

Cornelia Wilhelmina KOOLEN, *Reading beyond the female: The relationship between perception of author gender and literary quality*, Institute for Logic, Language and Computation, ILLC, 2018, ISBN 978-94-028-0951-0, 324 p.

Review by Anca ȘERBAN

Faculty of Letters, Babeș-Bolyai University

Cluj-Napoca, Romania

anca.serban10@gmail.com

The intersection of gender studies and digital humanities emerged as a necessary investigation of the field's patriarchal and hegemonic rhetoric that kept on furthering biased preconceptions, safely justified by the perception of an unmediated and complete access to knowledge. In an attempt to include and promote feminist scholarship, computational analysis is now investigating canonical theories regarding gender or 'genderness' in writing, which have always justified the exclusion and difficulty of women authors gaining literary prestige and entering the canon. Well known texts of the 20<sup>th</sup> century, from Woolf, to French feminism or classics such as Toril Moi, expanded the debate on what the limits of an existing 'feminine writing' could be. A century of women authors had to transcend correlations of domestic topics such as 'settling down', sensibility in writing, and 'feminine rhetoric' in order to adhere to acclaimed literariness. As much as one side of the debate wanted to force the inclusion of the newly emerging female authors through dropping the categories of men or women, the idealistic possibility of one to 'forget their gender when writing', that feminist pillars wrote about on repeated occasions, remains inaccessible today still. Or rather, although authors moved past the need to assign themselves a certain gendered writing style, reading cannot be done beyond the implications of gender or 'genderness', especially if literariness or literary quality is in focus.

C. W. Koolen's *'Reading beyond the female' (2018)* unravels the relationship between author's gender and the biased assignment of literary quality to a piece of work. The study focuses on two aspects: 1) the prestige of female authors and the perception of their work, in comparison to male authors; 2) the texts of the novels themselves, and what their content can reveal about the expectations and stereotyping of female authors.

These expectations are defined by Koolen as *the gendered lens* – a way of reading a text that is either subconsciously or consciously affected by the gender of an author.

The thesis is published in the context of the Dutch digital humanities project *The Riddle of Literary Quality*, founded by the Royal Academy of Arts and Science. The main focus of the project is the link between textual characteristics and judgement of literary quality, through a combination of computational methods of content analysis, and ratings of literary quality gathered from a national reading survey, having the same novels as subjects. For quality and accuracy's sake, the ratings had to be divided into literary quality and overall quality on a scale of 1 to 7, only for the novels that respondents had read from a list of 401 recently published – translated or originally Dutch works – fictional novels. Koolen's work uses the data in order to visualize patterns in the perception of the female versus male authors, in order to tackle issues considered long gone, in a literary field recently infused by the myth of a settled equality between the genders. A great part of the thesis is therefore focused on demystifying the constructed nature of these gendered lens, and ultimately, how can we *put on different glasses to make it more likely that female author's work is judged solely on its own merit* (33). In other words, how we can read beyond the imposed limits of *the female*.

The merits of computational analyses cannot overtake their lack of neutrality, and thus digital methods cannot supply an ultimate answer to the questions addressed. As methodological plurality is necessary for revisiting such an incomplete field, Koolen suggests a dual-perspective approach, meaning to avoid the analysis taking place in a social vacuum, full of biases and apparent neutrality. She brings forward a solid foundation from two separate disciplines: literary sociology and Natural Language Processing, dividing the study into two parts. The very first, 'The literary field: perceptions of literary quality', follows the literary sociology perspective, questioning the biased correlations between the gender and the perceived quality of an author's work in the Dutch literary field. The second, following a digital humanities perspective, revisits the assumptions discovered in the first part, in order to answer the question *to what extent can an author's writing style be ascribed to the author's gender?*

The first chapter settles the context (the study of literary quality, through the responses of a specific audience, the national reading survey), the key concepts (literary field, literary quality, genre, style), and the corpus: a series of recently published novels

that gained large audiences, and exist under the ‘best seller’ title. The meaning of *the literary field* takes from Bourdieu’s concept (1983), a field being represented by a number of actors from authors, readers, publishers, to award juries and educational institutions. Bourdieu’s literary field functions on two poles: one of publishers, who aim for large-scale production with the goal of gaining economic capital, and one of publishers strive for limited production, but of higher quality (social capital). As literary production changed, the current literary context places the most social capital on the literary best seller, a form of literary novel that might be placed between the two poles.

