Character Constellations
Smeets, Roel

Published by Leuven University Press

Smeets, Roel.


For additional information about this book
https://muse.jhu.edu/book/98048

For content related to this chapter
https://muse.jhu.edu/related_content?type=book&id=3191738
CHAPTER 3

CENTRALITY

3.1 INTRODUCTION: NARRATIVE CORNERSTONES

How does the centrality of characters co-shape the representation of the social group(s) characters function in? In this chapter, the literary representation of social groups will be studied through the concept of centrality. Centrality will be used as an umbrella term to refer to abstract notions such as importance, dominance, influence, and power. A central character is important, dominant, influential, or powerful in one way or another. The term centrality will be operationalized through both a narratological and a network theoretical approach. By interconnecting these two seemingly distinct methodological traditions, a model is developed to pinpoint what it means for a character to be central in a narrative structure. As such, this chapter aims to show how the cross-fertilization between the methodological toolkits of narratology and social network analysis contributes to a better understanding of the centrality of characters belonging to a certain social group in present-day Dutch language fiction.

First, it will be discussed how centrality is commonly defined and operationalized in both network theory and narratology. On the basis of that discussion, a method is developed to extract fictional social networks of characters from each of the 170 texts in the corpus. Building on the descriptive statistics of the population of 2,137 characters in the dataset as reported in chapter 2, a hypothesis is formulated as to which types of characters occupy central positions in each of the extracted fictional social networks. A data-driven, statistical model
is then developed to rank characters in each individual novel according to five common centrality metrics. Then, a multiple regression analysis is carried out to test which demographic categories predict a character’s place in the rankings. Finally, the statistical pattern resulting from this regression analysis is discussed in light of a close reading of Özcan Aykol’s *Eus* (2012), which is qualitatively assessed to determine how its depiction of female and migrant characters relates to the centrality of these types of characters in the corpus as a whole.

### 3.2 CENTRALITY IN NETWORK THEORY

The Russian-American mathematician and psychologist Anatol Rapoport is commonly regarded as one of the pioneers of social network analysis. In the 1950s, he voiced one of its central premises by pointing at the ‘well-known fact that the likely contacts of two individuals who are closely acquainted tend to be more overlapping than those of two arbitrarily selected individuals’ (Rapoport, 1954, as cited in Leinhardt, 1977, p. 75). Individuals, in other words, function in social networks in which some individuals are more closely related than others.

What is a network? Most generally, it can be defined as ‘a pattern of interconnections among a set of things’ (Easley & Kleinberg, 2010, p. 1). In network theory, those ‘things’ are commonly called ‘nodes’ and can consist of virtually anything: people of flesh and blood, molecules, trains, computers, Facebook profiles, commercial products, academic articles, cities, and fictional characters in literature or film. The term ‘edges’ is used to refer to the ‘interconnections’ between the nodes and can denote a variety of relations: cooperation, co-occurrence, affiliation, et cetera.

One of the founding articles of social network theory is ‘The Strength of Weak Ties’ (1973) by the American sociologist Mark Granovetter. Its influence is demonstrated by its citation score on Google Scholar: in a period of 25 years, it has been cited 47,761 times. By stressing the difference between strong and weak connections between individuals, Granovetter laid the foundation for viewing network relations in terms of their strength. Whereas it might seem obvious that strong relations between family or friends are important for individuals, Granovetter has emphasized that weak relations (e.g., between acquaintances) can have a significant influence on social cohesion as well (Granovetter, 1973). Of similar importance for the development of social network theory is Milgram’s small-world experiment, in which the average distance between all inhabitants
of the United States was shown to be only five or six (Milgram, 1967). Although Milgram never used the term ‘six degrees of separation’, this led to the now common knowledge that everyone is connected with everyone else through a maximum of six steps, which suggests that the social world is extremely connected.

One of the core assumptions of network theory is that the relations between actors in a network affect the relative importance of those actors. Network theory has therefore been occupied with the question of how to measure the centrality of nodes in a network. The centrality of a node can be measured in a number of ways to consider different aspects of the network structure. In 1978 the American sociologist Linton Freeman observed that there is ‘certainly no unanimity on exactly what centrality is or on its conceptual foundations, and there is very little agreement on the proper procedure for its measurement’ (Freeman, 1978, p. 217). He conceptualized three basic centrality measures – degree, betweenness, and closeness – which are still being used today, albeit frequently in revised form, and which are thought to ‘cover the intuitive range of the concept of centrality’ (p. 237). It is worth mentioning that Freeman’s intent was not ‘to “lock in” to any sort of ultimate centrality measure’ (p. 217), as centrality is a rather abstract concept and therefore hard to pinpoint statistically. Existing measures as those proposed by Freeman at best help to clarify what might be understood as central, but they do not necessarily give any definitive answers on which actors are most important in a network.

Before Freeman’s innovation, centrality was mainly viewed in terms of degree. In Figure 1, node A has an advantage over B, C, D, and E because it has more relations to others in the network: A has a degree of 4, B, C, D, and E have a degree of 1. The main limitation of degree centrality, however, is that it does not consider the overall structure of the network. A node can be related to many other nodes but located in the periphery of the network, which results in a situation where the node is far removed from the opposite side of the network.

As an alternative to degree centrality, closeness centrality is defined as the sum of distances to all other nodes in the network. An advantage of closeness is that it accounts for the relative access that a node has to other nodes in the network. In Figure 1, node A has a higher closeness than B, C, D, and E, as it is directly connected with its neighbors, whereas B, C, D, and E need to cross through A to reach a node other than A. The disadvantage of closeness centrality, however, is that it cannot properly be applied to networks that are not fully connected. By definition, nodes in two disconnected components of a network are unable to reach one another, and therefore closeness cannot be computed for the overall structure of a network with disconnected components.
Freeman was the first to propose *betweenness centrality*, which computes the extent to which a node lies on the shortest path between two other nodes. In Figure 1, node A has a high betweenness centrality because it connects all four nodes with each other. As it is applicable to networks with disconnected components, betweenness has an advantage over closeness. However, as a metric, it is limited because nodes are often not located on the shortest path between any two other nodes. Because of that, B, C, D, and E in Figure 1 all have a betweenness centrality of 0.

In some networks, edges between nodes have the same status. For instance, networks of Facebook friends, in which the nodes are people on Facebook who are connected by virtue of being Facebook friends, features binary edges: a Facebook profile either is or is not befriended with another Facebook profile. This is essential rather than gradual: such edges have a weight of 1 as there is no spectrum on which the relation can be positioned. Compared to real-world friendships, this is of course a highly reductive representation of affairs. In a circle of friends, not every person is befriended with everyone else in exactly the same way. Jan might have a closer bond with Piet than with Marie, whereas Piet and Marie can share childhood memories that strongly connect the two of them.
together. Viewing such a circle of friends as if it were a Facebook network distorts this gradual spectrum on which the relational structures exist. Specific weights have to be added between every two nodes: a weight of, say, 5 might be ascribed to the relation between Jan and Piet (their connection is moderately strong), the relation between Marie and Piet might have a weight of 10 (because of their childhood connection), and the relation between Jan and Marie a weight of 2 (they only see each other at Piet’s birthday parties).

In order to account for this, network theory makes a distinction between unweighted and weighted graphs. In a weighted graph, the edges represent the intensity with which two nodes are connected. As the basic centrality measures of degree, closeness, and betweenness are devised for application to unweighted, binary networks, alternative metrics have been proposed. Degree centrality has been redefined for weighted graphs by focusing not on the number of relations but on the sum of the weights of those relations (Barrat, Barthélemy, Pastor-Satorras, & Vespignani, 2004). Dijkstra's algorithm (Dijkstra, 1959), named after the Dutch computer scientist Edsger W. Dijkstra, has been used to redefine closeness and betweenness centrality by looking at the shortest paths in terms of distances (Brandes, 2001; Newman, 2001). As these new proposed metrics target primarily the weights and are less reliant on the number of relations, a second redefinition was needed to take into account both weight and number of relations (Opsahl et al., 2010).

