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=3191739
CHAPTER 4

COMMUNITY

4.1 INTRODUCTION: NARRATIVE CONNECTIONS

How do the communities in which characters function affect the representation of the social group(s) these characters are part of? The representation of social groups in present-day Dutch literature will be studied in this chapter through the lens of community. The concept will be used as an umbrella term denoting a variety of interrelated terms. It refers to a range of relational notions such as ‘clustering’, ‘coexistence’, ‘collective’, and ‘connectivity’, although each of these notions stresses a slightly different aspect of community. Whereas the next chapter on conflict focuses on the ways in which negative affiliations co-constitute the representation of social groups in literature, this chapter aims to understand how such representations are co-shaped through the bonds that characters form. While social groups are defined in terms of shared demographic characteristics (e.g., gender, descent, age, education), communities do not necessarily consist of individuals belonging to the same social group (e.g., a community can consist of both male and female characters, both migrant and nonmigrant characters, both older and younger characters, both lower and higher educated characters). The homo- or heterogeneity of communities of characters in narrative fiction can shed light on the extent to which social groups are either integrated or segregated into different communities.

What is a community? Classical social theorists such as Ferdinand Tönnies, Émile Durkheim, and Georg Simmel have been occupied with the question of why people unite in group-like structures. Tönnies famously made a distinction
between *Gemeinschaft* (commonly translated as community) and *Gesellschaft* (commonly translated as society). In his view, premodern social structures were typified by communities in the form of families and neighborhoods that were held together by a sense of belonging and a moral obligation to one another. Modernity witnessed the decay of these premodern communities and a transition to societies in which companies and states became the essential social structures (Tönnies, 1887/2005). To his regret, Tönnies observed that this transition also embodied a shift from intrinsic, morally connotated motivations for being together to social structures that merely serve as instruments to achieve joint goals.

An alternative view on the shift from premodern to modern modes of living is proposed by Émile Durkheim in *The Division of Labour and Society* (1893/2013). Whereas Tönnies ascribes rather positive features to his idea of the premodern *Gemeinschaft*, Durkheim has a more negative conception of such primitive social structures. These early communities are built upon what Durkheim calls mechanical solidarity. As there is little division of labor (people are mostly carrying out similar tasks), such communities subsequently share similar values and are conjoined in what he calls a ‘collective conscience’. Katherine Giuffre characterizes Durkheim’s idea of primitive community as ‘an entity that is more than the sum of its parts’ because ‘the moral force of the collective conscience acts on the members of the community to create the feeling of a shared identity’ (Giuffre, 2013, p. 22). In order to ensure that this collective conscience does not fall apart, repressive law is installed to sustain their shared identity and to prevent a divergence into distinct, individual identities. Individuality, in other words, is a threat to the very existence of the community.

Durkheim is more positive about modern, capitalist societies in which organic instead of mechanical solidarity is the binding mechanism. The labor in such societies is highly differentiated, a result of which is that there is great heterogeneity among its members. In contrast to primitive communities, individuality is not a threat but a driving force behind these advanced societies. Law is not repressive but restitutive, the point of which is ‘not to punish transgressions against the collective conscience – which has been weakened by the division of labor to the point where it can no longer provide a source of community cohesion – but to ensure the orderly functioning of the various differentiated “organs” of the community’ (p. 24). The strong collective conscience of primitive communities brings about a shared identity; the downside is that it needs repressive laws to smother individuality and heterogeneity. Modern societies have a weak collective conscience, but individuality can thrive for the greater good. For Durkheim, it is not sameness but difference that binds people in modern societies.
A fundamentally different take on communities can be found in the formal sociology of Georg Simmel. Whereas Tönnies and Durkheim utilize abstract notions such as moral belonging and shared identity, Simmel emphasizes the concrete pathways through which a community member’s individuality takes shape. A dual interplay between individual and social group lies at the heart of the theory that he developed in *Group Expansion and the Development of Individuality* (1908/1971): an individual is defined by the social groups it belongs to, a social group is defined by its individual members. ‘The uniqueness of the individual,’ Giuffré summarizes, ‘is based on her or his position at this nexus of a unique set of circles’ (Giuffré, 2013, p. 28). For Simmel, communities are equal to the sum of their elements; they are not entities on their own as is the case for Tönnies and Durkheim. Simmel’s approach is fundamentally relational, which is why he is commonly regarded as a forerunner of social network analysis. In network theory, communities are not defined through metaphorical notions such as ‘collective conscience’ but through the concrete relational ties between nodes. By adopting such a relational approach, a collection of nodes can be broken down into subgroups of nodes that are more densely connected with one another than they are with others.

This chapter takes up a Simmelian, network analytic approach to community combined with a narratological perspective. Most importantly, this approach enables a better measurability and quantifiability of the bonds between characters than the theories of Tönnies or Durkheim would. This does, however, not mean that Durkheim’s and Tönnies’s more metaphorical conceptions of community are not relevant for the purposes of this chapter. In thinking about the representation of groups of people in literature, it is unavoidable to refer to, for instance, the notion of a shared identity. Metaphors are, moreover, an indispensable part of literary fiction; and characters can comment on their social worlds through many different figures of speech. In general, there are fruitful analytic angles to be found in literary theory taking such literary-stylistic mechanisms into account. Because of their reliance on literature’s metaphorical and symbolic abilities, these theories are usually miles apart from the formal network analytical approach to communities as represented by Simmelian sociology. As the raw material of literary texts is language, most of these angles depend upon a particular idea of how literary language works. Such is the case in the writings of scholars associated with the Bakhtin circle that will be used in one of the subsequent sections.

In accordance with the twofold theoretical setup of this book, the empirically oriented approach of network theory will be put in dialogue with a more top-down perspective from literary theory. These theoretical frameworks will be
used to gain a closer understanding of how character communities play a part in shaping the literary work’s depiction of social groups. This is particularly relevant in light of the observation that the notion of community can gain insight into the extent to which social groups are either integrated or segregated (Blau, 1977). Which characters do and which characters do not belong to specific groups? Two models will be presented to assess the degree to which social groups in the present corpus are integrated or segregated. The first model uses community detection algorithms to trace which groups of characters cluster more strongly together than others. The second model computes how strongly characters with the same demographic profile tend to interact. Following Bakhtin, the results of both of these models are used to assess how homo- or heterogeneous – ‘polyphonic’ or ‘dialogic’ – the novels are in terms of their differentiation into different communities of characters (see paragraph 3.2). Combining these two statistical models with a Bakhtinian framework allows for an interpretation of the overall cohesion or fragmentation in the fictional populations of present-day Dutch literary fiction. Is there a con- or divergence of represented identities in the corpus? In order to assess the meaning of the statistical patterns generated by each model, the observed general trends are evaluated through close readings of Philip Huff’s Niemand in de stad [Nobody in the city] (2012) and Mensje van Keulen’s Liefde heeft geen hersens [Love has no brains] (2012).

4.2 COMMUNITY IN NETWORK THEORY

One of the seminal articles of social network analysis, Granovetter’s ‘The Strength of Weak Ties’ (1973), hypothesizes that strong edges exist within communities and weak edges between communities.