What came into focus about *literary quality* is its questionable merit, once the professionalization of the trade of novel writing emerged – is the quality of art a social construct? Nowadays, it could be seen as a combination of textual characteristics that professional readers tend to believe as a consensus. It also works for readers as a proof of literary establishment, therefore maintaining a combination of both extra-textual and textual qualities.

*Genre* is used as both a cognitive category and an organizational principle. The study understands it mainly through the eyes of publishers: *literary* is thus used as a genre in itself, contemporary applied as a marker of *general fiction*. Another key concept is *style*, which remains defined as a property of texts constituted by an ensemble of formal features, observed through the study quantitatively or qualitatively.

Finally yet importantly, *gender* is an obvious key issue of this debate. Although gender studies reached a point past the need of gender dichotomies, Koolen picks the most ‘workable of models’, that of female and male difference, although she reclaims the complexity of the gender spectrum. She revisits the concept of gender in Butler’s theory, as an identity which is different from biological sex and performed to the outside world. However, digital humanities still face a difficulty to operationalize gender. Computational methods mainly include a narrow view of gender, resuming to the binary model. Moreover, the authors are usually taken to be either female or male by publishers and readers, which is the perspective Koolen investigates, with the intention of remedying this perception.

A solid foundation for her approach is set through the second chapter, ‘Contemporary female authors and prestige’, in which Koolen convincingly disassemble the literary ‘ladder’, a five-step path that authors have to partake in order to gain

prestige, from professional authorship to the highest level of literary prestige: having won a literary award. The author defines *literary prestige* in terms of importance, while remaining unsure on what grounds to base this hierarchy. Although the Dutch literary field leaves the impression of a general improvement of the accessibility, Koolen not only questions the field's internal functionality, but also checks the percentages of the genders at each stage, proving the opposite. A decrease in the numbers is visible only when researching the previous ten years: after 2016, only 10% of women won the two biggest literary prizes while since 2007, female authors have not surpassed the 30% mark. Contrary to common perception, there has been no obvious progress since the conceptions of these awards. The generally accepted presumption is that the gender imbalance will resolve itself, being caused by a lack of professional female authors in the literary history. However, Koolen argues that, contrary to common perception, female authors were not outnumbered by male authors in the 19<sup>th</sup> century, but their novels were simply valued less and never canonized. As a consequence, once the beginning of 20<sup>th</sup> century hit, the sudden abundance of women's literature forced a shift of perspective: the overall quality of literary production will now be ruined. The difference in literary prestige is hence not caused by a lack of literary production; said female authors produce hundreds of literary novels per year. "Nor is it caused by the idea that women cannot write, since the general reader does not appear to believe that being female hinders an author from writing a literary text, as the survey indicated" (246).

Likewise, chapter three, 'Judging bestselling fiction', discusses the results of the National Reading Survey, examining the larger reading audience. The results are not positive either way: even though half of the novels in the corpus have been written by female authors, they are overrepresented in the genre of romantic fictions, which holds the overall lowest average rating. The author does not assume a causality between author gender and rating. Rather, she maps the likelihood of female and male authors to be in the position of being judged as the producers of highly literary quality works. The survey works accordingly on two scales: *literariness* and *overall quality*, as some books may seem very literary, but not very enjoyable, while some books are not very literary, but highly enjoyable. On that account, romantic novels are seen as being the least literary, but also of the least overall quality. The results show that genre can, in fact, precede or replace judgement of the actual text. If a novel is perceived to be a thriller or

a chick-lit novel, it can be automatically be excluded from a high rating, without further explication even being necessary.

That being the case, chapter four, 'Theoretical background: connecting (female) gender to text', introduces the theoretical framework. Genre is not the only element that influences textual judgement; topic, plot, style is also usually associated with femaleness, which serves as a dismissal. Whence the women's novel became the lowest possible qualification. Although chick lit is similar in structure, style and simplicity to thrillers or detective novels, the genre hierarchy remains as romantic novels remain still at the bottom. The 'emotional' style visible through choice of words and subject of a woman's novel equated to a large extent with being less literary though the entirety of 20<sup>th</sup> century (Van Boven, 1992), (Vogel, 2001). Contrary to the idealistic perspective, his association is visible in the survey's responses still.

