary piece to literature or subliterature is not defined by abstract qualities of the author, of the piece or of the public, but by a certain type of exchange"[4] (Escarpit 22). By the use of maps and statistics, Escarpit analyses the dynamics of generational shifts, the relevance of geographical settings, the circulation of books on the market, pricing issues, and the problem of the reading public in France.

The empirical method is strongly based on the discovery of existing patterns and on interpreting the results, while forming a coherent system.[5] This methodology serves as the basis of any quantitative approach to the literary field. However, Robert Escarpit has only launched partially-verified hypothesises about the influence of socio-political factors on the evolution of the literary structures. Not having had the technology to

---

[2] "Datele statistice permit reliefarea trăsăturilor de bază ale faptului literar" (My translation).

[3] "Aşa cum hărţile şcolare ne arată în mod fals o enormă Alaskă strivind un Mexic foarte mic, tot aşa doisprezece sau cincisprezece ani de Versailles strivesc, în secolul al XVII-lea, şaizeci de ani de viaţă literară franceză" (My translation).

[4] "Apartenenţa unei opere literare la literatură sau la subliteratură nu este definită de calităţile abstracte ale autorului, operei sau publicului, ci de un anumit tip de schimb" (My translation).

[5] See the essay "De la pattern la structură" (*From Pattern to Structure*) from Escarpit's *De la Sociologia literaturii la Teoria comunicării* (385-397) for a better analysis of the concept of "pattern" and its importance for the methodological approach to quantitative data.

exhaustively study the literary forms in order to grasp their possible historical patterns, he could not make proper comparisons between the social context and the inherent, thematic/stylistic elements of a literary piece.

Distant reading. Macroanalysis

After 1990, the figure of Franco Moretti is pivotal to the history of quantitative studies in literary studies. In *Atlas of the European Novel, 1800-1900*, he argues that literary forms evolve and change due to their circulation from a specific area to another. Literary forms are altered by their movement through different socio-cultural contexts in the everchanging world-system. In *Graphs, maps, trees*, his approach to the literary act is based on the idea that "a field this large cannot be understood by stitching together separate bits of knowledge about individual cases" (Moretti, Graphs 4), since "it's a collective system, that should be grasped as such, as a whole" (Moretti, Graphs 4). Therefore, a quantitative approach to literature is required, either by way of computational stylistics, indexing of databases or book history (in *Graphs...*, for example, his focus is mostly on the last one). He also encourages the use of inter-universitary cooperation as a core-element of quantitative studies, given the large corpus of data this kind of research requires. Therefore, literature is visualised as being distributed in historical cycles that work and develop according to observable patterns, which should be identified and studied. In order to do this, he makes use of graphs, maps and dendrograms, so as to comprehend the whole structure of a certain literary cycle, tracing its internal evolution so as to understand the reasons behind this process. For example, Moretti analyses the stages in the life of the British novel, pointing out to societal and political events that marked its rises and falls. However, he is aware that this hypothesis is outside the spectrum of quantitative research, since "it provides *data*, not interpretation" (Moretti 9), and the latter requires a completely different kind of research (sociological, stylistic etc.). While not explaining qualitative aspects of literature, quantitative studies are more faithful to the purely empirical facts. What Franco Moretti provides is the perspective and methodology. But the basis of the Morettian perspective is enabled and enlarged by the new technological developments, which brought significant nuances to this methodological approach, since "a desire for a

general theory of the new literary archive is slowly emerging in the world of digital humanities" (Moretti, Distant Reading 122).

