CZĘŚĆ DRUGA: PRZEGLĄD LITERATURY SOCJOLINGWISTYCZNEJ

JOANNA BYSZUK

Instytut Języka Polskiego Polskiej Akademii Nauk, Kraków
ORCID: 0000-0003-2850-2996

A REVIEW OF THE BOOK *READING BEYOND THE FEMALE:* *THE RELATIONSHIP BETWEEN PERCEPTION OF AUTHOR GENDER AND LITERARY QUALITY*

ABSTRACT

Presented text is a review of the book "Reading beyond the female: The relationship between perception of author gender and literary quality" by Dutch researcher Cornelia Koolen. Discussed book undertakes the issues of relations between the gender of the author, evaluation of literary quality of their work and actual features of the texts, thus fitting in the larger trend of research on gender stereotypes in language and literature. The innovative use of quantitative methods also grants it an important place within literature on stylometric, making it an interdisciplinary work.

'Female authors still remain back-benched when it comes to gaining literary prestige: novels by females are still much less likely to be reviewed, or to win a literary award' (Koolen 2018: 129) notes Cornelia Koolen, the author of *Reading beyond the female: The relationship between perception of author gender and literary quality*, adding that gender-labelled adjectives and expressions are commonly used in describing books written by female authors, whereas the situation is not the same for men. The presented book deals with the subject of the relationship between the gender of the author, an evaluation of the literary quality of their work and actual features of the texts, thus fitting in with the larger trend of research on gender stereotypes in language and literature. The author adopts an inter-disciplinary research approach, combining traditional feminist methodology and natural language processing, making the book an important source for digital humanists investigating gender in language.

Taking up the subject of the influence of gender on creativity, the author was aware of the problems and limitations associated with the binary perception of the socio-cultural

gender: difficulties in distinguishing one of the elements making up authorial style (gender vs social background, nationality, education) and bias persistent in many previous quantitative studies, addressing differences between the styles of men and women. In her text, she repeatedly pays attention to careful selections of datasets (texts only by authors clearly specifying their gender), making conclusions and avoiding generalisations. Despite her criticism of current research using natural language processing methods in gender-related style analyses, the author is against abandoning such studies, emphasising the still significant impact of the social perception of people and their creativity through the gender perspective, and the role of quantitative research in identifying areas of discrimination.

The book is divided into two parts and consists of 8 chapters. The first part (chapters 1–3) focuses on the results of the National Reader Survey carried out in the Netherlands in 2013, and presents an analysis of how readers evaluate books. The second part is dedicated to a discussion of empirical research, carried out using both computational and traditional methods on the corpus of novels discussed in the survey. Part 2a focuses on the detection and characterization of possible differences between books written by women and men, and part 2b focuses on the analysis of descriptions of the characters' outer appearances.

In Chapter 2, Koolen provides an outline of the dataset used in the study — the results of the National Reader Survey (Dutch: Het Nationale Lezersonderzoe), a part of the *The Riddle of Literary Quality* project, carried out at the Huygens Institute in the Netherlands, in which she participated. The aim of *The Riddle* project was to discover factors influencing the perception of a text as literary and general opinion on its value. This was possible thanks to results from a survey in which readers were asked to determine how literary a novel seemed to them and how much they liked it (on a 1–7 scale), and to provide a brief opinion on one randomly selected book from the ones they evaluated. The study was based on the 401 most often-bought or borrowed books, including 152 original Dutch works and 249 translations from other languages, varying in terms of genre and balanced in terms of the representation of male and female authors. During the 6-month period of data collection, 13,784 people were interviewed (9,791 women, 3,897 men and 96 people who did not give their gender), which, as the author notes, is a large and quite representative sample for a language used by a relatively small population. Although the respondents did not have access to the collection of texts, it was available to the researchers, and the author used it in further analyses, including the studies on gender and stereotypes related to it.

Chapter 3 examines opinions of the readers by means of linear regression. Koolen focuses on the relationship between the evaluation of a given book and the variables such as genre, the author's gender, and whether the book is a translation. Her main finding is that books written by women score lower both in terms of literariness and generally understood quality, even when other variables are taken into consideration. The books written by men were on average rated slightly (0.5 points) higher, and the genre of the text was rarely significant. The fact of being a translation did not have a significant

or permanent impact on how a given book was rated. However, as the author points out, while translation did not affect the evaluation of individual texts, it influenced the general perspective: the translated books were perceived by the respondents as better than those written by the Dutch.

In the same chapter, the author also looks at how the gender of a respondent influences their judgement, revealing that both in case of the general result and just literary texts, women rated male authors higher, and the other way round, and yet average overall scores for men were higher anyway. The general trend indicates that men rate books less critically, giving higher scores, and that representatives of both genders are more likely to rate books written by male authors higher. The author also notices a pattern showing that male authors are read more frequently. Particularly interesting for linguists is the analysis of the comments accompanying the scores, carried out using the keyword method. Koolen observes grammatical choices indicating that men are perceived as more active creators, with emphasis placed not on the gender of the author but on their creation, while women are described as more passive and in relation to their gender. She also finds that the comments on books by male authors were oriented more towards a reflection on style, whereas comments on female-authored books are about the feelings they engendered. This led to two new hypotheses: 1) readers judge texts they believe to be written by a woman as worse, and 2) female readers are more critical towards female writers. These hypotheses were examined in an additional study, which led to a rejection of these hypotheses.

