Andrew Piper, *Enumerations: data and literary study*, The University of Chicago Press, 2018, ISBN: 978-0-226-56889-8, 243 p.

Review by Ana ȚĂRANU
Faculty of Letters, Babeş-Bolyai University
Cluj-Napoca, Romania
taranuanamaria@gmail.com

Over the past two decades, computational criticism has emerged from the conjunction of literary scholarship and Digital Humanities, consisting in the employment of digital data research tools (text and data mining, big data, machine learning) as operative instruments of literary interpretation. In adopting practices that develop within empirical fields, computational criticism has often been deemed a form of intellectual colonialism, insofar as its attempt at positivist objectivity seems to contradict the very rationale behind literary practice and its main critical tool, subjective exegesis. Its rise marks the encounter of the two unreconciled (and apparently irreconcilable) epistemic traditions of the humanities and the (conventionally) quantitative disciplines, and has been enveloped in a consistently polemic rhetoric.

Viewed from within these epistemic frictions, Andrew Piper's richly theoretical volume represents an innovative intervention, issuing a proposal of methodological reorientation of computational modeling within literary scholarship. In the words of its author, *Enumerations: Data and Literary Study* sets out to resolve, across the most elementary dimensions of literature, "a fundamental, and as yet unanswered, question: What is the meaning of literary quantity?"(2).

The book comprises six chapters, each charting the quantitative dimensions of an elementary literary feature (punctuation, plot, topoi, fictionality, characterization and corpus) within a particularly relevant historical period. In doing so, it resorts to the application of a variety of increasingly complex computational techniques, from "grep", the extraction of regular expressions, for punctuation in the first chapter, to the employment of vector space models and social networks to approximate plot in the

second one, to using machine learning to determine the fictionality of a text in chapter four, among others.

Aiming to bring readers "into this world, but always with an eye to concerns within literary scholarship" (6), the volume seems to consciously avoid the opacity that often surrounds research methods within quantitative studies. Committed to an explicit exposure of its analysis and its conclusions, it makes critical use of diagrams and graphs. Through constant reminders regarding the inherent limitations of its data sets, corpus selection and algorithms, it counteracts the field's mythology of universal models and transcendental data, mapping a new understanding of literary quantity.

"An extended demonstration of the ways in which computation, when applied critically and creatively, can confirm, revise, but also invent new narratives about literary history" (6), *Enumerations* explicitly marks its departure from the dominant discourse of quantitative literary criticism in its theoretically dense *Introduction*. The influential notions of distance, bigness, or objectivity that have come to define the field rely, the book argues, on "models of reading that are at once overly simplified and deeply binary in nature (distant/ close, deep/ shallow, critical/ attached)" (18). Dependence on such models has led to a misrepresentation of computational reading as inherently novel and empirical, overlooking the important ways in which it is "inevitably tied to the norms and practices of the past" (3). Additionally, Piper addresses disciplinary reluctance to quantity by reformulating it against the age-old animosity between literacy and numeracy, asserting that, for many, "Numeracy's rise signals literacy's eclipse, and with it a host of highly charged concepts like subjectivity, individuality, creativity, or even agency." (4). He dismantles this stance of implied mutual exclusivity, and argues for an understanding of the computational turn in literary studies not as "something distinctively new or even alien", but rather "as part of the history of humanism itself" (4):

> Today, there is a new translational imperative at work, one that aims to move between letters and numbers. Translating texts into quantities has emerged as the overwhelming feature of our cultural moment. Rather than see this as a kind of fallen state, I think we would do well to reposition it within a longer tradition of translational humanism, to see

it as part of an ongoing intellectual drama that tries to understand the act of commensuration, of making different sign systems compatible with one another (5).

The relocation of the computational *within* the history of humanism serves among the central premises of *Enumerations*, facilitating its reframing of quantitative studies. This rhetorical effort is reflected by the theoretical references which commence each chapter and its associated demonstration, all extracted from conventional critical tradition. For example, the first chapter operates with the Bataillean concept of "general economy"; chapter three begins with Curtius' idea of *Topik*, only to further introduce topological reading as a reviewed form of the post-structural project of intertextuality, whereas chapter four functions entirely within Searle's speech-act theory, only to empirically dismantle it.