Linkage of micro and macro levels is thus no luxury but of central importance to the development of sociological theory. Such linkage generates paradoxes: weak ties, often denounced as generative of alienation [Wirth, 1938] are here seen as indispensable to individuals’ opportunities and to their integration into communities; strong ties, breeding local cohesion, lead to overall fragmentation. (Granovetter, 1973, p. 1378)
This idea is both intuitively clear and counterintuitive. It seems obvious that strong ties between members of a community are a conditio sine qua non, as otherwise it would be hard to speak of a community at all. A logical consequence of this observation is that such strong ties between subgroups are detrimental to the overall cohesion of networks. Paradoxically, strong ties lead to less cohesion. The storyworld of *Lord of the Rings*, for instance, consists of one large fictional population in which groups of elves, dwarves, wizards, men, and orcs are (in) directly connected through positive or negative associations. There are arguably stronger ties between the members within each of these groups, resulting in communities, than there are between, for example, groups of elves and dwarves. These relatively strong connections within each of these groups, however, do not lead to a cohesive and densely connected overall network. Quite the contrary: the *Lord of the Rings* network is fragmented into relatively separated communities consisting of groups of elves, dwarves, wizards, men, and orcs. Among these individual, fragmented communities of characters, a few elves, dwarves, wizards, and men – the fellowship of the ring – join forces to fight evil (orcs, Saruman, Sauron). In order to establish cohesion in an utterly fragmented world, they are obliged to transcend the boundaries of their individual communities. Fellowships (or communities) consisting of people from a wide variety of groups are, however, not the norm not in the fictional Middle Earth in *Lord of the Rings*, nor are they in real-world societies, regions, and countries. Most social worlds contain distinct, densely connected communities, which subsequently leads to fragmentation in the social world as a whole. It is, however, not at all an obvious task to draw lines between communities. What counts as relatively strong ties? What do ‘densely connected’ and ‘cohesive’ mean? One of the major challenges of network theory is how to assess the boundaries between different subgroups within a network.

### 4.2.1 Community Detection

Community detection is the header under which this methodological challenge is carried out. Confusingly, a variety of terms is used to refer to the objects of detection. *Community* is the most metaphorical of the terms, having different associations depending on the theory one adopts (e.g. Tönnies, Durkheim, or Simmel). (Cohesive) *subgroup* and *cluster* are more formal but also rather general terms denoting different things in different contexts as well. Clique is perhaps the most precisely defined among these terms, referring to a fully interconnected subset of nodes in a network. It was first coined by Luce and Perry in 1949 and represents the most stringent definition of group-like network structures.
Figure 1. Example of a clique. A, B, C, and D are all directly connected.

Figure 1 shows an example of a clique. A, B, C, and D all have direct links to all other nodes. While some definitions relax the criteria for cliques, the strictest definition ascribes that A-B-C-D ceases to be a clique when, for instance, edge A-D falls away.

The tight interconnectedness of A, B, C, and D does not mean, however, that A-B-C-D is fully isolated from other nodes. A construction such as shown in Figure 2 is quite possible.

Figure 2. Example of two cliques. Both A-B-C-D and E-F-G-H are fully connected. B and G form a bridge between the two cliques.
In the example of *The Lord of the Rings*, A-B-C-D might represent a clique of dwarves and E-F-G-H a clique of elves. Although both of these cliques share more dense connections with their own group than with the other group, there are no isolated entities within the overall *Lord of the Rings*. For instance, B may represent the dwarf Gimli and G may represent the elf Legolas; two characters connected through their co-participation in the fellowship of the ring.

It is debatable whether a stricter or looser definition of cliques should be adopted and whether linkages such as those between B and G should be taken into account. In order to avoid associations with the most stringent definition of clique, the looser term community will be used throughout the rest of this chapter to refer to any densely connected subset of characters within the overall character network in the novel. Not only is it a widely used term in different fields of inquiry, its metaphorical associations will prove to be useful in the interpretive parts of the chapter.

Thus, a community is generally defined here as ‘a cohesive group of nodes that are connected “more densely” to each other than to the nodes in other communities’ (Porter, Onnela, & Mucha, 2019, p. 1086). What ‘more densely’ entails is dependent on the method adopted for identifying communities of characters in literary texts. Ever since the first analyses of community structure (e.g. Homans, 1950; Rice, 1927; Weiss & Jacobson, 1955), different detection methods have been proposed in different fields of inquiry. Up to the present there is no generally accepted solution to the problem. Especially since the seminal paper by Girvan and Newman (2002), who first applied large-scale data handling and powerful computational tools to detect communities, the development and advancement of methods have been on the agenda of people working in applied mathematics and physics. Community detection gradually became ‘hip’ (Porter et al., 2009, p. 1083).

Community detection methods can be divided into agglomerative and divisive techniques. Agglomerative techniques start from the level of the individual node and gradually connect nodes into larger sub-communities of the full network. An example of an agglomerative method is linkage clustering (e.g. Johnson, 1967), in which ‘nodes are conjoined sequentially into larger clusters starting with the [most strongly connected pair]’. Conjoining the nodes into larger clusters is done stepwise by ‘recomput[ing] the similarities between the new cluster and each of the old clusters and [joining] the two maximally similar clusters, [continuing] iteratively until all clusters with nonzero similarity are connected’ (Porter et al., 2009, p. 1087). Conversely, a divisive technique
starts ‘with the full graph and breaks it up to find communities’ (ibid.) instead of starting from individual nodes.

One of the most popular community detection methods is optimizing so-called modularity (Prokhorenkova, Pralat, & Raigorodskii, 2016), which can be done either bottom-up by starting from individual nodes (agglomerative) or top-down by decomposing the overall network (divisive). Networks with high modularity have dense connections within communities and sparser connections between communities. The aim of modularity optimization algorithms is to ‘[seek] an arrangement in which the difference between the number of within-community relationships and the number anticipated in a randomly structured network is as large as possible’ (Marsden, 2011, p. 599). Such algorithms thus hinge on a comparison between the network as it is and a hypothetical network in which the nodes are randomly connected. For weighted networks (networks in which a particular weight is ascribed to edges), such methods ‘[measure] when a particular division of the network has more edge weight within groups than one would expect by chance’ (Porter et al., 2009, p. 1089). A disadvantage of using modularity optimization algorithms is that they have an inherent resolution limit (Fortunato & Barthélemy, 2007), the result of which is that ‘it misses communities that are smaller than a certain threshold size that depends on the size of the network and the extent of interconnectedness of its communities’ (Porter et al., 2009, p. 1091). Several algorithms therefore allow one to adjust a resolution parameter specific to the network size.

Approaching communities from a modularity perspective opens up interesting possibilities with regard to node attributes such as gender, descent, age, and education. The concept of modularity is often associated with assortativity, the notion that nodes/edges sharing features are more likely to be connected than nodes/edges that do not have these features in common. For the purposes of this chapter, assortativity relating to node attributes is especially relevant. Node attributes relating to, for instance, descent, can be used to determine the extent to which characters from the same descent are part of the same communities. Subsequently, this might be an indication of how fragmented the network is:

Fragmentation results when a network’s elements are clearly partitioned into subgroups, with few between-cluster relationships. If the network’s activities require little coordination among elements in its different parts, such loose coupling can allow clusters to function in appropriately autonomous and efficient ways. Fragmentation can compromise a network’s capacity to pursue joint tasks, however, especially when it
Community is polarized, with between-community antagonism alongside within-cluster solidarity. (Marsden, 2011, p. 600)

Networks with high modularity have dense connections within communities but sparse connections between members of different communities. In order to find an answer to this chapter’s question on how community-formation affects the representation of social group(s), it is worthwhile to assess how fragmented the fictional networks are in terms of gender, descent, education, and age. The extent of fragmentation, then, is an indication of how these fictional populations are depicted as being either integrated into a densely connected whole or as being segregated into distinct clusters – whether there is a convergence or divergence of identities. Quantitative, statistical analysis may be informative about the general nature of this fragmentation, but insight into phenomena such as polarization or ‘between-community antagonism’ can only be gained through qualitative close readings. In section 4.4.3, such a qualitative assessment of narrative communities will be conducted for Philip Huff’s *Niemand in de stad*.

### 4.2.2 Homophily

Communities tend to consist of members with a similar background (Marsden, 2011, p. 599). In network theory this mechanism is studied through the concept of homophily, ‘a principle of social organizing defined as people sharing similarities tending to have more social interaction’ (Seidel, 2011, p. 382). Often it is characterized through the proverbial expression ‘birds of a feather flock together’ coined by Lazarsfeld and Merton in 1954. Ideologically, homophily is an interesting sociocultural mechanism as ‘limited homophilous networks may serve as a structural barrier for minorities’ (Seidel, 2011, p. 383). Groups of similar members tend to exclude people who are not similar to them. Studying homophily in literature can thus be a means to gain insight into structural inequalities and hierarchies between characters with different demographic profiles.

