

# The Portrait of a Lady: Close and Distant Reading of Media Gender Bias

Abstract for short paper accepted at DH Benelux 2020

Astrid van Aggelen, Laura Hollink  
Centrum Wiskunde & Informatica, Amsterdam

11 March 2020

The Dutch media, we hypothesise, discuss men and women differently. This hypothesis is based on two key findings. First, women have been shown to appear in the media less frequently (e.g. [2, 6]) and in different contexts than men (e.g. as a vox pop rather than a voice of authority [5]). Second, language is known to reflect stereotypical beliefs [1]. For instance, we tend to highlight aspects that defy our expectations (a *career-oriented* woman [1]). Our ongoing study assesses whether such biases are present in descriptions in a Dutch newspaper of women in leading roles, and how these differences relate to stereotypical gender frames.

In the first part of our study, a computer system guesses the gender of the persons described in unseen texts, after being trained on a large set of gender-labelled examples. If it succeeds, this suggests the descriptions are gender-partial. In fact, the better the system is able to gender-label the portrayals used at test time, the more likely it is that the corpus as a whole displays gendered language. Unsupervised computational methods like this have proven capable of detecting stereotypical frames [3]. The classification study, which is nearly completed, is applied to at least 50k person descriptions from the newspaper NRC.

The portrayals are not presented to the computer as full texts; rather, we collect so-called *microporraits* [4] from a selection of syntactic features, including the labels and properties ascribed to the person, the actions they undertake and undergo, and the verbs used to quote them. The pronouns, given names and other obvious gender giveaways are used to derive the gender label that corresponds to a description, but are withheld from the feature set. The classifier learns the relationship between the features and the gender label on the bulk of the descriptions. The resulting model is then evaluated on a smaller set of descriptions.

In the second part of the study, which is in preparation, we collaborate with a social scientist to carry out an in-depth analysis based on close reading. The annotators are presented a set of women's descriptions and are asked whether they display each of the nine female stereotyping dimensions identified in [6].

The resulting assessments are contrasted with the accuracy and confidence level of the classifier on the same texts. If they match up, this suggests there are linguistic features that co-occur with recognised aspects of stereotyping. An inspection of these features and their gender-predictive value, then, will shed light on the “systematic asymmetries in word use” [1] brought about by stereotyping. Hence, the close- and distant-reading approach in this study have a synergetic purpose: close reading to corroborate distant approaches, and distant reading to deliver new insights for the social sciences.

## References

- [1] Camiel J Beukeboom et al. “Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies”. In: *Social cognition and communication* 31 (2014), pp. 313–330.
- [2] James Bryson and David Bunker. “Blurred lines: Exploring contemporary attitudes to gender portrayal in the media”. In: *Market Research Society Annual Conference*. 2015.
- [3] Dallas Card et al. “Analyzing framing through the casts of characters in the news”. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. 2016, pp. 1410–1420.
- [4] Antske Fokkens et al. “Studying muslim stereotyping through microportrait extraction”. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. 2018.
- [5] Joke D’Heer et al. “The bits and bytes of gender bias in online news: a quantitative content analysis of the representation of women in Vice. com”. In: *Feminist Media Studies* (2019), pp. 1–17.
- [6] Hanne Vandenberghe. “Representation of women in the news: Balancing between career and family life”. In: *Media and Communication* 7.1 (2019), pp. 4–12.