Every network thus demands a specific approach; there is no general method that applies to every network. The first question is which elements constitute the network, the second how those elements are related. Then, it should be decided if the network is binary and unweighted, or if the elements are gradually related to one another. The appropriate centrality metrics should be derived from the specific nature of the network (weighted/unweighted, unipartite/ bipartite) and the question through which it is approached, as not every centrality metric is relevant in all possible instances. There are cases in which degree centrality is most insightful, such as in the earlier mentioned binary, unweighted Facebook network. The Facebook profile with the most connections to other Facebook profiles is arguably a highly central actor in the overall Facebook network – the more connections, the more access to information on Facebook.

Conversely, in a weighted, real-world network of friends, degree might actually not be a good indication of someone’s importance in a circle of friends. Marie is connected to 35 other people, but the intensity of the larger part of those connections is very low (with a weight of only 1 or 2); only with Piet does Marie have an intense relation (with a weight of 10). Furthermore, the 35
people Marie is connected to all live in the same village, which makes her circle of friends geographically restricted. Piet, on the other hand, is befriended with only 10 people, but he has very strong connections with all of them, and they all live in different cities. For that reason, Piet’s circle of friends is both very strongly connected and geographically widely distributed. Degree centrality is in this case a less suitable indication of centrality, as it only focuses on the number of relations but not on the exact position in the network. Betweenness centrality might be more suitable in this case, as it is able to differentiate between nodes that are able to bridge different, dislocated parts of the overall network. Having a variety of friends in different cities, Piet is a so-called ‘broker’: he functions as a mediator, a bridge, between remote circles of friends in different places.

Whereas network theory provides tools to compute the centrality of nodes in a network based on statistical metrics, narratology offers insights into the ways in which characters occupy more or less central positions in a narrative structure, which is described in the section below.

### 3.3 CENTRALITY IN NARRATOLOGY

Narratology offers different instruments to analyze the centrality of characters in narrative fiction, of which this section will mention two of the most straightforward. A character’s position in the storyworld is already predetermined by some basic structural features of a literary text. The mode of narration is commonly a first indicator of how important a character is in the storyline. Handbooks of literature train first-year students to be aware of the embeddedness of certain narrative situations. Illustrative is the following introduction to the analysis of narrative texts from a frequently used handbook in Dutch literature departments:

Epic or narrative texts are characterized by an embedded language situation. They mainly pivot on spokesmen that enter into a dialogue with one another, just as in drama. These are the characters. But this dialogical situation is embedded in a textual frame that is produced by a narrating instance. This instance produces text that is not perceived by the characters. Therefore, the narrating instance is located on a higher textual level. He can tell something about the characters; he can represent their words directly by creating dialogues; he can also
summarize their conversations; he can even represent their thoughts, directly or indirectly. The narrating instance is thus above the world of the characters, he has insights into the world of the characters and he reports information on that world in the text.⁸ (Van Boven & Dorleijn, 2013, p. 33, my emphasis)

Although terms such as centrality, importance, or power are not explicitly mentioned, a hierarchical relation is posited between the narrating instance and the characters: the first ‘is located on a higher textual level’ and is ‘above the world of the characters’. The idea that a narrating instance is located at the top of hierarchically embedded narrative layers is an axiom of narratology. From a network theoretical perspective, this makes sense: the narrator is the one who controls the flow of information in a narrative and therefore occupies a key role in the depiction of events and description of characters. The main insight is that narrating characters are not on an equal footing with non-narrating characters. Mode of narration can therefore be taken as a point of departure for the study of the centrality of characters.

Another concept suited for the study of character centrality is focalization, which was coined by the French structuralist Gerard Genette to distinguish between who narrates and who perceives in a text (1980). Others have suggested revisions of the concept (e.g. Bal, 1977; Jahn, 1996; Nelles, 1990); the revision that has become most popular is that of the Dutch scholar Mieke Bal. She defines focalization as ‘the relation between the vision and that which is “seen”, perceived’ (Bal, 2009, pp. 145–146). An important difference with Genette’s use of the term is that Bal’s definition is able to discriminate between a focalizing subject (the one who perceives) and a focalized object (the one who is perceived). In this definition, focalization makes it possible to discern hierarchical relations between characters who occupy active focalizing roles and characters who are mainly in a passive position in which they are being focalized by other characters. The extent to which a character features in active focalizing roles is thus another indicator of its place in the character hierarchy.

Similar to the relation between the narrator and the narrated, the focalizer is in a hierarchical relation with the focalized:

If the focalizer coincides with the character, that character will have an advantage over the other characters. The reader watches with the character’s eyes and will, in principle, be inclined to accept the vision presented by that character. (Bal, 2009, pp. 149–150)
Even more than with the narrating instance, the focalizing instance has a major influence on the reader’s perception of the narrative. A character’s perception of an event or of another character usually goes hand in hand with a value judgment (cf. Jouve, 2001). The vision that is presented is not neutral but colored by a character’s disposition toward an event or fellow character. More importantly, the character who is being perceived is not in a position to put forward his or her own vision on the state of affairs. As such, the perceived instance, the focalized object, is subjected to a process of objectification. There is, therefore, an inherent power imbalance between the focalizer and the focalized, which is of particular relevance for a qualitative assessment of which social groups are depicted as more or less central in literary texts.

It is noteworthy that narrators and focalizer sometimes coincide in texts. This is typically the case for novels that are narrated from a first-person perspective. In such novels, the first-person narrator is usually part of the world of characters. Often, the first-person narrating character is the main focalizing instance in the story: the chain of events is presented through his or her vision, while the vision of other characters is mostly presented indirectly. In third-person narratives, the narrating instance is usually anonymous and not part of the world of characters. Although such anonymous narrators are sometimes also focalizing other characters, it does not make much sense to frame this as a power imbalance between the anonymous narrator and the characters. This is because the narrator is only connected to the characters on a meta level but is not part of the fictional social network of characters as such. As a concept for the analysis of character centrality in third-person narrated novels, focalization is thus primarily suited to analyze characters that are part of the fictional social network.

As the basic features of mode of narration and focalization affect the position of characters in the narrative, they are taken into account in this book’s method for extracting fictional social networks of characters from the 170 texts in the corpus. The corpus is divided into three sub-corpora based on their mode of narration and/or focalization: first-person narrated novels (63), third-person narrated novels (73), and multi-perspective novels (34). The next section describes in detail the method developed for each sub-corpus individually.
3.4 METHOD FOR EXTRACTING CHARACTER NETWORKS

In this section, a method is described to extract fictional social networks of characters from each of the 170 texts in the corpus.

3.4.1 Characters as Nodes

Each novel in the corpus can be regarded as a network with characters as nodes and relations between those characters as edges. In character studies, multiple definitions of the concept of characters circulate (see the introductory chapter, section 1.1.3). This book follows the most straightforward definition of The Living Handbook of Narratology as ‘a text- or media-based figure in a storyworld, usually human or human-like’ (Jannidis, 2013). In order to automatically extract character networks from the novels, characters first have to be detected in the texts, which requires a formalization of the concept of character. Following Van Boven and Dorleijn’s definition of characters as ‘people or creatures which to a greater or lesser extent are presented as human, existing of not more than a few linguistic features’ (2013, p. 335), the challenge is to define which linguistic features are essential characteristics of characters. The most eye-catching linguistic feature of a character is commonly its name, although not every character bears one. But a character is usually referred to not exclusively through their name but also through pronouns (‘he’, ‘she’, ‘I’) and coreferents (‘the man in the alley’, ‘the one who has been chosen’, ‘the mother of the child’). To this date, coreference resolution is an unresolved problem in Natural Language Processing (e.g. Clark & Manning, 2016). Previous studies (Vala, Jurgens, Piper, & Ruths, 2015) have shown that automatic detection of characters is difficult due to the poor performance of existing coreference resolution techniques. Because of this poor performance, the present study does not aim for full coreference resolution but instead uses a semi-automatic method that builds on a predefined set of characters. Building on Van Boven and Dorleijn’s formal definition (2013, p. 335), characters are defined here as people or creatures which to a greater or lesser extent are presented as human, existing of not more than a few linguistic features including one or more names.