The first written records associated with the notion of homophily date back to Aristotle and Plato, both of whom suggested that similarity binds people together. From the 1920s onward, systematic studies have been conducted that to a greater or lesser extent confirm the idea that similarity leads to stronger associations. Initially, studies focused on small groups such as friend circles at schools and colleges (e.g., Bott, 1928). Later, research on the topic witnessed a growth in scale because of the use of sample surveys that could be applied to societies as a whole (e.g. Marsden, 1987). More recently, even larger-scale
homophily studies have been conducted on online networks such as those of Facebook or Goodreads users (e.g., Bucur, 2019). There is one study that explores the idea of gender homophily in networks of characters in English-language fiction in depth (Kraicer & Piper, 2019), which is closest in nature to the analyses carried out in this chapter. Preparatory work for the present research was done in Volker and Smeets (2019), in which the Libris dataset was compared with actual personal networks of Dutch people through the framework of homophily. In this comparative analysis, we found that the character networks in the Libris corpus are less homophilous in terms of descent, education, and age than actual networks in Dutch society. The Libris dataset used in this analysis, however, only contained 1,292 characters. Currently, the Libris dataset has been augmented to a total of 2,137 characters and thus covers a larger portion of character interactions in the books, taking into account a fair number of less visible side characters as well. Moreover, the methodical setup of the present chapter differs from the research in Volker and Smeets (2019). The results reported in this chapter broadly confirm the findings reported in this previous study.

There are two categorical distinctions that are generally made within this field of inquiry. The first divides homophily into a baseline and an inbreeding category. Whereas baseline homophily refers to the opportunities that people have to engage in contact with similar others, inbreeding homophily denotes the internal motivation people have to interact with similar others. The baseline category is linked to the statistical chances of a network member to have the opportunity to interact with similar others. Being born in a particular situation, one might have a certain a priori chance to interact with similar people. Conversely, inbreeding refers to the interaction with similar others beyond this a priori chance. Regardless of the statistical possibility that someone connects with similar others, similarity of (for instance) race, gender, or social class might attract someone to find contacts outside of this opportunity structure.

The second categorical distinction is between status and value homophily (Lazarsfeld & Merton, 1954). Status homophily refers to similarity based on sociodemographic status. This status can either be ascribed, relating to characteristics such as gender, descent, race, ethnicity, or age, or this status can be acquired, relating to characteristics such as education, profession, or religion. Value homophily applies to network members who have similar thoughts, values, beliefs, and motivations regardless of their social status. One of the first observed examples of value homophily is a study reporting that people with similar intelligence levels are likely to get in contact (Almack, 1922). Political orientation is another feature that tends to fuel homophilous associations
Community (Verbrugge, 1977). For pragmatic reasons, this chapter only focuses on status homophily. Although forms of value homophily such as intelligence or political orientation are interesting from an ideological perspective, such characteristics are commonly features of characters that are not articulated, which makes it practically impossible to determine what every character’s level of intelligence or political orientation is. The present dataset contains metadata on the characters on the basis of which either ascribed (gender, descent, age) or acquired (education, profession) status homophily can be studied.

As in the previous chapters, the four demographic categories of descent, gender, age, and education will be the point of focus in the following analysis. The categories of race and ethnicity are the cause of the largest divides in contemporary societies. Strong baseline and inbreeding homophily on these features have often been reported (e.g., Marsden, 1987, 1988). Gender forms a considerably less large divide in societies. Although studies reported that, for example, schoolchildren tend to form gender homophilous circles of friends at an early age (Maccoby, 1998), no strong baseline or inbreeding homophily is found in contemporary societies such as is the case with race and ethnicity (Marsden, 1988). Homophily related to age differs per nature of the network but it is generally high in marriages and in networks of school children (McPherson et al., 2001, pp. 424–425). Education homophily has been reported to be high (Marsden, 1987; Verbrugge, 1977) and has a strong inbreeding component (Marsden, 1988).

What are the causes of such homophilous associations? There are at least two causes of why people engage in relationships with people that are like themselves. The main cause is geographic space. Although we live in a globalized world, our main relationships are structured locally in households, neighborhoods, and schools (Gans, 1968). Such functional local spaces foster meeting opportunities with people from similar backgrounds. For the social worlds depicted in novels, this raises questions about the role of narrative space: does the fact that two characters co-occur often in the narrative increase the chance of them being demographically similar? Another cause for homophily is family structure. Although family networks are rather non-homophilous with regard to gender (because of the dominance of heterosexual relations and the equal chance of having either sons or daughters), they tend to produce race, ethnicity, and religion homophily (Smits, Ultree, & Lammers, 2000). In light of the fact that family ties constitute the majority of the type of relations in the corpus (see chapter 2, ‘Data’), a possible hypothesis is that the unit of the family also fosters segregation in character networks. In the close reading of Mensje van Keulen’s *Liefde heeft geen hersens* in section 4.5.2, it will be assessed how geographic space and family
structure form a breeding ground for the homophilous associations related to age in the novel.

4.3 COMMUNITY IN NARRATOLOGY

Modern Western literature is arguably less about community than it is about the self. In The Political Unconscious (1981, pp. 137–171), Fredric Jameson emphasizes the role of the novel in ‘the emergence of the ego or centered subject’ in late capitalist societies (p. 140). On a concrete level, this is exemplified by the observation that most present-day novels center on the thoughts and feelings of particular subjects in the form of characters. One might even suggest, as Sandra Zagarell does in a contribution to Signs (1988), that such a focus on the subjective experiences of protagonists rather than on the shared experiences of groups, collectives, or communities fuels a sense of ego in contemporary culture.

There is, however, a literary tradition in which community rather than the self prevails. ‘Narrative of community’ is the term Zagarell uses to refer to this tradition. Novels in this tradition ‘take as their subject the life of a community (life in “its everyday aspects”) and portray the minute and quite ordinary processes through which the community maintains itself as an entity’ (Zagarell, 1988, p. 499, emphasis in original quote). Obviously, the self is not fully absent network of the community rather than as an individualistic unit’ (ibid.). Historically, this genre grew in response to processes of modernization, urbanization, and industrialization that were detrimental to premodern ways of living characterized by a stronger sense of community (see Durkheim, 1983/2013 and Tönnies, 1887/2005, and the introductory paragraph of this chapter). With its roots in the nineteenth century, traces of this literary tradition are arguably still visible in literary fiction today. The following two subsections explore in more depth narratological terms and theories applicable to the study of narrative representations of communities.

4.3.1 Syntagmatic and Paradigmatic Collectives

Whereas network theory has a wide range of concepts and tools to study communities in networks, narratology lacks obvious instruments for studying narrative communities. This is exemplified by the absence of the lemma ‘Community’ in Greimas and Courtés’s analytic dictionary of semiotics (1979).
The lemma ‘Collective’, however, describes elements that are indirectly relatable to representations of communities in literary texts. As always, the definition of the authors is rather dense: ‘An actant is called collective when he, being part of a collection of individual actors, is provided with a modal capacity and/or actions that are shared with all actors that he arranges’ (Greimas & Courtés, 1979, pp. 54–55). In the context of this chapter’s focus on the representation of social groups, the collective actant in this sentence can be considered as representing a group of characters and the individual actors as representing single characters. The ‘modal capacity’ can be viewed as a shared feature of characters belonging to a group; an all-female group of characters, for instance, share the modal capacity of being female. Thus, a collective of characters comes into being when characters have something in common. Indeed, this can be everything ranging from the same gender, country of descent, educational level, or age group, to shared histories, political beliefs, or religious views. Importantly, Greimas and Courtés make a distinction between two kinds of collectives. They define the ‘syntagmatic collective actant’ as ‘the actant in which the single actors, added up as ordinal numbers, alternate in the performance of one [narrative] programme (similar to the joint effort of single workers when building a house)’ (p. 55). Characters in a syntagmatic collective, then, are on an equal footing as they jointly operate in a shared effort. The authors describe the notion of the ‘paradigmatic collective actant’ as the hierarchical counterpart of the syntagmatic collective. This paradigmatic collective actant is ‘a more extensive and hierarchically higher distribution defined by classes’. Characters in a paradigmatic collective cluster together in different classes, such as characters with a high, middle, and low education. In most cases, a novel represents a paradigmatic collective as it contains hierarchies between characters belonging to a certain category. The population of characters in a novel can, for instance, be divided in different subpopulations of higher educated, middle educated, and lower educated characters. But specific syntagmatic collectives exist amidst the paradigmatic collective of the novel as a whole, which are represented by these different subpopulations. All members of a collective of higher educated characters are equal with respect to their education; a hierarchical ordering with respect to education only becomes apparent in relation to communities of characters with a different educational level.

