

# Character Centrality in Present-Day Dutch Literary Fiction

Roel Smeets

Radboud University Nijmegen, Department of Literary and Cultural Studies

Eric Sanders

Radboud University Nijmegen, Center of Language and Speech Technology

## Character hierarchies

A focus on character relations has lain at the core of ideological approaches to literature (e.g. Pattynama 1994, Murfin 1996, Meijer 1996, Pattynama 1998, Minnaard 2010, Sidi-Said 2010, Song 2015, Mushtaq 2010, Fatima et al 2015.). When researching e.g. gender, race or class representations in literary texts, a convenient point of departure is to look at the ways in which characters are related with respect to their demographic profile. On the basis thereof, hierarchies between e.g. male and female characters, Western and non-Western characters, lowly and highly educated characters, can be analysed and interpreted.

In the last decade or so, a growing number of researchers have been applying social network analysis to fictional storyworlds (Alberich e.a. 2002, Stiller e.a. 2003, Elson e.a. 2010, Lee & Yeung 2012, Karsdorp e.a. 2012, Agarwal e.a. 2013, Jayannavar e.a. 2015, Karsdorp e.a. 2015, Lee & Wong 2016). The method of social network analysis provides promising perspectives on the analysis of character hierarchies in literature as it enables a more quantitative, formalised and bottom-up view on complex cultural constellations.

The present research integrates an ideologically oriented, close reading approach to character relations with a data-driven, empirically informed approach using social network analysis on a sample corpus of 170 recent Dutch novels containing 2136 characters. For all those characters extensive demographic information (gender, age, place of birth, place of residency, level of education, profession) has been gathered in earlier research (Van der Deijl, Pieterse, Prinse, Smeets 2016), as well as thematic roles (family, love, colleague, friend, enemy) between all characters.

## Methodology

In order to discern relational patterns between characters in present-day Dutch literary fiction, key elements of character relations are formalised and quantified. The building blocks of a network are ‘nodes’ and ‘edges’: the first are represented by the characters in the novels, the second by the relations/interactions between those characters. This necessitates a formalisation of the concepts of 1) ‘characters’ and 2) ‘character relations’:

- 1) ‘Characters’ are defined as ‘people or creatures which to a greater or lesser extent are presented as human, existing of not more than a few linguistic features among which one or more names’.

As the issue of detecting characters in literary texts is complex (Vala et al 2015), the most pragmatic approximation is used. A list of character names is generated using Named Entity Recognition (NER). Characters names of which the frequencies are above a minimal threshold value are regarded as characters. Name variants of each character name are replaced in the text of the novels by a name-ID.

- 2) ‘Character relations’ are defined as ‘co-occurrence of character names in a window of N-words’.

The detection of character relations is customised on the basis of the narrative mode of each novel. The corpus has therefore been divided into novels per narrative mode, for each narrative mode another approach is used to measure the weight of the character relations:

- 1) *First-person narrated novels*. For novels with a single first person narrator, the weight of the relations of the first person narrator to the other characters is measured by occurrence frequencies of all the name variants of those characters in the novel as whole, as the whole novel can be regarded as the perspective of the ‘I narrator’. The weight of mutual relations between all the characters which are not the first person narrator are measured by co-occurrence frequencies in shifting windows of N-words.
- 2) *Third-person narrated novels*. For novels narrated from a third person perspective, the weight of the mutual relations between all the characters are measured by co-occurrence frequencies in shifting windows of N-words.
- 3) *Multiple perspective novels*. For novels narrated from multiple perspectives, the different perspectives are categorised as either first person or third person, after which the approaches of either 1 or 2 are applied to that specific perspective.

We will demonstrate the merits and pitfalls of these methodological choices by zooming in on three novels, one from each category.

## **Results**

In this talk, we will present and evaluate the results of our approach to model character relations by exploring a range of generic questions for the corpus, such as:

- 1) How strong are relations between characters of a certain demographic profile? E.g. relations between Western characters and non-Western characters, relation between old characters and young characters, relations between highly educated male characters and lowly educated female characters, et cetera.
- 2) How strong are relations between characters that pertain to a certain thematic role? E.g. relations between friends, between colleagues, between fathers and sons, between mothers and daughters, et cetera.

In order to answer these questions the weight of the character relations in all novels are aggregated. By transcending the natural confines of each novel, character relations can be analysed and interpreted on a larger scale than is usually done in close reading approaches to character representations. Statistical measures as degree centrality, betweenness centrality and closeness centrality are applied to study centrality of certain relation types in the corpus. The results suggest that certain relational patterns are most dominant in present-day Dutch literary fiction as opposed to others. We will critically reflect on how those patterns can be interpreted in light of the long standing research tradition on character representations.

Keywords: *social network analysis, character networks, Digital Literary Studies, Dutch literature*

## References

Agarwal e.a. 2013 – A. Agarwal e.a., ‘Automatic Extraction of Social Networks from Literary Text: A Case Study on Alice in Wonderland’. Online, 2013.