For each novel, a list of names is generated with Named Entity Recognition (NER); characters whose name frequency is above a normalized threshold value (based on the number of words of the text) will be regarded as characters. With this most viable approximation of a character’s presence in a text, the
detection of characters in the present study is thus restricted to occurrences of name variants of each character. Lists of all variants of a character’s name are generated and stored in a database called NAMES, which corresponds to the databases NODES, EDGES, and BOOKS (see Figure 2). BOOKS contains all relevant metadata of the novels, such as title, the name, gender, and age of the author, the publisher, and the filename of the digital version of the novel. NODES contains all relevant metadata of the characters, such as name, gender, country of descent, city of descent, country of residency, city of residency, education, and profession. EDGES contains all relevant metadata on the character relations, such as the specific nature of the relation (friend, family, enemy, lover, professional). NAMES contains all variants of a character’s name. All databases are linked to one another through a unique book id. NAMES, NODES, and EDGES are also connected through a unique character id. Based on these interlinked databases, the character networks are computed through an Object-Oriented model written in the Python programming language, consisting of three main classes: Book, Character, and Network (Smeets & Sanders, 2018).\(^\text{14}\)

![Figure 2. Visualization of database linkage.](image)

Each book in the corpus has a unique id ranging from 1 to 170. Every character in the corpus has a unique id that corresponds to this book id stored in database BOOKS. For instance, *De lichtekooi van Loven* by Ineke van der Aa is represented by the book id 1. In database NODES, character ‘Louise’ is represented by this same book id followed by character id 1 and her name (1_1_Louise). In database NAMES, this same unique identifier is followed by every name variant of the character. The name variants for this character are ‘Louise,’ ‘Louisje’ and ‘Louiseke’, which is represented in NAMES as 1_1_Louise Louise, 1_1_Louise_Louisje and 1_1_Louise_Louiseke. Each novel’s text is then automatically searched for every of these name variants, after which these variants are replaced by the unique character identifier.\(^\text{15}\) Figure 3 shows a piece of text from *De lichtekooi van Loven*,

---


\(^{15}\) Smeets, L., & Sanders, J. (2018). *Character Constellation: A network approach to character dynamics*.
in which name variants ‘Lili’ and ‘Louisje’ occur. Figures 3 and 4 show how these aliases are replaced by a unique character identifier.

As such, each text is marked with character identifiers representing the occurrences of each character. These markers are then used to map interactions between characters.

Figure 3. Original text snippet from De lichtekooi van Loven by Ineke van der Aa.

Figure 4. Text snippet from De lichtekooi van Loven by Ineke van der Aa, in which name variants are automatically replaced by unique character identifiers.
3.4.2 Character Relations as Edges

In networks consisting of fictional characters, nodes are perhaps quite obviously represented by the characters in the text, but it is less obvious how edges between these nodes are constituted. Earlier research on character networks differed in their approaches. One of the most used definitions of character relations frames connections between characters in terms of conversations or dialogues (Elson, Dames, & McKeown, 2010; Jayannavar, Agarwal, Ju, & Rambow, 2015; Lee & Yeung, 2012; Lee & Wong, 2016; Moretti, 2013; Stiller et al., 2003). The quantifiable unit of the conversation is, however, not the best indication for character interactions, as there are plenty of characters that do not enter into a conversation but are related to one another in some other way. For instance, two characters with family ties might never speak to each other, but such a relation should definitely be regarded as a character relation. Another way to define relational ties is in terms of co-occurrence in the same window of N words, sentences, paragraphs, or chapters (Alberich, Miro-Julia, & Rossello, 2002; Grayson, Wade, Meaney, & Greene, 2016). Defining character relations in terms of adjacency in the text will be able to capture more instances of character interaction than when it is defined in conversational terms. This is the most bottom-up definition of character relations, as characters do not have to communicate in a literal sense (as is the case in conversation networks) to be considered as having a form of interaction. Character relations are defined here as co-occurrences of character name variants in a window of N tokens.

This definition is based on the assumption that the strength of a character relation increases when a character occurs more often near another character. Nearness is defined here as characters occurring in the same reach, referred to as a window consisting of a specific amount (= N) of for instance words, sentences, paragraphs, or chapters (= tokens). Experiments were conducted with different window units and sizes for different types of novels to find the ‘sweet spot’ where not too many and not too few character interactions are detected (cf. Grayson et al., 2016). However, such a sweet spot is different for every novel. In order to be able to compare the novels, the same window unit and window size for every novel was used. As sentences are the smallest linguistic structures which are semantically meaningful in themselves (cf. Mann, William & Thompson, Sandra, 1988), sentences were as the window unit, which were tokenized using the Ucto software. The window size was set to two sentences, as semantic relations are known to extend over two sentences through connectives (cf. Blühdorn, 2010). A customized co-occurrence approach for each mode of narration was developed
(first-person, third-person, multi-perspective), which is described in detail in section 3.4.3.

However, character co-occurrence does not capture any thematic relations such as family or friend. For that reason, the dataset was manually enriched by two student assistants who gathered extensive relational information. Among all 2,137 characters identified in the corpus, the following thematic roles were annotated: friend, lover, colleague, enemy, family (specified through all possible subcategories such as mother, son, brother, grandmother, et cetera). 8,732 of these roles were stored in the EDGES database. As it was not always evident which labels apply for a relation between two characters, interpretive deduction was therefore sometimes unavoidable. In general, the roles were defined as narrowly as possible. Colleague, for instance, was used for every two characters who had a professional relation of some kind, whereas enemy was used for characters who were clearly hostile to one another. Note that ‘professional’ and ‘hostile’ are not objective categories but require interpretation. Changing relations between characters were also accounted for. In those cases double labels were assigned, such as Colleague_Enemy. Double labels were also assigned when the nature of the relation changed over time, such as friends becoming enemies.

3.4.3 Automatic Extraction of Character Networks

For each sub-corpus (first-person, third-person, multi-perspective) a slightly different co-occurrence approach was developed based on the specific mode of narration. Third-person novels are narrated by an anonymous narrator who follows one main character. First-person novels are narrated by an I-narrator. Multi-perspective novels are narrated by multiple narrators, either in third or first-person. For all novels, irrespective of their mode of narration, relations between characters are preestablished when they are annotated with one of the relational labels that are stored in EDGES (friend, family, lover, enemy, colleague). In all cases, the procedure below is used to establish the weight of the relations.

1. Third-person narrated novels [63 novels]
For every character in the novel, a sliding window approach is used in which co-occurrences of two characters are mapped in a window of two sentences. Whenever two characters occur in the range of the same two sentences, a relation between those characters is established. The more often such a co-occurrence takes place, the stronger their relation becomes.
2. First-person narrated novels [73 novels]
   a. As the first-person narrator has a priori high centrality in narratological terms, the relation of the first-person narrator with all other characters are simply defined by counting every occurrence in the novel of characters other than the first-person narrator. As every character is embedded in the narration of the first-person narrator, it can be argued that every character occurrence represents a relational tie with the first-person narrator. The more often a character occurs in the novel, the stronger its relation with the first-person narrator is.
   b. For every character other than the first-person narrator, a sliding window approach is used in which co-occurrences of two characters are mapped in a window of two sentences. Whenever two characters occur in the range of the same two sentences, a relation between those characters is established. The more often such a co-occurrence takes place, the stronger their relation becomes. Note that this approach will in most cases rightfully lead to relatively strong relations between the first-person narrator and all other characters, whereas this is not the case for the relations between and among all other characters.