Digital humanities have considerably facilitated quantitative research in literature by means of rapid mining of enormous quantities of data. Basically, the logistic disaster that threatened quantitative studies in humanities was surpassed. A short history of the context that prefigured the *digital humanities*' phenomenon shows that in the beginning the computer's applicability was limited to the strictly textual level, especially to the problem of authorial attribution. In 1851, Augustus de Morgan proposed a quantitative study of the vocabulary from the *Pauline Epistles* in order to identify the author. At the end of the 19th century, T.C. Mendenhall was the first to launch the idea that the style of an author can be quantified by measuring the length of the words and the frequency of different phrases in a given text.[6] This historical turning point is officially the basis of stylometry as we know it today. The so-called "computing humanists" had as a main objective exactly this: measuring and identifying the style of a particular author by studying factual aspects of the text. After the Second World War, the statistic research of style became increasingly viable and efficient, since the technical instruments of quantitative research were used by academicians exclusively for this. Another landmark in digital humanities was Roberto Busa's *Index Thomisticus*. In the '50s, Busa digitally indexed all the writings of Toma d'Aquino, enabling access to disparate parts of the philosopher's work. This could be seen as the beginning of metadata and digital indexes.

Despite technological progress and constant streamlining, the instrumentalization of the computer in literary research encountered reluctance from literary scholars. The positivist attitude towards the textual phenomena and the use of statistics in literary interpretation were the main reasons why the community of *computing humanists* were marginal in the last part of the 20th century in the United States. This is understandable, given the fact that these years were strongly marked by cultural studies, continental philosophy and close reading. Thomas Rommel observes that the primary contribution of digital humanities was in stylistics, therefore it was

---

[6] See Anthony Kenny's first chapter of *The Computation of Style. An Introduction to Statistics for Students of Literature and Humanities*, "The Statistical Study of Literary Style" (1-14) for a concise historical panorama of the use of statistics in humanities.

used only as a strictly intra-textual tool. At this point, digital humanities did not move beyond simple "stylo-statistics" (Rommel 94)[7]. Digital tools had a limited function, and we cannot compare these approaches to a flourishing computational quantitative paradigm.

At this point, the contribution of Franco Moretti and Matthew L. Jockers on digital humanities is relevant. "The existing tools have kept our focus firmly on the close reading of individual texts and have undoubtedly prevented some scholars from wandering into the realms of what Franco Moretti has termed 'distant reading'" (Jockers 17). The use of instruments that are specific to digital humanities and applying them to the literary macrostructures proposed by Moretti represents a significative step-forward in the history of quantitative studies, enabling the investigation of vaster literary zones. The digitalisation of an exhaustive corpus of literary works, as well as applying stylometry's techniques to more general literary phenomena might answer to questions such as: What are each literary genre's characteristic patterns? Is style determined by nation? Have sexuality, ethnicity or geography any stylistic influence? How are the literary forms related to the concomitant historical events? What is the stylistic difference between a marginal author and a canonical one? For these issues, Jockers proposes the practice of *macroanalysis*, appropriating Keynes' theory of *macroeconomics*. Instead of an atomised perspective on disparate literary pieces from which to draw general conclusions, the macro scale is more fruitful from a scientific point of view. Working on large and exhaustive corpuses and interpreting the results can open the way for new research areas: "The existence of huge data sets means that many areas of research are no longer dependent upon controlled, artificial experiments or upon observations derived from data sampling" (Jockers 7). Thus, literary research gets closer to scientific facts, to quantifiable phenomena and empirical observations. "The result of such macroscopic investigation is contextualization on an unprecedented scale" (Jockers 27), on the basis of which hermeneutics can be better applied:

> Combine a traditional literary training focused on close reading with the most
> common text-analysis tools focused on the same thing, and what you end up with is

---

[7] Thomas Rommel uses this concept in the chapter "Literary Studies" (88-96) from *A Companion to Digital Humanities* in order to describe the main uses of quantitative studies in this period.

enhanced search—electronic finding aids that replicate and expedite human effort but bring little to the table in terms of new knowledge (Jockers 17).

By combining the macroscopic view with the analysis of the internal factors of literary works, quantifying the range of influence and similarity between authors and charting the "linguistic macro patterns and macro trends present in a century's worth of novels" (Jockers 118), new breakthroughs in the literary research have been made.