Thus, the novel as a whole represents a paradigmatic, hierarchically ordered collection of collectives, in which each individual collective has a syntagmatic, nonhierarchical structure. At least, in this example. If an individual collective features characters with, for instance, both higher, middle, and lower educated characters, than there is a paradigmatic, hierarchical ordering within that collective
as well. The distinction between syntagmatic and paradigmatic collective will be used as a general conceptual framework in the close readings of the case studies in sections 4.4.3 and 4.5.2 of this chapter.

4.3.2 Dialogic Interaction and Polyphony

More generally, the ideas of Russian scholars associated with the Bakhtin circle cater to a conceptualization of communities in the novelistic genre. This is especially true for the early twentieth-century work of Valentin Nikolaevich Voloshinov and Mikhail Mikhailovich Bakhtin himself. Although they do not explicitly focus on the community concept, their ideas about the assets and mechanisms of literary language provide a fruitful analytic angle for the study of narrative communities.

In *Marxism and the Philosophy of Language* (1929/2003), Voloshinov philosophizes about the workings of language in general. Voloshinov’s main point, and the theoretical point of departure for most of those of the Bakhtin circle, is that all utterances are inherently a form of dialogic interaction:

Utterance, as we know, is constructed between two socially organized persons, and in the absence of a real addressee, an addressee is presupposed in the person, so to speak, of a normal representative of the social group to which the speaker belongs. The *word is oriented towards an addressee, toward who that addressee might be: a fellow-member or not of the same social group, of higher or lower standing (the addressee’s hierarchical status), someone connected with the speaker by close social ties (father, brother, husband, and so on) or not. There can be no such thing as an abstract addressee, a man unto himself, so to speak. With such a person, we would indeed have no language in common, literally and figuratively. Even though we sometimes have pretentions to experiencing and saying this *urbi et orbi*, actually, of course, we envision this ‘world at large’ through the prism of the concrete social milieu surrounding us. In the majority of cases, we presuppose a certain typical and stabilized *social purview* toward which the ideological creativity of our own social group and time is oriented, i.e., we assume as our addressee a contemporary of our literature, our science, our moral and legal codes. (Voloshinov, 1929/2003, p. 58; emphasis in the original text)

The fundamental idea here is that utterances between people only make sense ‘through the prism of the concrete social milieu’. This served as an important
framework for the more ideologically oriented strands of literary criticism from the 1960s onwards, which foregrounded the specific social lenses through which literature can be approached. A word is a bridge thrown between myself and another (ibid.): language is a fundamentally dialogic, interconnective, interactive, and interrelational mechanism of communication. Language has therefore the potential to connect people, identities, and cultures. This insight forms the conceptual backdrop for a series of essays by Bakhtin on the nature of the novelistic genre.

In the essay ‘Discourse in the Novel’ (1935/2003b), Bakhtin coins the concept of heteroglossia to refer to the stratification of language into multiple social discourses, of which each represents a specific ideological system of beliefs. He distinguishes between the centripetal forces of unitary language striving to verbal unification and centralization, and the centrifugal forces of heteroglossia striving to verbal deunification and decentralization. Whenever Bakhtin writes about the centripetal and centrifugal forces of language, he is not so much referring to a (de)unification and (de)centralization of language in a purely linguistic sense, as he is first and foremost describing the ideological mechanisms underlying linguistic utterances. According to Bakhtin, literary language is one of the so-called ‘heteroglot’ languages striving for verbal and thus ideological deunification and decentralization and is itself also stratified into different heteroglossia. For instance: although most of the novels from the Libris corpus are written in the same standard Dutch, each novel is made up of language that does not represent one specific – unified and centralized – ideological position but rather represents a variety of different, possibly conflicting ideological messages. From the standpoint of Bakhtin’s theory, each of those novels has the potential to escape the ideological constraints of unification and centralization to which other, nonliterary, linguistic utterances are subjected to.

What this exactly entails for the study of literature becomes most clear in Bakhtin’s analysis of Dostoevsky in Problems of Dostoevsky’s Poetics (1929/2003a). In this essay he makes a widely recognized distinction between monologic novels such as those of Tolstoy and dialogic novels such as those of Dostoevsky. In the monologic novel, character voices are subordinated objects of the standpoint of the author, whereas the dialogic novel features character voices that exist alongside the authorial point of view. Bakhtin argues that Dostoevsky rejects the objectifying authorial viewpoint and replaces it with interactions between the different identities of the characters, thereby creating ‘a great dialogue of interacting voices, a polyphony’ (Morris, 2003, p. 89). Whereas the term heteroglossia represents stratified social languages in general, the term polyphony stands for the individual
or collective voices in the novelistic genre. A character in a Dostoevsky novel is no ‘mouthpiece for the author’s voice’ (Bakhtin, 1929/2003a, p. 89) as is the case for Tolstoy’s characters. Rather, Bakhtin sees Dostoevsky’s novels as platforms where a diversity of characters and themes coexist and interact. A Dostoevskian character is ‘a carrier of a fully valid word and not the mute, voiceless object of the author’s words’ (p. 93).

Moreover, the ‘voices’ orchestrated in such polyphonic novels are not necessarily represented by single, individual characters. Commonly, these characters are to a greater or lesser extent defined by the community to which they belong. Consequently, the dialogue between all these different voices can be seen as an interaction between character communities. Importantly, the polyphony orchestrated by Dostoevsky’s dialogic novels contains no inherent hierarchy between represented identities according to Bakhtin. The consciousness of the hero of a story exists next to ‘other consciousnesses with rights equal to those of the hero’ (ibid). In monologic novels such as those of Tolstoy, there is no ‘connection between consciousnesses’ as all characters are subordinated to the objectifying authorial consciousness (ibid., p. 95). Conversely, dialogic novels feature ‘great dialogue, but one where the author acts as organizer and participant in the dialogue without retaining for himself the final word’ (p. 96).

Such statements invoke an image of the polyphonic novel as a radically democratic genre in which a variety of different communities are harmoniously integrated with one another. In Bakhtin’s essay ‘Discourse in the Novel’ (1935/2003b), mentioned earlier, the novel is defined as ‘a diversity of social speech types (sometimes even diversity of languages) and a diversity of individual voices, artistically organized’ (Bakhtin, 1935/2003b, p. 114). It is, however, rather unclear how the extent of diversity of such ‘social speech types’ can be assessed in a novel. As Bakhtin only provides a general framework in which novels have the potential to create a multi-voicedness, it is up to present-day scholars to find suitable methods to test this in principle alluring idea. Building on this classic literary theory, the following sections operationalize two models through which the extent of polyphony can be determined. How polyphonic are present-day Dutch novels, and by which specific criteria? As the previous two paragraphs have shown, network theory and narratology offer different criteria by which to assess the structure, function, and meaning of communities. Conjoining the empirical approach of network theory with the theoretical perspectives of Greimas and Courtés, and the Bakhtin circle, the subsequent two analyses aim to provide an integrative account of communities in literary texts with quantitative findings entering into a dialogue with qualitative readings of two individual texts.
4.4 MODEL I: COMMUNITY DETECTION

This section presents a quantitative model to statistically assess the degree of polyphony in literary texts by means of the automatic detection of communities. Community detection algorithms can be used to break down a character network into separate clusters based on statistically significant cut-off points (see section 4.2.1 of this chapter). These clusters, then, represent communities of characters who are more strongly connected to one another than they are to characters from other communities. In light of Bakhtin’s concept of polyphony, such communities can be regarded as representing a specific social speech type or voice. In this view, a larger number of communities in a novel might indicate a higher degree of polyphony because there are simply more social speech types or voices present.