3. Multiple perspective novels [34 novels]
   A student assistant annotated for each of these novels where a character perspective begins and ends in the text. These annotations also contain information on the narrative mode and focalization: a first-person or third-person narration was annotated as such, and for third-person narration the main focalizer was annotated. On the basis of those annotations, each novel was divided in separate sections. For sections narrated respectively from first- or third-person, the first- or third-person method was applied. After that, the co-occurrence counts between characters were aggregated for all the separate sections.

All these relations are symmetrical and thus undirected. This means that the character relations are not regarded in terms of directionality, which is a logical consequence of the co-occurrence approach, as adjacency is an a priori symmetrical issue. Furthermore, the resulting networks, with characters as nodes and character relations as edges, are both undirected and weighted. Not every relation between any two characters will have the same status as the strength of a relation is increased when two characters occur more often in the same window.
Of major importance is the definition of ‘window’ that is used in the sliding window approach. The goal is to automatically compartmentalize the narrative in order to detect character interaction in a precisely delineated context. This delineation can be done on the basis of three quantifiable linguistic units, from small to large:

1. characters (i.e., the letters of the alphabet and symbols, and not in the sense of ‘fictional characters’)
2. words
3. sentences

Which of these linguistic units is most suited for mapping character interaction depends on linguistic-semantical considerations. The smallest building blocks of a text are characters, the symbols that represent a number or a letter. These are not semantically meaningful in themselves but gain their significance only when formed into words, which are the second smallest building blocks of the text. Words carry meaning in different ways: function words (‘the’, ‘a’) only perform a grammatical function in a text, whereas content words (‘animal’, ‘person’) carry meaning in themselves.

One can argue that words only become meaningful in relation to the syntactical structure in which they are embedded. A sentence is commonly regarded as the smallest syntactical structure that is meaningful as a closed-off system in itself (cf. Mann et al., 1988). Furthermore, as well as the linguistic units used, the size of this window (of characters, words, or sentences) will greatly influence the results (Grayson et al., 2016; Wade & Grayson, 2016; Zadeh & Handschuh, 2014). A window size too small will capture too few character interactions, whereas a window size too large will capture too many. The smallest possible window size would be 1 (character, word, or sentence). Essentially no interaction takes place in a window size of 1 based on characters or words. When the unit of a window is a sentence, interaction is possible: two characters can occur in one sentence. The biggest window size would be as big as the total number of units (characters, words, or sentences) in the novel. A window this big would lead to nonsensical results as every character will be connected to every other character as often as they occur in the novel. Therefore, it is essential to find a so called ‘sweet spot’: a window size of N tokens (characters, words, or sentences) that leads to a network that is not too small or too large, but reasonably reflects the intensity of interaction between any two characters.
Although this methodological problem could have been approached in a bottom-up, data-driven way, it was tackled top-down by taking into account linguistic theoretical considerations. As mentioned above, sentences are the smallest syntactical structures in which the linguistic elements are semantically related to another (cf. Blühdorn, 2010). Whenever two characters occur in the same sentence, one can therefore be sure that they are somehow related to one another. This is not necessarily the case for two characters appearing in the same window of N characters or words, as those characters or words are not always part of an overarching semantic framework. For this reason, the window unit was set to sentences. All novels from the corpus are thus split into sentences by tokenizing the texts with the Ucto software.

Furthermore, linguists have shown that semantic connections in language are formed not only in a sentence but also between two adjacent sentences (cf. Blühdorn, 2010), for instance through connectives in the form of conjunctions (as, and, but, if, or), prepositions (at, by, in, to), relative pronouns (who, which, what, that), conjunctive or relative adverbs (hence, when, whence, where, why). It can therefore be argued that characters occurring in the same two sentences have a relational tie. Because of these considerations the window size was set to two sentences. Interestingly, windows defined in this way come close to the window size used by Grayson et al. (2016). Although they use words instead of sentences, the sizes are similar. Grayson et al. (2016) use varying window size that are not much smaller or bigger than around 50 words, which approximates the number of words in two sentences, considering that the average sentence has around 15–20 words.

Figure 5 shows an example window of two sentences in which two characters, represented by their unique identifier, occur. Whenever the algorithm finds co-occurrences of two characters in this context, the weighted relation between characters X and Y is incremented by 1. The generated weights are subsequently used to compute the degree, betweenness, closeness, and other centrality metrics for each character in the corpus.

Figure 5. A window of two sentences in De lichtekooi van Loven by Ineke van der Aa.

In order to determine the extent to which this approach to character network extraction matches up human intuitions, a sample of extracted character networks
Centrality was qualitatively assessed. For each of these cases, the extracted character networks were compared with a reading of the novel. In all instances, the characters that seemed central units in the narrative also ended up as central characters in the character networks, while the side characters ended up in more peripheral positions. This cursory validation suggests that the presented approach to character network extraction at least conforms to common sense perceptions about protagonists, main characters, and side characters.

3.5 MODEL I: CHARACTER RANKINGS

For each of the 170 novels in the corpus a unipartite, undirected, weighted network is thus extracted based on the method explained in the previous section. With Python’s software package networkx, the resulting networks for each individual novel are used to rank the characters on the basis of five centrality metrics (see section 3.2 of this chapter for an overview of centrality metrics). These rankings provide a view on how central a certain character in the corpus is according to one of the centrality metrics. Among those metrics are the above-described degree, betweenness and closeness centrality, as well as eigenvector and Katz centrality, two metrics on which Google’s PageRank algorithm is based. PageRank is used by Google’s search engine to rank web pages by relevance. PageRank, eigenvector, and Katz are all based on the same, seemingly circular assumption that a node in a network becomes more important when it is connected to other important nodes (Page, Brin, Motwani, & Winograd, 1999). Unlike eigenvector centrality, Katz centrality tends to be more useful for networks that are not strongly connected. For Katz centrality, the default options were used as free parameters. The computation of all these metrics is based on the weighted edges.

3.5.1 Results Multiple Regression Analysis

A regression analysis was carried out to determine the extent to which the demographic variables (gender, descent, age, education) predict a character’s place in the rankings. Because of the exploratory nature of the present study and the absence of prior research on this topic, there were no strong suggestions for a hypothesis about which demographic factors would possibly determine a character’s place
in the rankings. However, it was preferable to not just enter all possible variables into the regression equation as this would possibly obscure the results of the analysis. Therefore, a nonformal hypothesis was formulated based on qualitative, nonstatistical research in the critique of literary representation. Several studies suggest that female characters and/or characters of mainly non-Western descent are often represented in a stereotypical manner and are therefore likely to be featured in less central, more marginal positions in literary texts (e.g. Meijer, 1996a, 1996b, 2011; Minnaard, 2010; Pattynama, 1994, 1998). Gender and descent are therefore possible predictors of a character’s position in the rankings. Based on descriptive statistics on the dataset (see chapter 2 ‘Data’), it can be suspected that male and nonmigrant characters will end up as more central, since these types of characters are simply more present in the dataset (see Figure 6). More precisely, it is hypothesized that both male characters and nonmigrant characters will score higher on the centrality metrics than female characters and characters with a migration background.

Figure 6. Gender and descent distributions among characters in the corpus (N=2,137). The numbers are based on types, not on tokens. Missing data are not displayed in this figure.
For each of the five centrality metrics (degree, betweenness, closeness, eigenvector, Katz), a multiple linear regression was conducted to predict characters’ centrality scores based on their gender and descent. Gender is coded as 0 for male and 1 for female. Descent is coded as 0 for nonmigrant and 1 for migrant. As the aim was to generalize over all novels, the division between sub-corpora (third-person, first-person, multi-perspective) is not included in the statistical model. This division was only used for the computation of the extraction of networks and does not need to be accounted for in the outcome of the regression model.

No significant results were found for betweenness, closeness, and eigenvector centrality. Gender and ethnicity are thus no predictors for characters’ scores on betweenness, closeness, and eigenvector centrality.