The Romanian case

In Romania, quantitative studies reached neither the amplitude, nor the depth of the previously mentioned cases. In most of the cases, the contact of the Romanian critics, historians and theoreticians with quantitative methods was established either via structuralism (especially linguistics), or through the influence of French literary sociology and comparative literature. Digital humanities have not become a popular trend and distant reading has been generally dismissed for not being a viable way of comprehending the literary domain. When it comes to large-scale quantitative studies, they were mostly used as means of analysing the superstructure of the literary institution and the sociological dynamics inside it. Other usages of similar quantitative studies were in the domain of traductology and the circulation of translations as ways of documenting on literary movements. More recent local attempts at quantifying literary phenomena were used for similar reasons. The intrinsic study of large and exhaustive corpuses of literary works has not yet been established. However, the cases that shall be discussed in the further paragraphs prove that there has been a positivist and scientific tendency in the Romanian literary criticism. The scarcity of the examples and the minimal amount of ambition of these projects can only be perceived as symptomatic of a culture in which the most vocal tradition has been the impressionistic one, reluctant to the adoption of a single theoretical paradigm and mostly preoccupied with the infliction of aesthetic verdicts and feuilletonism.[8]

Before beginning the discussion about quantitative studies in Romania, we must analyse the structuralist case, as there is a strong correlation between it and the

---

[8] For more details on this particular issue, see Alex Goldiş's *Critica în tranşee. De la realismul socialist la autonomia esteticului,* Bucharest, Cartea Românească, 2011.

appropriation of quantitative methods in the autochthonous literary criticism. It must be said that unlike literary sociologists (Constantin Crişan, Mihai Dinu Gheorghiu), who were opposed to the local impressionistic tradition of criticism, Romanian structuralism was rapidly institutionalised and co-opted in the metaliterary discourse. That happened because "after all, theoretical poetics and aesthetic criticism are both proposing the isolation of the literary irreducibleness"[9] (Stan 354). This phenomenon of unrooting from the initial, completely different program of Western structuralism might be characteristic of the Romanian case. Solomon Marcus' *Lingvistica matematică* (*Mathematical linguistics*, 1963) and *Poetica matematică* (*Mathematical poetics*, 1970) can be interpreted as attempts at forming an algebraic poetical language, being similar to the early projects of computation of stylistic elements in 20th century's incipient forms of digital humanities. However, his usage of algorithms and the "coding" of semantic information can be only shallowly added within quantitative studies' spectrum, since "the most scientific from the Romanian structuralists [*i.e.* Solomon Marcus] uses the 'raw' language of mathematics and linguistics only to better emphasize, as with a contrastive solution, the literary 'ineffableness'"[10] (Stan 355). Since quantitative methods are closely related to the empirical and materialistic determination of the literary object, it is problematic whether to consider local structuralism's quantifications of words and narratological schemes proper "quantitative studies".

Therefore, the object of analysis in this part of the essay is the sociological area. A nodal point in the history of literary sociology of our country is the conferences volume *Probleme de literatură contemporană şi sociologie literară* (*Problems of Comparative Literature and Literary Sociology*), published by the Academy of Social and Political Sciences of the Socialist Republic of Romania in 1970, in which, along with traditional comparative essays based on parallelisms between individual local authors and international literary phenomena (Ionică Tăutu and Volney, Ioan Budai-Deleanu, Casti etc.), there are also studies with larger objectives. Thus, Engel Carol discusses the "Circulation of French Books in Transylvania during the Second Half of the 19th

---

[9] "[...] în definitiv, poetica teoretică şi critica estetică îşi propun ambele să izoleze ireductibilul literar" (My translation).

[10] "[...] cel mai scientist dintre structuraliştii români foloseşte limbajul ,dur' al matematicii şi lingvisticii doar pentru a pune mai bine în evidenţă, cu o soluţie de contrast, ,inefabilul' literar" (My translation).