Within a single community, however, multiple voices can also coexist, as a community can consist of characters from different social, cultural, or economic backgrounds. Both lower and higher educated, migrant and nonmigrant, older and younger, male and female characters can be part of the same community. Therefore, a great variety of voices within a community can also be seen as an expression of polyphony.

In other words, there is a tension between computationally detected communities and the demographic backgrounds of the characters within those communities (i.e., the social group to which they belong – men versus women, migrants versus nonmigrants, the old versus the young, the higher versus the lower educated). Communities can be seen as distinct entities representing a collective voice, but single characters also represent distinct individual voices within communities. There are ways imaginable in which the individual voices of characters cause friction with the collective voice of the character community of which they are part of. Is it possible, for instance, to speak of polyphony when a community is extremely unbalanced in terms of gender, descent, education, or age distribution? If we define the social speech types Bakhtin refers to in terms of the social, cultural, and economic backgrounds of characters, then it is hardly possible to speak of polyphony when there are, for instance, predominantly higher educated, older, Dutch, male characters in a community. In other words: the demographic composition of a community will affect its degree of polyphony. In the following two subsections a method will be presented for clustering the character networks into separate communities. The demography of these communities will then be scrutinized further in order to get a sense of how polyphonic, or dialogic, the corpus is. Building on Bakhtin, polyphonic
Character Constellation

4.4.1 Clauset-Newman-Moore and Girvan-Newman Algorithms

None of the generic algorithms for community detection in networks, described broadly in section 4.2.1, are built specifically for analyzing character networks. It is thus a challenge to find the most suitable algorithm for detecting communities in literary texts. As there is no existing research tradition of community detection in novels, it is also not possible to fall back on best practices reported in other work.

Experiments were conducted with a variety of state-of-the-art community detection algorithms. Unexpectedly, the first experiments yielded negative results. The Clauset-Newman-Moore greedy modularity maximization algorithm (Clauset, Newman, & Moore, 2004) is based on optimizing the modularity of separate clusters of nodes (see section 4.2.1 for an explanation of modularity algorithms), and within the dataset invariably yields a number of communities equal to the size of the network. A novel with 20 characters, for instance, is broken down by this algorithm into 20 separate communities, each of which contains only one character. It is obviously not meaningful to analyze such small-size communities. The Girvan-Newman algorithm (Girvan & Newman, 2002), another state-of-the-art algorithm, aims to remove edges with high betweenness centrality to find a cut-off point between clusters (see chapter 3 ‘Centrality’, section 2, for an explanation of different centrality metrics). Applying this algorithm to the corpus results in similar problems. This algorithm yields only two communities for every book, one of which contains one single character, the other containing all other characters – such clusters cannot meaningfully be used in the subsequent analyses. Other state-of-the-art algorithms produce similarly useless clusters.

Why do state-of-the-art algorithms fail to detect meaningful clusters of characters? Some descriptive statistics on the network structures of the novels shed light on this. Table 1 shows the means for a range of basic network features, such as the number of nodes and edges, density, clustering coefficient, and triadic closure.
Table 1. Descriptive statistics of number of nodes, number of edges, density, clustering coefficient, and triadic closure in the corpus (N=170).

<table>
<thead>
<tr>
<th>Model</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>170</td>
<td>3</td>
<td>29</td>
<td>12.57</td>
<td>5.146</td>
</tr>
<tr>
<td>Number of edges</td>
<td>170</td>
<td>1</td>
<td>212</td>
<td>35.67</td>
<td>29.84</td>
</tr>
<tr>
<td>Density</td>
<td>170</td>
<td>0.06</td>
<td>1.00</td>
<td>0.46</td>
<td>0.19</td>
</tr>
<tr>
<td>Clustering coefficient</td>
<td>170</td>
<td>0.00</td>
<td>0.87</td>
<td>0.23</td>
<td>0.13</td>
</tr>
<tr>
<td>Triadic closure</td>
<td>170</td>
<td>0.00</td>
<td>1.00</td>
<td>0.64</td>
<td>0.19</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The number of nodes (i.e., characters) in novels in the corpus ranges from a minimum of 3 to a maximum of 29, with a mean of 12.57. Compared to networks used in other strands of network analysis (e.g., biological or real-world social networks), these are extremely small network sizes. One explanation of the inapplicability of, for example, the Clauset-Newman-Moore greedy modularity maximization algorithm and the Girvan-Newman algorithm is that those algorithms are designed for analyzing larger network structures. Character networks are simply too small to detect a fair number of communities with these algorithms.

Another explanation for the poor performance of these community detection algorithms is the relatively high density of the networks. Density indicates the completeness of the network; it is a measure of ‘the extent to which all possible relations are actually present’ (Scott, 2000, p. 32). A seminal study on real-world social networks reported that most personal networks have a low density, only one-fifth of the studied networks having a density higher than 0.50 (Wellman, 1978, p. 1215 as cited in Scott, 2000, p. 78). The character networks in the Libris corpus have a mean density of 0.46, meaning that they are relatively tightly knit. Probably, the high density of these networks makes it more difficult to separate the network into distinct, dense clusters. When everyone is connected to everyone, there is simply nothing but one tight community.

The low mean clustering coefficient (scale 0–1) of 0.23 strengthens this interpretation. This measure indicates the extent to which nodes in a network
cluster together. A low clustering coefficient value ‘can indicate a network with relatively small clusters compared to the overall size of the network’ (Bucur, 2019). Overall, novels in the Libris corpus contain dense character networks, but they are not easily separable into dense subnetworks. Furthermore, this is supported by the scores on triadic closure, or transitivity, which measures how interconnected a graph is in terms of the ratio of actual over possible connections. The mean triadic closure of 0.64 (scale 0–1) is considerably high, and again, just as the high density, shows how tightly knit the networks are – and thus probably hard to group into distinct communities.

4.4.2 Kernighan-Lin Bisection Algorithm

Because of the low number of characters, high density and triadic closure, and low clustering coefficient, it is not meaningful to break down the character networks into a fair number of separate communities. However, it is possible to enforce a segmentation of a novel’s social network into two clusters, for instance of equal size, by optimizing a separation criterion. The Kernighan-Lin algorithm (Kernighan & Lin, 1970) bisects a network into two clusters by ‘iteratively swapping pairs of nodes to reduce the edge cut between the two sets’.

Although this results in only two communities for each novel, it seems to be the only feasible way to group the character networks into separate clusters. This is a fairly rough clustering technique, but it relies on a statistically significant cut-off point as is the case with the Clauset-Newman-Moore and Girvan-Newman algorithms.

The Kernighan-Lin algorithm thus detects two communities of equal size for every novel. Philip Huff’s *Niemand in de stad*, for instance, contains 22 identified characters, and the algorithm separates these 22 characters in a community A and a community B, each containing 11 characters. Following Bakhtin, each of these communities can subsequently be regarded as a collective voice. In order to assess the diversity of individual voices – the extent of polyphony – within these communities, the gender, descent, age, and education distribution for community A and community B is computed for every novel in the corpus.

Solely at the level of the individual text, this bisection allows me to qualitatively determine whether or not community A and community B are segregated in terms of gender, descent, age, and education, which will be done for Huff’s novel in the close reading carried out in section 4.4.3. A hypothetical example: 20 characters were identified for novel X, of which 10 are male and 10 are female. If community A of this novel features 9 male characters and 1 female character, whereas community B consists of 9 female characters and 1 male
character, then it is likely that novel X is segregated by gender: almost all male characters are grouped together in community A, and almost all female characters are grouped together in community B.

Another hypothetical example: novel Y has 10 identified characters – 1 female character and 9 male characters. Community A of this novel contains this single female character and 4 male characters, whereas community B contains all the other remaining 5 male characters. Is novel Y segregated by gender? Contrary to the hypothetical example of novel X, the character population of novel Y is already extremely unbalanced in terms of gender (90% male, 10% female); clustering these characters into distinct communities A and B will only reflect this general unbalance. Both hypothetical examples of novels X and Y illustrate that an assessment of the multivoicedness, the polyphony, cannot rely on a comparison of community A and B mutually, as it is of paramount importance to compare the gender, descent, education, and age distributions in A and B against the general distributions in the novel as a whole.