However, significant results were found for degree and Katz centrality. First, for degree centrality, a significant regression equation was found ($F(2, 2128) = 6.424, p < 0.01$), with an $R^2$ of 0.006. Characters’ predicted degree centrality is equal to a $B$ value of $0.428 + 0.024 \times \text{GENDER} + 0.059 \times \text{DESCENT}$ (see Table 1). This means that, on degree centrality (on a scale from 0 to 1), female characters scored 0.024 higher than male characters, and migrant characters scored 0.059 higher than nonmigrant characters.

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized coefficient</th>
<th>Standardised coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B: 0.438, Std. Error: 0.006, Beta: 0.065, Sig: 0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Descent revised: 0.058, Beta: 0.065, Sig: 0.003</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>B: 0.428, Std. Error: 0.008, Beta: 0.066, Sig: 0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Descent revised: 0.059, Beta: 0.066, Sig: 0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gender: 0.024, Beta: 0.043, Sig: 0.048</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1. Linear model of predictors of degree centrality.*

Secondly, for Katz centrality a significant regression equation was found ($F(2, 2128) = 6.124, p < 0.01$), with an $R^2$ of 0.006. Characters’ predicted Katz centrality is equal to a $B$ value of $0.272 + 0.009 \times \text{DESCENT} + 0.007 \times \text{GENDER}$ (see Table 2). This means that, on Katz centrality (on a scale from 0 to 1), migrant characters scored 0.009 higher than nonmigrant characters, and female characters scored 0.007 higher than male characters.
<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized coefficient</th>
<th>Standardised coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>0.273</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.007</td>
</tr>
<tr>
<td>2</td>
<td>(Constant)</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Descent revised</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Table 2. Linear model of predictors of Katz centrality.

These findings suggest that the initial hypothesis, based on qualitative critiques of literary representation, should be rejected. Contrary to what was expected, female characters and migrant characters scored higher, at least on two of the five centrality metrics used in the analysis. Based on these results, it can be argued that a higher frequency distribution of a character type does not necessarily lead to a more central position in a character network, as the results of the regression analysis have shown. Although male and migrant characters are more present in the corpus, they do not end up as more central in network analytic terms. These results invoke questions with regard to the notion of centrality in narrative fiction. While the frequency of occurrence of certain types of characters impacts their presence, their visibility, in narrative fiction, characters who occur relatively less frequently (and thus are less present, less visible) can still occupy central positions in their social networks. Does this mean that authors ascribe more central roles to social groups that are less present? In order to better understand what this finding means for centrality of characters in narrative fiction, the pattern is qualitatively assessed in the next section by close reading one novel from the corpus.

3.5.2 Close Reading: Centrality, Gender, and Descent in Özcan Akyol’s *Eus* (2012)

As it is unclear what the significance of the quantitative representational patterns generated by the regression model is for the critique of literary representation, a narratological evaluation of these patterns is warranted to pinpoint what these patterns mean at the level of the individual text. In concrete terms, the finding
for degree centrality is that female and migrant characters have significantly more relations than male and nonmigrant characters. More specifically, women and characters with a migration background often co-occur with a wider range of fellow characters in the novels. The higher scores of female and migrant characters on Katz centrality indicate that they often co-occur with characters who also have relatively high Katz centrality. In sum, female and migrant characters have both more relations in general and more relations with important characters.

In order to make sense of this pattern, a narratological exploration of character centrality in one novel from the corpus was conducted and confronted with the results of the statistical analysis. The data-driven model presented in the previous section generated a statistical baseline: the finding that female and migrant characters occupy relatively central positions in fictional social networks. Individual works can be compared to this baseline in order to determine the extent to which they conform to or deviate from it. This statistical baseline thus functions as a norm of representation in light of which this section’s case study was qualitatively assessed. As there is a field of tension between the usage of ‘centrality’ or ‘importance’ in network theory and its usage in narratology, special attention was paid to the various meanings of these terms in both disciplines and the potential conflicts between them.

As a case study, a novel was used in which both gender and descent are thematized, as these were also used as variables in the regression analysis. For the sake of clarity, only the two earlier mentioned basic concepts of narration and focalization were used as points of departure. Note that there is a wide variety of other narratological concepts and perspectives that might potentially lead to alternative insights. How do narration and focalization shape the centrality of female and migrant characters in this particular narrative? And how does the relative importance of these types of characters in this novel relate to the statistical baseline that female and migrant characters are relatively central in the corpus as a whole?

_Eus_ (2012) by Özcan Akyol is a semi-autobiographical first-person narrated novel, in which the reader follows the life of first-person narrator Eus. The novel clearly conforms to certain conventions of the (Dutch) picaresque novel (‘schelmenroman’) in the tradition of Jan Cremer’s _Ik, Jan Cremer_ (1964): Eus is a sly social outcast with loose sexual morals who gets involved in criminal activities. Unlike most novels in this genre, however, the migration background of Eus’s family plays a crucial role in his story. He is the son of Turkish migrants living in Deventer, a small city in the Netherlands. Because of his criminal activities he ends up in jail, where he starts a writing career. This plotline thus foregrounds the
theme of upward social mobility: a character with a migration background from the lower social classes, who initially has a hard time finding his way in Dutch society, eventually finds his creative ambition and becomes a successful author.

Oppositions between people with either Dutch or non-Dutch cultural backgrounds are thematized by Eus's emphasizing of the socioeconomic hierarchies that exist between these social groups. At the beginning of the novel, Eus states that he and his friends ‘didn’t dare to go to the better neighborhoods’, although they ‘knew that they existed’ (Akyol, 2012, p. 24). An implicit opposition is thus postulated between ‘better neighborhoods’ populated by Dutch, higher class people and Eus’s own, apparently lesser neighborhood inhabited by a wide range of people with a migration background. Later on in the novel, his characterization of the ‘indigenous youth, rich kids’ as ‘white scum’ (p. 120) explicitly shows that Eus expresses a negative disposition towards people in these ‘better neighborhoods’.

Another less prevalent, but latently present theme in the novel is the way in which men (with a migration background) engage with (Dutch) women. Throughout the novel, women are treated with little respect by Eus and his friends. Female characters are either object of sexual desire or considered a man’s possession. They are repeatedly referred to as ‘whore’ (pp. 36, 58, 85, 145) and variants on the term ‘slut’ (pp. 43, 57, 62, 86, 145, 157, 176, 253). The male characters seem mostly interested in whether or not a woman is ‘fuckable’ (p. 163). More generally, women tend to be sexualized by the men in this novel, of which the following quote is a clear illustration:

Sometimes I stared out of the window for hours, in search of the hottest girls in school, about whom I then started fantasizing. How beautiful they were! Nice tits! Nice ass! (p. 50)

On the basis of such thematic cues, one could argue that at least two binary oppositions take shape in the narrative: between characters with and without a migration background, and between male and female characters. The first binary opposition is anchored in the premise of the book: a street-smart boy from a family with a migration background fights his way up in the social hierarchy of Dutch society. The second binary opposition emerges from the sexualization and objectification of female bodies by Eus and his male friends. The binary oppositions between these social groups – characters with versus characters without a migration background, male versus female characters – will be taken as a point of departure in the analysis below. How to assess which of these social groups are represented as more or less central in this novel?
As was described in the introductory chapter (section 1.2.1 ‘Representation and ideology’), scholars such as Susan Suleiman (1983), Philippe Hamon (1984), Liesbeth Korthals Altes (1992), and Vincent Jouve (2001) have devised narratological models to trace the ‘ideology-effect’ or ‘value-effect’ in texts. Such models can help to pinpoint which characters are more central than others in the normative hierarchy of each single narrative structure. In *Authoritarian Fictions* (1983), Susan Suleiman grounds her analysis of the ideological dimensions of texts on a model that represents the different constituents and levels of narrative texts (pp. 156–157). This model breaks down the narrative text at the ‘Level of Story’ and the ‘Level of Discourse’. The first level of story relates to the components of the narrative content and contains events as experienced by characters in a specific context. The second level of discourse refers to how this narrative content is put into discourse (i.e., ‘the way in which the story is presented to the reader or listener’ (p. 156)), which is done through narration, focalization, and temporal organization.