Century" (117-128), Alexandru Duțu comments on the "Translation and Reconfiguration in the Romanian Culture during the Enlightenment" (155-160), and Victor Iancu talks about "The Receptivity of Romanian Literature in the German Culture" (231-241). Analysing the process of book circulation and of translation on large periods of time, the selected authors are not discussing isolated cases, but larger-scale, transcultural phenomena, on the basis of quantifiable data. Paul Cornea's study "'Demand' and 'Supply' in the Determination of the Translations' Profile in the Second Half of the Previous Century" examines the problem of the relevance of translations in the 19th century, based on the quantification and interpretation of data from this period. After mentioning Escarpit's theories, Paul Cornea decides to "eliminate the fastidious side of the documentation", mostly focusing on interpreting the results of his research. However, this reference is important.

The case of Paul Cornea is especially relevant, being one of the Romanian literary historians and critics whose enthusiasm for the empirical study of literature as a collective and social phenomenon was manifest in his theoretical approach to literature. Therefore, the "close"-"far" (*aproape-departe*) dichotomy that he conceptualises in one of his studies emphasises two methodologies of literary research that can be compared to the quantitative-qualitative dyad:

> If one wants to observe physiognomies and details, colours and tones, one is moving in the proximity of the object; conversely, if one is interested in big configurations and surfaces, not only texts but also the whole of the literary movement, one must take distance and have a perspective[11] (Cornea, Aproapele 5).

He emphasises the panoramic view of the literary macrostructures: "the perspective of this volume is one of *species* and not of *kin*, one of the *collective versant* of literature (genres, currents, typologies, concepts) and not of individual works and authors"[12] (Cornea, Regula 7). More than this, Paul Cornea explicitly proposes the quantitative

---

[11] "Cine vrea să observe fizionomii și detalii, culori și timbre, se deplasează în proximitatea obiectului; invers, dacă e interesat de marile configurații și reliefuri, nu numai de texte, ci și de mișcarea literară în ansamblu, trebuie să ia distanță, să-și ofere șansa perspectivei" (My translation).
[12] "[...] perspectiva asupra literaturii din volumul de față e cea a *speciilor* și nu a *spețelor*, a *versantului colectiv* al literaturii (genuri, curent, tipologii, concepte) și nu a operelor și autorilor individuali" (My translation).

methodology in the study of literature, which "must turn to supple methods, to use structural or dialectical analysis in order to explore the individual and statistics to explore the multiplicity"[13] (Cornea, Regula 21).

It can be seen that Paul Cornea opts for the interchangeability of methods, depending on the object of study. However, he is critical against Lucien Goldmann's "metaphysical" system (unlike Mihai Dinu Gheorghiu, whose case will be discussed below). From a sociological point of view, Cornea prefers Robert Escarpit's methodologies, especially when it comes to the sphere of the public reception of literature, and the fundamental characteristic of literary sociology is found, according to him, to "the utilisation of, as much as possible, quantitative methods"[14] (Cornea, Regula 24). However, Cornea uses quantitative studies only when quantification is accessible and possible. For example, the article "The Constitution of a Genre. Between 'Romance' and 'Novel': the Romanian Novel in the 19th Century" starts from the basic assumption that "when society is in the stage of capitalist take-off, [...] the novelistic creation is born in an artisanal way, out of the multiplication of individual initiatives"[15] (Cornea, Regula 265). Therefore, the lack of a novelistic mass production facilitates the counting of Romanian publications in the 19th century. This solitary counting of books would have been impossible in a hegemonic and productive culture such as the French one. This is also the case for the essay "Translations and Translators in the First Half of the 19th Century", from *De la Alecsandrescu la Eminescu. Aspecte – Figuri – Idei*. After presenting a short history of the usage of statistics in international literary historiography, referencing Havelock Ellis, Henry Peyre, Jacques Bousquet and, finally, Robert Escarpit (therefore explicitly manifesting his theoretical interests), Paul Cornea succeeds in explaining the contradictory synchronicity of classicism and romanticism in 19th century's Romanian literature by analysing the number of translations and the frequency of the translated foreigner writers in this epoch. Paul Cornea also takes into account the "paraliterary" phenomenon, often ignored by contemporary critics, in

---

[13] "[...] trebuie sa recurgă la metode suple, sa utilizeze analiza structurală sau dialectică pentru explorarea individualului și concursul statisticii pentru explorarea multiplului" (My translation).