Part 2a of the book (Chapters 4 and 5) is dedicated to a computer analysis of texts contained in the examined corpus, with the aim of verifying whether there is a relationship between the literary and linguistic quality of texts written by men and women and their reception. In Chapter 4, the author examines how readers perceive authorship, genre, writing style and subject through direct analysis of comments, which allows her to discover, among other things, that the terms related to the female gender (e.g. 'this is a fairly feminine book') are used to describe the text negatively, whereas the terms related to male gender are never used in this way. Koolen notes that this is a harmful situation for literature: forcing women to adapt to the male standard instead of writing with their own voice.

Chapter 5 presents an exploration of the corpus using several computational methods: 1) the LIWC program which analyses a text on the basis of the distribution of words and conceptual categories, 2) machine learning, a classification based on the lemma frequency distribution using the Support Vector Machine algorithm, and 3) topic modelling. Importantly, a large part of this chapter is dedicated to a critique and potential pitfalls of these methods and approaches, showing the deep reflection of the author. The chapter brings interesting observations, noting that while all issues seem to be addressed by writers of both genders, they have very different proportions: e.g. the thematic category "body" is usually explored by women to a fairly similar extent, while men tend to either use it a lot or almost not at all. The author also notes that male authors were more easily recognised by the classifier, which shows a great

potential risk of bias, and should be examined in future studies also involving other datasets and languages.

In Part 2b, the author focuses on the image of a woman and a man in excerpts from a smaller corpus built of literary texts and the so-called chick-lit novels — easy popular literature, written typically by women for women. The analysis of sentences describing the physical appearance of character revealed that the gender of the protagonist significantly affects the style of description. Koolen distinguishes four image patterns: overt criticism of the female protagonist by female authors; the "perfect" man in the eyes of chick-lit female authors; a "destructive nymph" female type pictured by male literary authors; and, strikingly, a lack of pattern in how men portray men. The author emphasises the lack of diversity in what is considered beautiful across the whole corpus. Not signing into the canon of being slim is always an element of criticism of the appearance of a heroine.

The book is undoubtedly interesting and worth the attention of all engaged in research on gender stereotypes and methods of text analysis, especially style. Its unquestionable advantage lies in its comprehensive approach to the problem: by undertaking a series of studies varying in methodology, the author resists the temptation to draw easy conclusions arising from individual experiments, and describes the problem from many perspectives, warning against hasty interpretation of results. Extensive and accurate descriptions of conducted computational experiments (including explanation and criticism of the methods) and linking them to online repositories containing the code show a great sensitivity to Open Scholarship culture. Finally, from a research perspective, the study also makes an important contribution to the development of stylometric research methodologies in relation to gender studies.

A slight weakness of the work is the rather small collection of book data, which, however, is noted by the author, emphasising many times that the results concern a narrow setting. While the gathered corpus of opinions is sufficiently large to make some generalisations, for books, depending on the experiment, the number of the considered texts varies from 30 to 378, which allows for observations to be made about the collection itself, or about a certain part of popular literature in the Netherlands, but not about real tendencies in language or literature in a broader sense. As mentioned, the author is aware of these limitations and carefully formulates her conclusions. Nevertheless, the results form an important contribution to the study of stereotypes in language and literature, and the significance of revealing the absence of major stylistic and qualitative differences in the literature created by men and women cannot be understated: the refutation of this accusation is an important contribution to the debate on the reasons behind the rare awarding of prestigious literary prizes to women.

REFERENCES

Koolen C.W. 2018: *Reading beyond the female: The relationship between perception of author gender and literary quality*, ILLC Dissertation Series, Amsterdam, open-access, online: <https://dare.uva.nl/search?identifier=cb936704-8215-4f47-9013-0d43d37f1ce7> [1.10.2019].

STRESZCZENIE

Recenzja książki Cornelii Wilhelminy Koolen *Reading beyond the female: The relationship between perception of author gender and literary quality*

Słowa kluczowe: recenzja, gender studies, analiza dyskursu, stylometria, literaturoznawstwo.

Niniejszy tekst stanowi recenzję książki *Reading beyond the female: The relationship between perception of author gender and literary quality* holenderskiej badaczki Cornelii Koolen. Prezentowana publikacja podejmuje tematykę relacji między płcią autora, oceną jakości literackiej jego lub jej twórczości i rzeczywistymi cechami tekstów, wpisując się tym samym w nurt badań nad stereotypami płciowymi w języku i literaturze. Dzięki innowacyjnemu zastosowaniu ilościowych metod analizy tekstu stanowi też istotną pozycję w zakresie metodologii stylometrycznej, co nadaje całości pracy interdyscyplinarny charakter.