[http://www.cs.columbia.edu/~apoorv/Homepage/Publications\\_files/MAIN.pdf](http://www.cs.columbia.edu/~apoorv/Homepage/Publications_files/MAIN.pdf)

Alberich e.a. 2002 – R. Alberich e.a., ‘Marvel Universe looks almost like a real social network.’ Online, 2002. <https://arxiv.org/abs/cond-mat/0202174v1>

Deijl, Lucas van der, Pieterse, Saskia, Prinse, Marion & Smeets, Roel. 2016. ‘Mapping the Demographic Landscape of Characters in Recent Dutch Prose: A Quantitative Approach to Literary Representation.’ In: *Journal of Dutch Literature* (7:1).

Elson, David, Dames, Nicholas & McKeown, Kathleen. 2010. ‘Extracting Social Networks from Literary Fiction’. In: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (ACL 2010), Uppsala.

Fatima, Aziz, Jamil, Anil & Hanif, Sidrah. ‘Othering of Africans in European Literature: A Postcolonial Analysis of Conrad’s “Heart of Darkness”.’ In: *European Journal of English Language and Literature Studies* (3.5): 2015, pp. 40-45.

Jayannavar e.a. 2015 – P. Jayannavar e.a., ‘Validating literary theories using automatic social network extraction’. In: *Proceedings of the Fourth Workshop on Computational Linguistics for Literature* (2015), p.32-41.

Karsdorp, Folgert, Kranenburg, Peter van, Meder, Theo & Antal Van den Bosch. 2012. ‘Casting a spell: Identification and ranking of actors in folktales.’ In: F. Mambrini, M. Passarotti, and C. Sporleder (eds.), *Proceedings of the Second Workshop on Annotation of Corpora for Research in the Humanities* (ACRH-2), pp. 39–50.

Karsdorp, Folgert, Kestemont, Mike, Schöch, Christof, & Bosch, Antal van den. 2015. 'The Love Equation: Computational Modeling of Romantic Relationships in French Classical Drama.' In: *Proceedings of the Sixth International Workshop on Computational Models of Narrative*, pp. 98-107

Lee & Yeung 2012 – J. Lee & C.Y. Yeung, 'Extracting Networks of People and Places from Literary Texts'. In: *Proceedings of the 26th Pacific Asia Conference on Language, Information and Computation 2012* (2012), p.209-218.

Lee & Wong 2016 – J. Lee & T.S. Wong, 'Hierarchy of Characters in the Chinese Buddhist Canon'. In *Proceedings of the Twenty-Ninth International Florida Artificial Intelligence Research Society Conference*,(2016), p.531-534.

Meijer, Maaïke. 'De verschrikkelijke sneeuwman: projectie, geweld en nieuwe mannelijkheid in het werk van Jan Wolkers'. In: *Het omstreden slachtoffer: geweld van vrouwen en mannen* (R. Romkens & S. Dijkstra, eds.). Baarn (Ambo): 1996, pp. 39-58.

Minnaard, Liesbeth. 2010. 'The Spectacle of an Intercultural Love Affair: Exoticism in Van Deysse's *Blank en geel*'. In: *Journal of Dutch Literature* (1:1).

Moretti, Franco. 2013. *Distant Reading*. London: Verso.

Murfin, Ross C. (ed.) *Heart of Darkness (Case Studies in Contemporary Criticism)*. Boston/New York (Bedford/ St. Martin's): 1996.

Mushtaq, Hammad. 'Othering, Stereotyping and Hybridity in Fiction: A Postcolonial Analysis of Conrad's *Heart of Darkness* (1899) and Coetzee's *Waiting for the Barbarians* (1980).' In: *Journal of Language and Literature* (3): 2010, pp. 25-33.

Pattynama, Pamela. 'Oorden en woorden. Over rassenvermenging, interetniciteit, en een Indisch meisje.' In: *Tijdschrift voor Vrouwenstudies* (15.1): 1994, pp. 30-45.

Pattynama, Pamela. 'Secrets and Danger: Interracial Sexuality in Louis Couperus's *The Hidden Force* and Dutch Colonial Culture around 1900.' In: *Domesticating the Empire. Race, Gender, and Family Life in French and Dutch Colonialism* (Julia Clancy-Smith & Frances Gouda, eds.). Charlottesville/London (University Press of Virginia): 1998, pp. 84-107.

Sidi-Said, Fadhila. *Cross-cultural and Ideological Representations of the "Other" in Selected Works by Herman Melville and Joseph Conrad*. Mouloud Mammeri University of Tizi-Ouzou (Algeria): 2010.

Song, Angeline M.G. 2015. *A Postcolonial Woman's Encounter With Moses and Miriam*. New York: Palgrave Macmillan US.

Stiller e.a. 2003 – J. Stiller e.a. 'The Small World of Shakespeare's Plays.' In: *Hum Nat* 14 (2003) 4, p. 397-408.

Vala, Hardik, Jurgens, David, Piper, Andrew & Ruths, Derek. 2015. 'Mr. Bennet, his coachman, and the Archbishop walk into a bar but only one of them gets recognized: On the difficulty of detecting characters in literary texts.' In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 769–774, Lisbon, Portugal, Association for Computational Linguistics.