[14] "[...] utilizarea, pe cât posibil, a metodelor cantitative, supunerea la imperativul verificării" (My translation).

[15] "[...] când societatea se găseşte în stadiul decolării capitaliste, [...] creaţia romanescă se naşte artizanal, prin multiplicarea unor iniţiative individuale" (My translation).

"Again about Romanian Romanticism" from *Delimitări şi ipoteze. Comunicări şi eseuri de teorie literară şi studii culturale*. However, his use of quantitative studies is limited to simply inventorying literary works and drawing conclusions from facts related to their frequency and circulation, but not about actual influences or shifts in literary forms.

Mihai Dinu Gheorghiu's approach to the sociology of literature is closer to Pierre Bourdieu's institutional analysis than to Escarpit's studies about the literary agents' social environment and the literary market. The Romanian sociologist is vocally critical to the Escarpitian quantitative approach for the reason that it encourages "the aesthetic element to sustain its metasocial nature"[16] (Gheorghiu 103). He mentions:

Ignoring the uniformizing quantitativism, the empirical relative of semiotic textualism, sociology has discovered a method which, beside the fact that it remains 'purely' sociological, axiomatically states the autonomy of the cultural: institutional analysis[17] (Gheorghiu 110).

Thus, Gheorghiu's perspective is deeply enrooted in the institutional basis of the literary act: his method is no longer a means of studying literature as an object, but to observe the relationships that form inside the sociocultural world and the transfer of capital that occurs within it. He is not reducing the literary phenomena to socioeconomic factors but rather merely observes the ways in which the sociocultural domain functions. That does not mean that he denies the use of statistics. What he observes is a lack of data for an exhaustive quantitative approach of the Romanian literary institutional phenomenon. The study "Notes on the Institutionalisation of Romanian Literature in 19th Century" from *Anuarul Centrului de ştiinţe sociale Iaşi* (1988), offers empirical evidence by means of statistics of the way in which national literature has developed as an institution and were later used by Mihai Lazăr.

In the field of institutional analysis of literature, Mihai Lazăr's study *Paradoxuri ale modernizării. Elemente pentru o sociologie a elitelor culturale româneşti*

---

[16] "[permite] esteticului să-şi susţină natura metasocială" (My translation).
[17] "Evitând cantitativismul uniformizator, ruda empirică a textualismului semiotic, sociologia a descoperit o metodă care, pe lângă meritul de a rămâne 'pur' sociologică, enunţă axiomatic autonomia culturalului: analiza instituţională" (My translation).

(*Modernity's Paradoxes. Elements for a Sociology of the Romanian Cultural Elites*) also uses statistics in order to portray the process in which the literary object had progressively become autonomous as a separate social branch in the 19th and 20th centuries. By analysing statistical data about the dynamics of the literary object in local publications in relationship to other cultural publications, he observed the way in which literature had become a hegemonic part of the cultural world in all Romanian regions. His use of graphs and factorial planes of analysis are useful for observing, for example, the relationship between the Capital and the periphery, taking into account each region's openness to literary modernity in the 19th century, as seen from the profile of local publications, proving there is a strong connection between modernism and the autonomic status of the literary institution. What Lazăr's study also aims at portraying is that the separation of the political discourse from the literary one was indeed a key-element in the consolidation of the modern literary tradition, but it was also the reason why the literary act had lost its central position in the State. Moreover, statistics on the economic and political status of writers in 19th century proves that, while at the beginning of the Romanian literary institution there was a strong connection between literature and politics and, by extension, feudal aristocracy, the latter two's influence on the first one progressively dropped. Therefore, political capital and cultural capital, at first interbedded, became by the end of the 19th century and the beginning of the 20th century two separate entities. This could explain, from a sociological point of view, literary phenomena such as the emergence of the "autonomy of the aesthetic" thesis or the appearance of Traditionalism in the 20th century. Marius Lazăr's study has the objective of presenting the sociological realities of the literary act and of relating them to its internal, artistic dynamics.