As follows, chapter five, six and seven apply computational analysis to the selected corpus in the following manner: chapter five through a bottom-up analysis, starting with the full texts, while chapter six uses a top-down analysis, starting with a single topic. Chapter five implies questioning the simplistic links between author gender and texts, examining subject, sentiment and overall style. The 'gender differences' appear to be large at first. However, Koolen brings forth for questioning the typical lack of rigor when building a corpus for similar studies: once the analyses are performed on a different corpus, (where the novels were all nominated for a large literary prize, instead of a selection of best-sellers) the gender difference turns out to be significantly smaller. In consequence, the computational methods cannot supply a complete, thorough result and a manual rigorous selection and analysis is required.

That being said, final two chapters perform researches starting with specific topics and subjects identified in the previous chapters as 'feminine', the most noticeable being the excessive attention to physical appearance. The methods used are firstly, automated extraction of sentences that have 'a description of physical appearance of a human character' as their main subject, through basic frames of sentences written manually (Appendix B), and secondly, machine learning, where the computer learns which sentences contain such a description based on sentences marked manually that contain such a physical description. The results dismantle most of the arguments that kept female authors outside the professional authorship realm: the shallow interest for

physical appearance seems to be given more importance in the literary novels than in the chick lit novels.

Consequentially, chapter seven analyses the phrases individually: a set of 200 sentences, which contain a description of a character's appearance. For this purpose, the concept of 'male gaze' (Laura Mulvey, 1975) comes in hand, signifying a break from narrative action, a pause as it were, to reflect on something that the focalizer in the novel sees. When reflecting the point of view of a male character, the gaze subjects the others to a 'controlling and curious gaze', while *looking* becomes a source of sexual pleasure and the person looked upon a victim, a passive entity, enacting the viewer/object or active/passive dichotomy. The concept of male gaze is useful in identifying 'viewing behavior' patterns that could tell about *who looks at whom and how?* By this point in the study, it is safe to say that the gender of the author is important, but not crucial. More importantly becomes the gender of the main character, a factor which determines the nature of the pattern. Koolen establishes three recurrent patterns: 1) the Caring Adonis, the beautiful man who pampers the female protagonist, being mirrored in literary novels by 2) the Destructive Nymph: a beautiful woman, always delicate and childish in appearance, who seduces and betrays the male protagonist. The third is 'the gaze-upon-herself', a series of descriptions where main characters scrutinize themselves in comparison to beauty standards or other characters. A few distinctive factors place the first two in anti-thesis: The Destructive Nymph is sorely described from the male protagonist's perspective, without the ability to look back and assume the role of the viewer themselves, while chick lit is more democratic with sharing the gaze between the Caring Adonis and the main character. As a consequence, men are looked at as potential love interest mainly in chick lit, while hardly being described in literary novels. As a result, algorithms will depict descriptions of appearance as more 'feminine' in *literary novels* taking into consideration that women are most often the subjects of these descriptions and the wording and general style is adapted for their frail and feminine figures. Thus, chick lit doesn't adhere to the preconditions of the study.

Another marker for femininity in writing is the 'settling down' topic, which, contrary to common belief, is similarly present in male-authored novels than 'masculine topics' such as military. That being the case, barely a few set preferences in writing style have been indicating a significant link to a specific gender: female authors do have a

preference to dialogue over narrative, characteristic found in half of the texts by female authors. Incidentally, these texts have been correctly classified by the algorithm. With almost 83% accuracy, the machine learning classifier predicted the author gender based on the text of the novels. However, as the computational methods were lacking the necessary context and interpretation, femininity kept being more easily identifiable in literary novels through descriptions of character's physical appearance.

Now, it can be reasonably said that femaleness is not an automatic consequence of an author's gender, the study proving a more significant importance of the main character's gender. The majority of presumptions that the study begins with tend to be demolished: appearance is recurrent in literary novels more than romantic literature, while author's gender is not indicative of the 'way of looking', except for the simple difference that the male protagonists do not look at themselves, but look at women. As a result, literary authors spend the largest part of the sentences describing women. On the other hand, chick lit authors reduce the complexity of this self-reflectiveness through what Koolen named 'the gaze-upon-herself': a fear of looking a certain way, generated through the same lens men look upon women in the literary novels. In other words, chick lit cannot be seen as separate from literary production.

The author makes a persuasive case for dissecting the dominant perspective on gender differences in writing style and the subconscious biases that have a bigger influence than some like to believe. She repeatedly states that the issue lies in preemptive categorization: not the difference in writing style or topic, but oversimplification and stereotyping.

The final suggestion is "putting on a different pair of glasses as it were" (251), allowing us to read beyond the female and thus doing justice to authors of all genders, while keeping literature connected to, but not solely dependent on gender.