To cater to such a comparison between communities at the corpus level, so-called 'difference scores' were computed for each individual novel by subtracting the number of characters of a specific type in a community from the number of this type in the character population of the novel as a whole. For gender, for example, the following six difference scores were computed based on absolute numbers:

- the difference between the number of male characters in community A and the novel as a whole
- the difference between the number of male characters in community B and the novel as a whole
- the difference between the number of female characters in community A and the novel as a whole
- the difference between the number of female characters in community B and the novel as a whole
- the difference between the number of characters with gender unknown in community A and the novel as a whole
- the difference between the number of characters with gender unknown in community B and the novel as a whole.

Similar difference scores were computed for descent, education, and age. These scores indicate the extent to which both community A and community B reflect the overall gender, descent, education, and age distributions in the novels as a
whole and more generally points at the extent to which the communities are segregated in terms of these demographic categories. If the difference between, for instance, male characters in community A and male characters in the novel as a whole is higher than the difference of male characters in community B and male characters in the novel as a whole, this may indicate that community A is more dominated by the male social speech type or voice than community B. Note that these difference scores are solely based on the bisection of communities within each of the 170 novels individually and have nothing to do with the composition of communities in other novels in the corpus.

Of course, this approach requires a move from the differences scores in communities of particular novels to generalizable statements about community formation and segregation at the corpus level. In order to assess the extent of multivoicedness or polyphony in the corpus as a whole, it was tested whether there is a significant difference between the difference scores of each of the 170 individual novels. In the case of gender, for instance, a test was conducted to see whether there is a statistically significant difference between the difference scores for male characters in community A and male in characters in community B, as well as between female characters in community A and community B, and between characters with gender unknown in community A and community B. If there indeed are significant differences between these scores, then this shows that the gender of characters is dependent on the type of community (A or B) which they are a part of – that the novels in the corpus, in other words, are segregated by gender.

The appropriate statistical test to conduct for these purposes is a repeated measures ANOVA, which can be used to detect overall differences between the mean difference scores of all the novels in the corpus for gender, descent, education, and age. Four repeated measures ANOVAs were conducted on the dependent variable community, one for each of the variables gender, descent, education, and age (see Appendix C for all the main and interaction effects). No statistically significant interaction between gender and community was found, \( F(1, 169) = 1.002, p = 0.318 \). There was also no statistically significant interaction between education and community, \( F(1.776, 300.196) = 2.036, p = 0.138, \varepsilon = 0.897, \text{ partial } \eta^2 = 0.012 \). The absence of statistically significant interactions between gender and community and between education and community suggests that the novels in the corpus are not segregated by gender and education.
Significant interactions were found, however, between descent and community, and between age and community. The interaction between descent and community reached significance, \( F(1.595, 269.625) = 3.521, p = 0.041, \varepsilon = 0.798 \), but with a very small effect size (partial \( \eta^2 = 0.020 \)), meaning that the size of the difference is relatively small.\(^{16}\) Breaking down these effects by the different levels of nonmigrants, migrants, and unknown descent, Figure 3 demonstrates that the mean difference for nonmigrant characters between the communities A and B in the corpus is bigger than it is for migrant characters and for characters with descent unknown (i.e., the distance between the dark purple line [community A] and the lilac line [community B] is smaller for those categories).

![Figure 3. Interaction between descent and community of all characters (N=2,137) in the corpus (N=170). Error bars: 95% CI.](image)

Something similar holds for age. A significant interaction effect between age and community was found \( F(1.905, 321.989) = 4.156, p = 0.018, \varepsilon = 0.953 \), but again with a small effect size (partial \( \eta^2 = 0.024 \)).\(^{17}\) Figure 4 shows that the mean difference for younger characters between the communities A and B and the mean difference for older characters between the communities A and B is bigger than it is for characters with age unknown.
In sum: although significant interactions effects have been found for both descent and age with community, the size of this effect is limited. Both descent and age are causes of segregation in the fictional characters’ populations, but the extent of this segregation is low.

What does this mean with regard to the degree of polyphony in the novels? As there are no remarkable differences between the gender and education distributions within the communities compared to the books as a whole, it is tempting to conclude that the majority of the novels in the Libris corpus are polyphonic or dialogic in terms of these categories. After all, the detected communities are not significantly segregated into either male and female communities or higher and lower educated communities. Overall, both male and female voices thus tend to be simultaneously orchestrated in the same communities, as do the narrative voices of the higher and the lower educated. For descent and age, however, the tests suggest otherwise: there are significant differences between the descent and age distributions within communities A and B compared to the novels as a whole. As these results indicate that the detected communities are segregated into migrant and nonmigrant characters, and into younger and older characters, it seems legitimate to state that the corpus is *not* polyphonic or dialogic in terms of descent and age. Characters with and without
a migration background tend to function within communities of their own, as do younger and older characters – connections between these social speech types are weaker than between men and women and people with different levels of education.

The results of the repeated measures ANOVA tests should be seen in relation to the negative outcomes of the tests with the Clauset-Newman-Moore and Girvan-Newman algorithms. As pointed out in paragraph 4.4.1, the in-applicability of these algorithms can be explained by basic network features of the novels, such as the low number of characters, their high density, and low clustering coefficient. But possibly, the ineffective outputs of these algorithms (e.g., a large number of extremely small size communities) are a symptom of a more fundamental problem studies on character communities face. If it is so hard to group the social networks of characters into a fair number of distinct clusters, is it even meaningful to call the bisections into two groups of characters of equal size – by the Kernighan-Lin algorithm – communities?

Perhaps not. One argument is that the small fictional worlds of present-day Dutch novels are simply communities in themselves. As the chances are relatively high that two random characters in a novel engage in some kind of a relationship, the chance that a novel can be broken down into distinct groups of characters decreases. This is exactly what is suggested by the high density of the novels: the characters are too interconnected to be meaningfully segmented into communities. Given these observations, this argument makes the case that there is just one community for every novel. From a Bakhtinian perspective, in which communities are defined as collective voices, the absence of a multitude of communities can be regarded as nonpolyphonic. Although novels with only one community might still contain a variety of narrative perspectives or social speech types, these novels are self-contained social systems in the sense that they lack a variety of collective voices. Obviously, within each single community there are interconnections between individual voices, with characters interacting on the plot level and co-occurring on the sentence level. But is this equal to a ‘connection between consciousnesses’ (Bakhtin, 1929/2003a, p. 95)? The existence of and interaction between multiple communities, each representing a collective voice, might be a necessary prerequisite for a novel to be genuinely polyphonic. Compared to novels such as those of Dostoevsky, containing a fairly large number of characters and perhaps a lower density, present-day Dutch novels are possibly of a radically different nature.
A second argument reverses the above line of reasoning. Instead of interpreting the high density of the networks as a negation of polyphony (because it hampers the clustering of characters into communities), it can conversely be seen as an argument for polyphony. Arguably, dense networks are the perfect breeding ground for a ‘connection between consciousnesses’ (ibid). In the tightly knit character networks in the Libris corpus, there is ample opportunity for characters with different social, economic, and cultural backgrounds to engage in contact. Although it is hard to group the characters into communities representing a collective voice, the interplay between their individual voices fosters a dialogic environment. In this second argument, polyphony, or dialogism, is a function of the demographic composition of the novels as a whole and not of the collective voices represented by distinct communities. The repeated measures ANOVA tests, furthermore, have shown that the corpus is not segregated into communities of male and female and into lower and higher educated characters, which suggests that there are sufficiently many moments of contact between characters from different genders and educational levels or social classes. However, the repeated measures ANOVA tests have also shown that the Libris corpus is segregated into migrant and nonmigrant communities and into younger and older characters, which suggests non-polyphony in terms of descent and age.