Picking up from Marius Lazăr's study, Ioana Macrea-Toma's study aims at analysing the function of the literary institution during a more controversial timespan. *Privilighenția. Instituții literare în comunismul românesc* (*Priviligentsia. Literary Institutions in Romanian Communism*) has the objective of examining the status of literature under the Communist regime, observing via statistics the economic and social position of the literary agents inside the Writers' Union of Romania, which formed in 1949 and had a fundamental role in the way literature had functioned and been publicly received during the last century. Macrea-Toma proves that there had been significant

economic advantages of being a member of the Writers' Union and that the position inside this institution had had an important influence in the development of literature. Analysing the shifts from the Stalinist regime (the '50s) to the temporary liberalisation of culture (the '60s), followed by the implementation of July's Theses (1971), *Privilighenţia* succeeds in portraying an accurate panorama of the Communist cultural life. The main point of Macrea-Toma's statistical approach to the literary institution is to show that the literary institution had kept being an autonomic social and economic entity even through the socialist realist period, by observing the writers' incomes, their cultural capital, and the importance of national literary prizes. Therefore, the advantages of being inside the literary institution during the Communist regime are emphasized: social and symbolic gratification and separation from the extra-literary space due to potential political risks had shaped the literary world in a fundamental way, making any explicit subversive movement impossible (Macrea-Toma 323-333).

However, these studies, although they use quantitative methods as a means of explaining the literary phenomenon, leave out important details about the literary object. Their usage is, therefore, not used for studying literature *per se*, but for studying the context in which literature appears and develops. But *how* exactly does literature develop, formally speaking? A Morettian approach to Romanian literature has not yet been made, but recent uses of the quantitative methods may lead into that direction.

For example, Ştefan Baghiu's article on "Translating Novels in Romania: The Age of Socialist Realism. From an Ideological Center to Geographical Margins" in *Studia UBB Philologia*, LXI, 1, 2016 (5-18) applies quantitative research methods in order to explain the way in which novel translations had circulated during the Stalinist period of the communist regime in Romania. The point of Baghiu's study is to describe the function of translation during the socialist-realist decade and how it had facilitated our country's contact with cultural diversity, often outside the Western world. This proves that, if before the rise of communism in Romania literary tendencies had been exclusively Eurocentric and modernist, a broader palette of influences was now available, despite the totalising tendency of Soviet literature. The study traces translations of novels from different continents and also from neighbouring countries. This may explain the development of the novel in the communist age by tracking down possible influences from distinct (sometimes exotic) cultures. Where exactly this jigsaw-

like influence is manifest or whether it had such a significant role in the novel production of this period requires further examination. Another quantitative study is Daiana Gârdan's "The Evolution of the Erotic Novel from the First Half of the 20th Century. Between Exercise and Canonization" in *Transilvania*, 7/2018 (5-10). It demonstrates that the erotic novels that had been established as canonical are mere refined versions of the popular erotic novel, which had a significant role in the novel production at the beginning of the 20th century. Statistical study of the frequency of literary forms, narratological plots and the evolution of the novel during the interwar period shows that what appeared to be an organicist development of the novel was actually an evolutionary process of stylisation of the popular erotic novel. What Daiana Gârdan's study lacks is the factor of translation and the circulation of foreign erotic novels, which could give a better insight into the production of this particular literary form. It should also be noted that both of these studies were based on data from dictionaries of publications and translations[18]. A digital library that could make possible the intrinsic study of these novels in order to understand influences and formal strategies, as in Jockers' studies, is not yet existent in our country, so local quantitative studies are limited to previous indexations.

Conclusion

Firstly, regarding the evolution of quantitative research on large corpuses of data, a change in the usage of this methodology can be observed. If literary sociologists used quantitative methods to study extra-literary factors such as institutional issues, book market, social environment or the reading public (this is the case of Robert Escarpit), the latest technological developments resulted in their appropriation in the intrinsic study of literary systems (Moretti and Jockers). In short, there is a formalist conceptual basis with the addition of methods from the sociology of literature and digital humanities, which resulted in a more comprehensive study of forms and their potential relationship with the historical context. More nuances to this theory shall be added in a further study.