The *Introduction* further establishes four new models of computational reading (as opposed to the binary tendencies of past practices): repetitive, implicated, distributed and diagrammatic. Each of them is linked to an essential concept or practice of quantitative analysis (repetitive linguistic structures, computational model building, distributional semantics and visual representation of data through diagrams), as it is reframed under the scope of literary methodology. Combined, they form a reviewed theoretical framework, not only for computational criticism, but for literary practice in general. In this sense, the volume exposes the ways in which a critical engagement of quantitative methods can help reorient the literary field without threatening its interpretive *ethos*.

For example, awareness of the repetitions of language can reveal "the ways in which quantity impinges upon meaning" (18), shifting the traditional preoccupation of literary criticism with "notions of breaks, ruptures and singularities toward questions of stability and duration, to see the deep grooves or furrows of literature, culture, or writers' lives"(19). Therefore, quantity gains its significance by providing us with the ability of rendering visible the semantic configurations of cultural practices that "manifest themselves in repetitive, often predictable, and sometimes excessive ways" (3).

Computational modelling is among the main modes of knowledge production that quantitative analysis relies on, and, given its origin in empirical disciplines and the

opacity of its algorithmic form, it has been mystified as a dehumanizing form of enclosing knowledge – it is precisely this rhetoric trend that the volume aims to dismantle. By acknowledging "both the subjectivity inherent in modeling and a basic instrumentality to reading that has been operative for centuries", the book argues for modelling as a "process of contingent world-making"(12), the main feature of which should be explicitness: "We bring our subjectivity and creativity to a model just as much as we cede portions of our subjectivity to it. Unlike the critic in exile, we are implicated in the very structures and networks through which we build our representations of those structures and networks" (19).

In light of these theoretical assessments, the first chapter traces the manifestation of punctuation within twentieth-century poetry, charting its excesses and lacks as its main points of interpretation. Despite the general antipathy towards the normative use of literary punctuation that characterized last century's poetics, an exponential increase in periodization emerged. The meaning of this phenomenon is assembled through a computational reading of a corpus of 75,000 poems written in English by 452 poets, looking to identify "commonalities of expression that surround such excessive punctuatedness" (31). The empirical results of the research ultimately undergo a transition into literary interpretation, contained within the concluding statements of the chapter: "the period, in its excessive poetic state, does appear to have a particular kind of meaning associated with it. It asks us to look both ways, to be in the moment but also think more generally, to move from the very similar to the very dissimilar, to hold a paradox in our heads without clear ends" (40).

Employing distributional semantics, the second chapter analyses emplotment by focusing on the relationship between narrative discourse and its content. Considering novels that are interested in exploring a form of social constraint as linguistically experienced, it ultimately encourages "a fundamental rethinking about plot away from high-level motifs or content and toward a distributional understanding of language" (65).

The last four chapters pose inquiries into how topic models reconstruct the spatial nature of language, the ways in which machine learning can provide insight into the linguistic nature of fictionality, the problem of character as narrative space and, ultimately, of poetic corpus as an aging body. They constitute relevant interventions into

their associated literary debates, providing examples of a literarily sensitive employment of quantitative methods.

The *Conclusion* of the volume constitutes a self-aware attempt at authorial implication. Admitting that "it is precisely our inability to account for ourselves in the analytical process that empties our work of its critical force" (179), this last section is dedicated to a quantitative self-assessment of the author's work, as it stands among similar publications, and of the field of the humanities, its institutional inequality rendered visible. Ultimately, Piper calls for an instrumentation of data that supports the ideal of epistemic equality, assigning computational humanities an ethical imperative: "Through this work we are trying to imagine new forms of algorithmic openness, where computation is used not as an afterthought — as a means of searching for things that have already been preselected and sorted — but as a form of forethought, as a means of generating more diverse ecosystems of knowledge" (185).