In order to assess the value of these statistics-based interpretations, a close reading of Philip Huff’s *Niemand in de stad* (2012) will follow evaluating the degree of polyphony in narratological terms. This novel is selected because it explicitly thematizes community life of student fraternity members; it will be read against the background of the reported quantitative observations. Do the detected communities in this novel make sense from a qualitative perspective, and how polyphonic are they in both statistical and narratological terms?

### 4.4.3 Close Reading: Communities in Philip Huff’s *Niemand in de stad* (2012)

*Niemand in de stad* has an obvious intertextual relation with Nescio’s short story *De uitvreter* (1911), one of the classics of twentieth-century Dutch literature. Just as Nescio’s characters Japi, Bavink, and Koekebakker, the circle of friends around Philip, the protagonist of Huff’s novel, is depicted in a rather strong opposition to ‘society’ as a whole. In *De uitvreter*, the work-shy main characters look down upon common, working-class people because they conform to society’s rules and norms. In *Niemand in de stad*, student fraternity life is represented as the last free haven before having to subsist under the yoke of society.
This sentiment is most accurately described in a key scene in which first-person narrator and protagonist Philip discusses student life with his fellow fraternity members Jacob and Matt. Idealizing his student years, Jacob characterizes the present moment as ‘the autumn of [their] youth’ and ‘[their] time in the Garden of Eden’ (Huff, 2012, p. 104). Furthermore, he postulates a strict opposition between fraternity and the rest of society:

People assume that there are no windows [in the fraternity clubhouse] so that they cannot look inside. But that is not the reason why. At least: it is only partly true. It is also about how visible the outside is seen from the inside. The club does not have any windows so that we do not have to look outside. We are hiding from the world, from the sheep shearsers with their knives and the hunters with their guns, from our fathers and their expectations, and from our mothers and their safety-net. For just a little while. In this small, artificial, and unimportant society of ours.

(p. 105)

Metaphorizing the absent windows in the fraternity’s clubhouse, Jacob highlights the antagonistic relation between the insiders (the fraternity members) and the outsiders (‘the world’). Remarkably, the last sentence in this quote is ambiguous: does ‘this small, artificial, and unimportant society of ours’ refer to society as a whole (‘the world’) or to fraternity life specifically? The impossibility of disambiguating between the first and the second option obstructs an analysis in terms of Greimas and Courtés’s distinction between paradigmatic and syntagmatic collectives (see section 4.3.1 of this chapter). If fraternity life is the object of reference in this sentence, then the narrative world of Niemand in de stad can be framed as a paradigmatic collective in which society as a whole is placed higher in the hierarchical order than the community of fraternity members, the latter being merely a ‘small, artificial, and unimportant’ suborder of that society. When the sentence is interpreted as referring to society in general, then the fraternity is simply part of ‘this small, artificial, and unimportant society of ours’. In this interpretation, the novel conversely depicts a syntagmatic narrative world in which there is no explicit hierarchy between the fraternity and ‘the world’ as such. Because of the sentence’s ambiguity, both options are left open.

There are, however, less ambiguous text fragments that reinforce a strong, hierarchical opposition between fraternity and society. Most obviously, this is illustrated by the fraternity’s strong emphasis on the importance of being part of a group, a collective, a community. As the most eloquent member of the group, Jacob points out that Philip and his friends should use ‘[Their] right to surrender
to the group’ (p. 27). Authenticity and individual identity, in other words, have to be sacrificed for the sake of the collective. In fraternity, Jacob contends, people are ‘the role [they] play’. ‘Yes, of course’, Philip passively complies (ibid.). One of the main goals of this collective is to remain intact. This is exemplified by Matt’s dissatisfaction with the quality of the beer in the clubhouse, about which he jokingly fears that ‘in this way, we will not be able to drink enough to feel connected’ (p. 19).

The importance of the social cohesion within the fraternity collective conflicts with the social ties Philip has with people outside of it, most importantly with his high-school girlfriend Elisabeth. When Elisabeth visits the ‘Weeshuis’, the residency of Philip and his fellow fraternity members, the clash between inside fraternity and outside is foregrounded. Hannes, Jacob, and Bart encounter Elisabeth and Philip having sex when they enter his room unannounced, after which they make ironic jokes about Elisabeth’s pubic hair. Elisabeth is, quite expectedly, not amused. But more strikingly, she seems to hold Philip responsible for her confrontation with the rudeness of his friends: ‘Elisabeth turned her face to the wall. […] When I touched her shoulder, she pushed away my hand’ (p. 33). Being part of this community, Philip indeed is in a sense responsible for the norms and values he implicitly upholds by being a member of it. One of these implicit norms is to have a variety of sexual contacts. Because he has a girlfriend, Philip has a hard time conforming to this expectation. Again, this highlights the hierarchical opposition between fraternity life and the outer world, here represented by Elisabeth, who is described by Philip as ‘[his] home’ (p. 36). In the course of the novel, the fraternity gradually replaces Elisabeth as Philip’s ‘home’.

Network visualizations of the two communities as detected by the Kernighan-Lin algorithm (see section 4.4.2) help to gain a more general insight into the novel’s segmentation into distinct groups. Figures 5 and 6 show the networks of the two detected communities.

With some exceptions, the bisection of the novel’s character network into these two communities by the algorithm is understandable from a close reading perspective. Community A (Figure 5) contains most of the fraternity insiders. Philip, Jacob, Matt, Paulus, Bart, Hannes, and Tom are all residents of the ‘Weeshuis’. Karen is also part of the student organization and has an affair with Philip. Simon is the organization’s praeses. Only Elisabeth, Philip’s girlfriend, and Tessa, Matt’s girlfriend, are not part of the same organization. However, as Elisabeth and Tessa interact frequently with both Philip and Matt, it seems logical that they are grouped into the same cluster by the algorithm.
Figure 5. Network visualization of community A as detected by the Kernighan-Lin bisection algorithm. Node size is equal for all nodes, node color indicates gender (light color = male, dark color = female). Edge size indicates weight.

Figure 6. Network visualization of community B as detected by the Kernighan-Lin bisection algorithm. Node size is equal for all nodes, node color indicates gender (light color = male, dark color = female). Edge size indicates weight.
Most of the outsiders are part of community B (Figure 6). Sasha and Henk are coworkers of Philip and Matt at their part time job at the casino. Mark and Job are both older ex-fraternity members. Peter is Matt’s older brother, and Daphne is his stepmother. Carel is Jacob’s father, and Justin is Jacob’s nonfraternity friend with whom he supposedly has a romantic relation. Eva is a stripper whom Philip encounters in Prague. Less logical from a thematic point of view is that Rosanne and Kirsten are grouped in this community as they are part of the same students’ organization. But as they have only a minor role in the novel and thus do not interact frequently with the main characters, there are probably no statistical reasons for the algorithm to group them in the other community.

Looking at these network visualizations, it is immediately clear that there is a strong difference in the extent of interconnectedness between community A and B. This is supported by the density of both networks: the insiders’ community has a density of 0.70, whereas the ‘outsiders’-community has a density of only 0.07. In other words, people belonging to the circle around Philip and his fellow fraternity members are drastically more interconnected than people less close to this circle. Narratologically, this can be explained by the simple observation that the outsiders are relatively minor characters compared to the insiders: they are not as fully characterized as the insiders and are less visibly dispersed in the narrative. As a key descriptive statistic related to the network structure, the density of the communities thus reflects the importance of social cohesion and collective spirit thematized by the fraternity members in the novel.