---

[18] Ştefan Baghiu's sources are from *The Chronological Dictionary of Translated Novels in Romania from the Origins until 1989*, Editura Academiei Române, 2005, while Daiana Gârdan's data are collected from *The Chronological Dictionary of the Romanian Novel from the Origins until 1989*, Editura Academiei Române, 2004.

Secondly, the quantitative approach to literature has been reluctantly viewed by Romanian criticism. This is interesting, since there has been a locally established tradition of formalism and structuralism and the usage of statistics in the sociological studies on literature was appropriated by names such as Paul Cornea, Marius Lazăr or Ioana Macrea-Toma. The sociology of literature limited itself to observing either the circulation of books and translations in order to explain the origins of certain literary phenomena, or the dynamics inside literary institutions. More recent studies are coming closer to the Morettian perspective, engaging in speculation about the formal and evolutionary aspect of literature based on quantified data. These attempts are valuable. Nonetheless, there are not enough data sets to inquire into a deeper investigation of literary forms.

References

Baghiu, Ştefan. "Translating Novels in Romania: The Age of Socialist Realism. From an Ideological Center to Geographical Margins." *Studia UBB Philologia*, LXI, 1, 2016: 5-18.

Cornea, Paul. *De la Alecsandrescu la Eminescu; Asptecte – Figuri – Idei*. Editura Pentru Literatură, 1966.

Cornea, Paul. *Regula jocului*. Editura Eminescu, 1980.

Cornea, Paul. *Aproapele şi departele*. Editura Cartea Românească, 1990.

Cornea, Paul. *Interpretare şi raţionalitate*. Polirom, 2006.

Cornea, Paul. *Delimitări şi ipoteze. Comunicări şi eseuri de teorie literară şi studii culturale*. Editura Polirom, 2008.

Dima, Alexandru, editor. *Probleme de Literatură Comparată şi Sociologie Literară*. Editura Academiei Republicii Socialiste Românã, 1970.

Escarpit, Robert. *De la Sociologia literaturii la Teoria comunicării. Studii şi eseuri*. Editura ştiinţifică şi enciclopedică, 1980.

Gârdan, Daiana. "Evoluţia romanului erotic din prima jumătate a secolului al XX-lea. Între exerciţiu şi canonizare." *Transilvania*, 7/2018: 5-10.

Gheorghiu, Mihai Dinu. *Scena literaturii. Elemente pentru o sociologie a culturii româneşti*. Editura Minerva, 1987.

Jockers, Matthew L. *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press, 2013.

Kenny, Anthony. *The Computation of Style. An Introduction to Statistics for Students of Literature and Humanities*. Pergamon Press, 1982.

Laurenson, Diana T., Swingewood, Alan. *The Sociology of Literature*. Schocken Books, 1972.

Lazăr, Marius. *Paradoxuri ale modernizării. Elemente pentru o sociologie a elitelor culturale româneşti*. Editura Limes, 2002.

Macrea-Toma, Ioana. *Privilighenţia. Instituţii literare în comunismul românesc*. Casa Cărţii de Ştiinţă, Cluj-Napoca, 2009.

Marcus, Solomon. *Lingvistica matematică*. Editura Didactică şi Pedagogică, 1963.

Marcus, Solomon. *Poetica matematica*. Editura Republicii Socialiste România, 1970.

Moretti, Franco. *Atlas of the European Novel, 1800-1900*. Verso, 1998.

Moretti, Franco. *Graphs, maps, trees. Abstract Models for a literary History*. Verso, 2005.

Moretti, Franco. *Distant reading*. Verso, 2013.

Schreibman, Susan, Siemens, Ray, Unsworth, John. *A Companion to Digital Humanities*. Blackwell Publishing, 2004.

Stan, Adriana. *Bastionul lingvisitc. O istorie comparată a structuralismului în România*. Editura Muzeul Naţional al Literaturii Române, 2017.