For the present analysis, the processes of narration and focalization are particularly useful to determine which types of characters occupy more central positions than others in the narrative of *Eus*. A closer insight into its narration (‘who is telling the story, to whom, under what circumstances?’ (ibid.) and focalization (‘from whose perspective(s) is the story “seen” or experienced?’(ibid.)) helps to pinpoint the relative position of the characters in the hierarchy of values that the text communicates. More specifically, it helps to assess how central characters with a certain gender and descent are in this particular narrative.

First of all, the novel is narrated by Eus, which means that he controls the flow of information in the narrative. It is a logical consequence of the first-person narration that Eus decides which events to either report or leave out. When he, for instance, reports that ‘I was born and raised in Koekstad, a small town by the IJssel, exactly on the border of two eastern provinces’ (p. 13), he chooses to use an alias (‘Koekstad’) for a town which the reader might know as Deventer. As an I-narrator, Eus is thus able to manipulate the narrative at will. Furthermore, his specific position in the narrative structure enables Eus to perform one of the various functions Suleiman ascribes to narrators, which is the ‘interpretive function’: ‘to analyze, interpret, formulate judgements about characters, events, or contexts’ (p. 157). Unlike all of the other non-narrating characters in the narrative, Eus has the opportunity to evaluate and judge his fellow characters directly through his first-person narration.
Furthermore, he is also the main focalizer: the narrative events are filtered through his perceptions. This means that the description of events are not neutral but colored by the vision and judgement of Eus, which is exemplified by the following quote:

The coming four years I went to the Hegius school, amidst the beautiful, posh girls who followed the highest level of education. According to the rumors, these girls were above average interested in foreign boys because they never saw those types of boys. (p. 37)

Lumping together a group of girls, Eus ascribes the features ‘beautiful’ and ‘posh’ to these other non-narrating characters. By foregrounding their physical appearance and their alleged poshness, Eus suggests that they are spoiled rich kids, whose most interesting features are their looks. While the clause ‘According to the rumors’ suggests that the statement made in the next sentence should be taken with a grain of salt, Eus chooses to foreground the rumor that these girls are sexually interested in boys with a migration background. These two sentences contain an extremely colored representation of a specific type of character (in this case: female, higher educated). They ascribe features to characters that are not verifiable within the context of the first-person narration; readers can either take his word for it or adopt a critical stance toward Eus’s description of events and other characters. Either way, such descriptions of other characters say more about Eus than about them. As such, the quote is a manifestation of what Suleiman calls the ‘interpretive function’ of narrators (1983, p. 157): Eus’s descriptions of women denote his values and his judgements about other characters.

These basic narratological observations are key to the interpretation of character centrality in the novel. As first-person narrator and main focalizer, Eus is part of both the social group of people with a migration background and the social group of men. These simple observations suggest that the non-Dutch and (heterosexual) male perspectives are a priori more dominant than the Dutch and female perspectives. Taking into account that most of Eus’s friends and fellow criminals (Kosta, Ata, Meltem, Mahir) are also both male and from a non-Dutch descent, one could argue that the center of gravity lies with both male characters and characters with a migration background.

However, a closer look at ‘the interpretive function’ (Suleiman, 1983, p. 157) of Eus as a narrator complicates this preliminary conclusion. While Eus seems to embody the perspective of people with a migration background, the values which he communicates do not tend to coincide with the values he ascribes to other characters with a migration background. More specifically, his judgments
on characters with a Turkish descent are often outspokenly negative. This is most notably exemplified by his descriptions of his father Turis, a first-generation Turkish migrant, who is characterized by Eus as a drunken ‘tyrant’ (Akyol, 2012, p. 12) and a work-shy social parasite who is ‘Rather lazy than tired’ (p. 13). Through such judgements on people with whom he shares his cultural roots, Eus adopts a critical stance toward their values. Moreover, Eus does not seem to identify with Turkish people more generally. This is illustrated by his friendship with his Dutch friend Karelje, with whom he joins a football club consisting of ‘fifteen gypsies’ (p. 87). Although most Turkish boys play football at the ‘club for all Turks in the city’ (p. 88), Eus would never think of joining that club:

I would never play football there. I had nothing in common with those people. They didn’t even serve beer in the cantine. Only tea.

Such negative judgements on the values of Turkish people make Eus a very atypical Turk. While the fact remains that he has a Turkish background, he does therefore not automatically represent the typical Turkish perspective. According to Eus, he even has ‘nothing in common with those people’ as he feels more strongly connected to the group of Dutch gypsies: ‘[t]hey did not see me as a migrant. I was one of them’.

Despite his adversarial attitude toward people with a Turkish background and his connections with Dutch gypsies, Eus cannot avoid being stigmatized as belonging to the social group of people with a migration background. This is best exemplified by the ways in which he is treated in the Dutch school system. Despite his excellent performance in primary school, his teacher did not allow him to follow the highest educational level because of his Turkish background (‘She said that I would have had a very hard time as a Turk’ [p. 36]). This stigmatization continues in high school when his math teacher pejoratively calls him and his friend Metin ‘de Hasans’ (p. 42). While Eus’s judgements on the social group of migrants give the impression that he does not identify with this group, the actions and words of other characters make painfully clear that he cannot escape this social identity. From a narratological perspective, however, it still is the question whether or not the first-person narration of Eus therefore ascribes a more central role to characters with a migration background. As a narrator and main focalizer, Eus represents a complex sociocultural identity that resists being part of binary categories such as the Dutch or the non-Dutch. Based on his dominance as a narrator, it is clear that his perspective is most central to the narrative. However, his resistance toward a fixed sociocultural identity makes it problematic to claim that therefore the center of gravity lies with the migrant perspective.
Things are less ambiguous with regard to gender. Almost without exception, Eus's first-person narration and dominant focalization communicates an extremely masculine, heterosexual worldview. Throughout the novel, he and his friends encounter a variety of girls whose sole function is to fulfill their sexual desires. In high school, Eus and his friend Metin arrange a double date with Eef and Levine. The conversation they have prior to this date serves as a clear example of the dominance of the male, heterosexual perspective:

In the afternoon Metin and I had discussed the terms and conditions of our meeting with the girls. We agreed that he would take Eef and I would take Levine. For him, double dates were a routine job. 'The one with the big tits is more compliant,' he said, 'I can immediately tell. She won't be making a fuss. You can take the other one. You like serious girls.' (p. 59)

Because of the first-person narration, the novel is structured in such a way that only sparse attention is dedicated to the female perspective represented by girls such as Eef and Levine. Most information on the female characters is indirect, for instance through conversations between Eus and his friends such as this one. These conversations tend to follow a similar routine: a self-confident bravado, machismo, an objectification and sexualization of female bodies, lack of respect, lack of interest for anything other than bodily traits or sexual performance, et cetera. There is simply no female counterview present in the novel to nuance, criticize, or reverse the images of women as 'whores' (pp. 36, 58, 85, 145), 'sluts' (pp. 43, 57, 62, 86, 145, 157, 176, 253), or 'preys' (p. 60) as represented by the first-person narration of Eus and the focalization of the main character and his male friends.

