

Response to the Reviewers' Comments

Marieke van Erp, Niels Dekker, and Tobias Kuhn

Dear Editor and Reviewers,

We are happy that our previous revisions were largely satisfactory to the reviewers and we are grateful for the additional comments and suggestions to improve our article. We have revised our manuscript accordingly. We have addressed the typos and missing punctuation marks spotted by Reviewer 1 and proofread the text again. Furthermore, we have followed Reviewer 3's comments on strengthening the positioning of our work. Please find detailed answers regarding these modifications below.

I. DETAILED RESPONSE TO REVIEWER 3

[...] I still think that the positioning of the paper can be even more specific. This increased specificity begins with the first paragraph to the research questions, in the introduction. The very first paragraph attempts to position the study in a larger field, but in doing so it not only incorrectly positions the contribution and scope to making contributions to literary theory (which the paper does not do), but also casts the scope so large to quantification of literary studies; when in fact your study is much more specific contributing to knowledge specifically to character network identify through NLP techniques. As the Game of Thrones example proves and which you state on page 13 line 461 "This is a design choice that [462] will depend on the type of research question one wants to answer by analyzing the social networks". Yes! Yes! This is your contribution! in order to understand the validity and reliability of these character networks and thus our understanding of social structures in novels, this starts from the nature that the nodes and edges were automatically identified using the NLP/NER tools. This is so fundamentally important, and your approach gives scholars great insights into which tool is best suited for their study or research question. Thus drop the entire framing of contributions to literary theory, claim your findings and their boundaries and roll with them!

Response 1. Thank you for your kind words and insightful analysis of our work. We have amended the introduction to draw more attention to the motivation behind evaluating the robustness of automatic extraction methods in order to know how they behave on different kinds of texts.

Thus, I would suggest a re-positioning. This would mean that I would highly recommend removing the first paragraph completely, and instead start out with why the understanding of social networks in novels are important in literary studies. Then narrowing your scope to how these networks can be identified, and the NLP approaches differ and how. Thus you then have one more general research question: To what extent do NLP techniques using NER differ in identifying social networks in novels? / as to how to operationalization networks (really what is a node and edge) and how that translates to our understanding of the projection of the networks and thus analysis of the networks. Where you first look at the NER process and compare it, and then the network itself and its relationship to the extraction/identification.

Response 2. We have removed the first paragraph and rewritten the second in order to bring forward the focus on the analysis of NLP techniques.

In addition, this suggests that your framing of the importance and selection of the two time periods instead becomes part of your case selection, not part of your aim. It is not your aim to contribute to our understanding of different periods, you even say that no differences have been found and no other research has been done this (starting at page 2). Thus I would highly recommend you delete this claim/positioning. Instead mention the time periods and genres in the methods section as a variable you consider to insure that it is applicable to "all" periods; much like you compared these networks to real world networks. This serves to further ensure the validity of the findings, but it is actually irrelevant to the approach and contribution you aim to make.

Response 3. We selected the two periods because we realised that in prior work we had not and have not yet seen these types of analyses done on contemporary literature and one of our questions initially was whether the tools would behave differently. The fact that we did not find these differences in our opinion does not preclude us from explaining this in the introductory section.

We have amended our text to put less stress on the comparison between the two periods, but still mention it and keep the distinction between the two corpora, as we deem it important that the corpora on which NLP tools are tested are as diverse as possible.

This then suggests that the title becomes more general and thus more powerful! evaluating social network extraction tools for fictional literature.

Response 4. Thank you for your suggestion, we have updated the title to reflect the focus on the evaluation of the extraction more.

In addition, given one of your core audiences is literary scholars, I still adamantly oppose to you using the word classics, these are canonical works as you also mention now in a footnote. But no literary scholar is going to take your work seriously if you from the outset incorrectly categorize this core term in their field.

Response 5. As we state in the manuscript, as well as in our previous response letter, we are following Charles Augustin Sainte-Beuve in our nomenclature. Furthermore, we do not refer to our corpus of 19th century novels 'classics' but 'classic novels'. We also consulted with a literature studies professor at one of our institutes who did not object to this. We therefore think that we sufficiently distinguish our dataset from the ancient Greek and Roman classics.