But how seriously should the reader take all this macho bravado? Eus is clearly inspired by Jan Cremer's *Ik, Jan Cremer* (1964), one of the classics of Dutch picaresque novels. Not only does it have an intertextual relation with this novel through its thematization of being a social outcast, it also imitates the repetitive descriptions of sexual intercourse with attractive women. Although it can be argued that the sexualization and objectification of women in Eus is simply a convention of the genre of the picaresque novel, it nonetheless expresses an extremely masculine, heterosexual worldview. While the reader can choose to read the sexual escapades of the womanizing protagonist as an ironic allusion to the genre of the picaresque novel (‘Eus, the biggest player of the East. See how he rolls!’ [p. 62]), this does not change the fact that the female perspective is subordinate to the male perspective. Moreover, there is a wide range of scenes
with female characters in which irony is hard to find. When Eus, for instance, finds out that Levine had had sexual intercourse with another man, his double moral standards regarding men and women are unambiguously expressed. While Eus has a variety of sexual contacts, Levine is not supposed to do so. In his eyes, she is ‘a whore, a piece of filth’ (p. 85), and ‘a slut who has loose sexual morals’ (p. 86). Such scenes show how the ‘interpretive function’ (Suleiman, 1983, p. 157) of narrators co-shapes the masculine worldview presented in the narrative by expressing judgements on other characters. Narratologically, it thus seems safe to say that the male perspective occupies a more central, important, dominant, influential, and powerful position in the narrative than the female perspective.

How do these narratological observations relate to the statistical baseline that female characters and characters with a migration background are relatively central in the corpus as a whole? First of all, it is insightful to determine to what extent Eus conforms to or deviates from this pattern in a statistical sense. Does it live up to the pattern or does it form an exception to the rule? An answer to that question can help to contextualize the narratological analysis of centrality based on narration and focalization in light of the statistical baseline. Table 3 shows the characters in the novel ranked by their scores on degree centrality. The character ranking in this table demonstrates that the particular narrative of Eus conforms to the general pattern as observed in the multiple linear regression (see section 3.5.1 of this chapter) only with regard to descent. Of the 21 identified characters, 12 have a migration background, and they are higher in the rankings than the Dutch characters, which is in line with the general pattern according to which characters with a migration background have a significantly higher degree of centrality.

However, with regard to gender, the novel deviates from the pattern. Of the 21 identified characters, 14 are male, and they occupy higher positions in the rankings on degree centrality, indicating that the male characters in Eus have more relations than the female characters.

For Katz centrality, a similar pattern emerges. Table 4 lists the characters in the novel ranked by their scores on Katz centrality. Here, too, both characters with a migration background as well as male characters occupy higher positions in the rankings than nonmigrant characters and female characters. The first types of characters are thus connected to more important characters than the latter.
<table>
<thead>
<tr>
<th></th>
<th>Name</th>
<th>Gender</th>
<th>Descent</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kosta</td>
<td>male</td>
<td>immigrant</td>
<td>0.65</td>
</tr>
<tr>
<td>2</td>
<td>Kareltje</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.55</td>
</tr>
<tr>
<td>3</td>
<td>Eus</td>
<td>male</td>
<td>immigrant</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>Turis</td>
<td>male</td>
<td>immigrant</td>
<td>0.40</td>
</tr>
<tr>
<td>5</td>
<td>Meltem</td>
<td>male</td>
<td>immigrant</td>
<td>0.40</td>
</tr>
<tr>
<td>6</td>
<td>Ata</td>
<td>male</td>
<td>immigrant</td>
<td>0.40</td>
</tr>
<tr>
<td>7</td>
<td>Mahir</td>
<td>male</td>
<td>immigrant</td>
<td>0.35</td>
</tr>
<tr>
<td>8</td>
<td>Selma</td>
<td>female</td>
<td>immigrant</td>
<td>0.30</td>
</tr>
<tr>
<td>9</td>
<td>Metin</td>
<td>male</td>
<td>immigrant</td>
<td>0.30</td>
</tr>
<tr>
<td>10</td>
<td>Haakneus</td>
<td>female</td>
<td>immigrant</td>
<td>0.30</td>
</tr>
<tr>
<td>11</td>
<td>Levine</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.30</td>
</tr>
<tr>
<td>12</td>
<td>Theo</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.15</td>
</tr>
<tr>
<td>13</td>
<td>Nathan</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.15</td>
</tr>
<tr>
<td>14</td>
<td>Eef</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.15</td>
</tr>
<tr>
<td>15</td>
<td>Inez</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.1</td>
</tr>
<tr>
<td>16</td>
<td>Ömer</td>
<td>male</td>
<td>immigrant</td>
<td>0.1</td>
</tr>
<tr>
<td>17</td>
<td>Angelo</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.1</td>
</tr>
<tr>
<td>18</td>
<td>Vinny</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.1</td>
</tr>
<tr>
<td>19</td>
<td>Osman</td>
<td>male</td>
<td>immigrant</td>
<td>0.05</td>
</tr>
<tr>
<td>20</td>
<td>Daphne</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.05</td>
</tr>
<tr>
<td>21</td>
<td>Moeder Eus</td>
<td>female</td>
<td>immigrant</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 3. Characters in Eus (2012) ranked by degree centrality score.
<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
<th>Descent</th>
<th>Katz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Kosta</td>
<td>male</td>
<td>immigrant</td>
<td>0.218218982823223</td>
</tr>
<tr>
<td>2 Mahir</td>
<td>male</td>
<td>immigrant</td>
<td>0.21821840342212764</td>
</tr>
<tr>
<td>3 Eus</td>
<td>male</td>
<td>immigrant</td>
<td>0.2182184034212662</td>
</tr>
<tr>
<td>4 Kareltje</td>
<td>male</td>
<td>nonimmigrant</td>
<td>0.2182182875401764</td>
</tr>
<tr>
<td>5 Turis</td>
<td>male</td>
<td>immigrant</td>
<td>0.21821822960063142</td>
</tr>
<tr>
<td>6 Ata</td>
<td>male</td>
<td>immigrant</td>
<td>0.218218113719957</td>
</tr>
<tr>
<td>7 Meltem</td>
<td>male</td>
<td>immigrant</td>
<td>0.2182179398985992</td>
</tr>
<tr>
<td>8 Selma</td>
<td>female</td>
<td>immigrant</td>
<td>0.21821782401818632</td>
</tr>
<tr>
<td>9 Haakneus</td>
<td>female</td>
<td>immigrant</td>
<td>0.218217824018094</td>
</tr>
<tr>
<td>10 Levine</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.21821782401787862</td>
</tr>
<tr>
<td>11 Metin</td>
<td>male</td>
<td>immigrant</td>
<td>0.2182178240177709</td>
</tr>
<tr>
<td>12 Theo</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.2182177660783798</td>
</tr>
<tr>
<td>13 Nathan</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.21821770813789643</td>
</tr>
<tr>
<td>14 Eef</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.21821765019709</td>
</tr>
<tr>
<td>15 Inez</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.21821759225711432</td>
</tr>
<tr>
<td>16 Angelo</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.218217592257022</td>
</tr>
<tr>
<td>17 Vinny</td>
<td>male</td>
<td>non-immigrant</td>
<td>0.218217592257022</td>
</tr>
<tr>
<td>18 Ömer</td>
<td>male</td>
<td>immigrant</td>
<td>0.21821759225692972</td>
</tr>
<tr>
<td>19 Daphne</td>
<td>female</td>
<td>non-immigrant</td>
<td>0.21821753431670787</td>
</tr>
<tr>
<td>20 Osman</td>
<td>male</td>
<td>immigrant</td>
<td>0.21821753431661559</td>
</tr>
<tr>
<td>21 Moeder Eus</td>
<td>female</td>
<td>immigrant</td>
<td>0.21821747637639374</td>
</tr>
</tbody>
</table>

Statistically speaking, *Eus* thus conforms to the baseline only with regard to descent and not with regard to gender. As is the case in the corpus as whole, characters who are categorized as having a migration background occupy relatively central positions in the fictional network of this novel. Novels such as *Eus* arguably contribute to the relatively central network position of characters with a migration background in the corpus as a whole. *Eus*, however, is an outlier with regard to gender: while female characters tend to be relatively central in the corpus as a whole, they are statistically less central in the fictional network of this novel.

How does the statistical analysis of network centrality in *Eus* relate to the qualitative, narratological analysis of centrality, gender, and descent in the novel? Interestingly, the rankings of characters as presented in Table 3 and 4 are partly in line with the narratological analysis. This is most clearly exemplified for character centrality with regard to gender. While the narratological analysis has made the case that the narration and focalization in the novel co-constitute an extremely dominant male perspective, male characters also occupy more central positions in the fictional network of this novel. Furthermore, the fact that *Eus* deviates from the overall finding that female characters are more central in the fictional social networks in the corpus as a whole — i.e., the baseline — ascribes even more significance to the narratological finding that the female perspective is subordinated to the male perspective in the novel. Statistically speaking, *Eus* is a peculiar case with respect to how central female and male characters are as it deviates from the statistical baseline. As such, the qualitative, narratological assessment of the extremely masculine worldview is backed up by the quantitative, statistical finding that male characters have higher scores on degree and Katz centrality. The dominance of the male perspective in the novel, as reported by the narratological analysis, thus deviates from a statistical norm, and is therefore even more salient. It suggests that *Eus* does not just follow a norm of representation (i.e., female characters occupy relatively central network positions), but forms an exception to the rule. In light of its deviation from this norm, this baseline, the qualitative observation of this dominant masculine view stands out more than it would have without a comparison against this norm.

While the findings of the statistical analysis and the narratological analysis are complementary with regard to gender, they are less so with regard to descent. Statistically, *Eus* ascribes relatively central positions to characters with a migration background, which is in line with the baseline that was generated for the corpus as a whole. Narratologically, the centrality of the sociocultural identity of ‘the migrant’ is more complex. Although the qualitative assessment of narration and focalization in *Eus* demonstrate that Eus as a first-person narrator is in any case
the most central actor in the narrative, it has also underscored how Eus resists being part of a social group of either ‘the Dutch’ or ‘the migrants’. As such, the narratological analysis highlights a fundamental challenge of statistics-based approaches to literary representation. While a statistical analysis requires clear categorizations, literary texts have the potential to disrupt, criticize, or deconstruct such seemingly fixed boundaries. Although the migration background of Eus is indeed part of his identity, the novel more generally thematizes the nuances and complexities of categorizing people in binary categories, such as migrants and nonmigrants.

In sum, this narratological evaluation of the statistical baseline highlights two important points with regard to the interpretability of a statistical analysis of literary representation in general and the centrality of characters specifically. 1) A qualitative, narratological analysis of an individual text can provide a more nuanced backup of a statistical argument. Mode of narration and focalization in Eus illustrate the dominance of the male perspective, which is supported by the character rankings for the novel. In light of the statistical finding that female characters occupy more central network positions in the corpus as a whole, the deviation of this particular narrative from this pattern underscores the abnormality of the narratological finding that the novel communicates an extremely masculine worldview. 2) However, narratological observations might also nuance or conflict with statistical findings, which is the case with regard to the centrality of characters with a migration background. Characters with a migration background score higher than Dutch characters in the novel in terms of network centrality, which is in line with the statistical baseline. But despite the novel’s seemingly conformation to this overall pattern, the very notion of a fixed sociocultural identity is problematized in Eus. As such, the narratological analysis highlights that individual narratives have the potential to challenge statistics-based patterns of representation. More generally, this insight emphasizes the various, sometimes conflicting meanings that ‘centrality’ or ‘importance’ can have in qualitative narratology as opposed to quantitative network theory.

3.6 CONCLUSION TO THIS CHAPTER

This chapter aimed to find an answer to the question of how the centrality of characters co-shapes the representation of the groups in which they function.
Informed by insights from both network theory and narratology, a method was developed to extract fictional social networks of characters from the 170 novels in the research corpus. Based on these fictional social networks, a model was proposed to rank characters according to five centrality metrics. In order to determine which groups of characters tend to end up high in these rankings, and thus are central in statistical terms, a regression analysis was conducted to test which demographic features of characters predict their places in the rankings. The output of this data-driven, statistical model on the whole corpus, combined with the close reading of centrality, gender, and descent in the case study of Özcan Akyol’s *Eus*, indicates that centrality co-constitutes the representation of social groups in present-day Dutch literary fiction in at least these two ways.

First, the frequency of distributions of characters belonging to a certain social group such as reported in the previous chapter are not sufficient for analyzing how central men as opposed to women, migrants as opposed to nonmigrants, the higher as opposed to lower educated, and the young as opposed to the old are in narrative fiction. Although female and migrant characters are less present in the corpus, statistically speaking, they take up a more central position in the social networks of present-day Dutch literary fiction than nonmigrant and male characters. This remarkable outcome requires an explanation, particularly in light of the highly imbalanced frequency distribution of migrant and nonmigrant characters in the corpus. For those characters in the corpus whose descent is known, almost 90% do not have a migration background (i.e., a Dutch or Belgian background), but the regression model suggests that migrants are more central in the networks than nonmigrants. These higher centrality scores might be explained by the probability that novels that thematize descent, and stage a higher number of migrants, also ascribe more central roles to them, which was demonstrated in the close reading of *Eus*. Overall, the corpus contains fewer migrant characters (only around 10% of all characters have a migration background), but these migrants score higher on degree and Katz centrality. Something similar holds for female characters: there are fewer female characters than male characters in the corpus (almost a 40–60 ratio), but they have relatively high centrality values. In order for migrant or female characters to be central in network theoretical terms, a high frequency of occurrence is not a necessary prerequisite as long as they interact with a high number of other (central) characters.

Whereas frequency of occurrence of characters belonging to a certain social group is the most straightforward indication of their centrality in the text (how present they are), the model developed in this chapter proposes an additional, more sophisticated measure of centrality. Other than simply counting frequency
of occurrences of characters with a certain demographic profile, the network analytic approach developed in this chapter views centrality in a fundamentally relational sense. As result of this, centrality of characters is defined in relation to their interactions with other characters. While a certain social group (male, without a migration background) occurs more often in the corpus, they occupy fewer central positions in the social networks of their fictional worlds. This is not to say that their frequency of occurrence does not affect their centrality. How often a certain social group is depicted definitively affects its visibility (and thus centrality) in literary fiction, but equally important is the position characters from these groups have in the social structure relative to the position of other types of characters.

Second, the close reading of case study *Eus* has demonstrated how centrality in quantitative, statistical terms (frequency of occurrence, position in the network) relates to, and sometimes conflicts with, centrality in qualitative, narratological terms (thematic structure, mode of narration, focalization). Whereas the high number of characters with a migration background in *Eus* is not representative of the corpus, the centrality of these types of characters in the novel’s social network probably explains why migrant characters in the corpus take up relatively central positions in the networks. The novel thematizes descent, and ascribes a central role to the migrant perspective through its mode of narration and focalization, although it also challenges the notion of a fixed socio-cultural identity. As a consequence of this, migrant characters in this novel not only occupy more central roles in the network structure, statistically speaking, but arguably also leave their mark more profoundly on the narrative structure. As the novel’s first-person narrator represents the perspective of the social group of migrants, his perception of the events and the other characters in the novel is colored by this specific migration background. Because of this, the centrality of the first-person narrator in a narratological sense (through his narration and focalization) reinforces his centrality in the social network. However, the narratological analysis also shows that the narrator is resistant toward being categorized in a binary category such as ‘migrant’ or ‘nonmigrant,’ which highlights the potential of individual texts to disrupt, distort, nuance, or criticize inevitably reductive statistical patterns.

Qualitative assessment of an individual novel does, therefore, also show how individual narratives can not only support the statistics-based patterns found for the corpus as a whole but also have the potential to escape or transcend these patterns. Whereas statistical trends might indicate general patterns of literary representation in large collections of texts, they can subsequently serve as an
analytic backdrop for the individual analysis of particular novels. Using such statistical patterns as a baseline for comparison, the extent to which a single novel either conforms to or deviates from them can be used to determine the particularity of a certain aspect of representation. Contrasting the narratological analysis of centrality in *Eus* with the pattern that migrant and female characters occupy relatively central positions in each of the 170 networks, helps to pinpoint how peculiar or deviant this particular narrative is in this respect.