Is the thematic opposition between fraternity and society reflected in the gender, descent, education, and age distributions of the two communities? An opposition in terms of education or descent is not apparent. Tables 2 and 3 show that the descent and education distributions in both the insiders’ community (A) and the outsiders community (B) are close to the overall distributions in the book as a whole. This is a result of the fact that migrant and lower educated characters are fully absent in the novel. The only differences in descent and education distribution in the communities is due to the degree of descent and education that is unknown for certain characters.
<table>
<thead>
<tr>
<th>Community</th>
<th>Descent</th>
<th>Relative distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>Nonmigrant</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>Migrant</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.09</td>
</tr>
<tr>
<td>Community A</td>
<td>Nonmigrant</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Migrant</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.00</td>
</tr>
<tr>
<td>Community B</td>
<td>Nonmigrant</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Migrant</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.18</td>
</tr>
</tbody>
</table>

*Table 2. Relative descent distributions for Niemand in de stad as a whole compared to the two detected communities (A and B).*

<table>
<thead>
<tr>
<th>Community</th>
<th>Education</th>
<th>Relative distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>High</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.18</td>
</tr>
<tr>
<td>Community A</td>
<td>High</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.09</td>
</tr>
<tr>
<td>Community B</td>
<td>High</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.27</td>
</tr>
</tbody>
</table>

*Table 3. Relative education distributions for Niemand in de stad as a whole compared to the two detected communities (A and B).*
It is tempting to regard this as an indication of the absence of polyphony, or
dialogism, in terms of descent and education. After all, the novel only includes
higher educated and Dutch characters, which can be interpreted as an obstruction
of the diversity of social speech types with regard to these demographics. An
important side note, however, is that the homogeneity in terms of descent and
education in both communities is a direct result of the homogeneity of the book
as a whole. Subsequently, this overall homogeneity results in a narrative world
in which there is no possibility for a segregation into communities of higher and
lower educated, and migrant and nonmigrant, characters. When all characters
in a novel are higher educated and of the same descent, no (hierarchical)
opposition between groups of characters can arise. However, this is only true
in the most direct, literal sense and does not account for the full complexity
of the represented identities in the novel. Philip and Jacob, for instance, are
both annotated in the database as higher educated, but there is an obvious class
difference between the two: Philip comes from a broken family and receives
little financial support, whereas Jacob has a wealthy family with its own seal ring.
Reducing them both to higher educated characters does not do justice to such
differences. Statistically, there is no opposition in terms of descent and education
in the detected communities, but it would be reductionistic to conclude that
these identity categories play no part at all in the thematic opposition between
fraternity and society as presented in Niemand in de stad.

A slightly different image emerges for age and gender. Tables 4 and 5 show that
the relative age and gender distributions in the insiders’ community (A) and the
outsiders’ community (B) deviate more strongly from their overall distributions
in the novel as a whole than is the case with descent and education. The insiders’
community is relatively young and male in light of the age and gender distribution
in the novel’s narrative world as a whole. Whereas almost all characters in this
community are below the age of 25, the outsiders community has more variance in
age categories. The novel’s overall male-female ratio is 0.64–0.36, whereas 73%
of the characters in the insiders’ community are male as opposed to only 55% of
the characters in the outsiders’ community.
<table>
<thead>
<tr>
<th>Community</th>
<th>Age</th>
<th>Relative distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>&lt;25</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>26–35</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>36–45</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>46–55</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>56–64</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>65+</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.14</td>
</tr>
<tr>
<td>Community A</td>
<td>&lt;25</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>26–35</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>36–45</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>46–55</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>56–64</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>65+</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.09</td>
</tr>
<tr>
<td>Community B</td>
<td>&lt;25</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>26–35</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>36–45</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>46–55</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>56–64</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>65+</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.18</td>
</tr>
</tbody>
</table>

*Table 4. Relative age distributions for Niemand in de stad as a whole compared to the two detected communities (A and B).*
Table 5. Relative gender distributions for Niemand in de stad as a whole compared to the two detected communities (A and B).

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Relative distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>Male</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.00</td>
</tr>
<tr>
<td>Community A</td>
<td>Male</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.00</td>
</tr>
<tr>
<td>Community B</td>
<td>Male</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Although being young seems to be a crucial part of student culture, it is not specifically thematized in the novel other than through an idealization of youthful fraternity life in general such as exemplified by Jacob in the quote above. Being male does not seem to be a necessary prerequisite for entering this supposedly last free haven of fraternity life. However, *Niemand in de stad* is a story about boys rather than about girls. 'Boys we were, but nice boys,'¹⁸ the opening line from Nescio’s short story *De Titaantjes* (1911/2018), and one of the most famous lines from twentieth-century Dutch literature, constantly looms in the background. The comings and goings of Philip, Jacob, Matt, and other fraternity members should first and foremost be read in light of the book’s nostalgic yearning for a time when boys could just be boys without having to bother about careers, marriages, parenting, and mortgages.

This image of naive boyhood trickles down to the way in which the female characters are depicted. Most women in the novel function as catalysts for the burgeoning sexual desires of Philip and his friends. This is first of all apparent in the characterization of the female characters. Although physical descriptions are often a key element of characterization, the bodily features of the women in the novel are remarkably more foregrounded than those of the men. For instance, Philip introduces Elisabeth to the reader by emphasizing her appealing looks rather than her fine character traits: ‘She was small, blonde, large breasts, and wore black glasses. Especially those glasses excited me.’ (p. 29). Furthermore, the first scene she features in is the above-described uncomfortable sex scene, highlighting
her physicality rather than her psychology. Philip's fascination with Karen is also primarily a fascination with her body, which is exemplified by his description of her looks (e.g., ‘She has full, striking lips that highlight her sensuality’ [p. 71]), as well as by Matt’s reaction on Philip talking with such a ‘hot chick’ (p. 75): ‘Last night, you were having extremely lengthy conversations with those tits of Karen Ricks’ (p. 74). When Philip is worried about Elisabeth finding out about his affair with Karen, he subtly makes clear that he has an affair with her body rather than with her mind: ‘And Karen’s body was not here to calm me down’ (p. 263).

The foregrounding of the bodily features of Elisabeth and Karen does not stand on its own; it is exemplary of a pattern. Philip and his friends conform to the fraternity stereotype of striving for as much sex as possible. Although Philip is rather troubled about having an affair, the dominant morale of fraternity is accurately described by Matt stating that ‘you have to fuck those chicks out of your system. Satisfying the need’ (p. 184). When Philip and Matt visit Matt’s stepdad, this again becomes clear through Philip’s sexualization of Daphne, Matt’s stepmom: ‘She wears a short, purple dress without a bra. Her large, heavy breasts are firm and look good in the fabric’ (p. 175). In line with the present analysis, Philip is aware that this desire is part of a pattern: ‘It seems as if Karen opened up a door of desire to strange, unreachable women’ (p. 178).

Whereas descent and education are not particularly highlighted, and age only partly, as a component of the thematic opposition between fraternity life and society, gender definitely creates a distance between the insiders’ community and the outsiders’ community. Not only does the insiders’ community feature more men than the outsiders’ community, female characters also have a different status in the novel than male characters. The idealization of the free-haven of fraternity goes hand in hand with an objectification and (hetero)sexualization of female bodies. ‘Boys should be boys’ here also denotes a permission to surrender to raw, hormonal desires for sexual fulfilment. Furthermore, the frustration of this desire has the potential to result in covert aggression toward female characters. Most strikingly, this is illustrated in a scene where Philip is robbed by stripper Eva in a private session. Being frustrated in his sexual desire and upset about the robbery, Philip aggressively masturbates while thinking about Eva in terms of a ‘Bitch’ (p. 66), a ‘Goddamn whore’ (p. 66), and a ‘Stealing whore’ (p. 67). In a similar vein, Philip is hostile toward Elisabeth when she criticizes his friend Matt for having such loose sexual morals (‘That boy is a Neanderthal’ [p. 118]). As a result of his girlfriend’s criticism of the fraternity’s implicit norm of sexuality, a quasi-ironic desire to kill her emerges in his mind: ‘I feel like pushing Elisabeth. Boy kills girlfriend in hotel room’ (ibid., emphasis in original quote). As if criticism
against the collective he belongs to desexualizes, Philip then promptly loses his sexual interest in Elisabeth (p. 119).

The male voice or social speech type is arguably more dominant in *Niemand in de stad* than the female. It is narrated from the perspective of a young man who finds himself in the middle of a male-centered, masculine community in which women are primarily regarded in terms of their sexuality. Prioritizing the male perspective, the novel can be taken as being quite non-polyphonic, or non-dialogic, in terms of gender: a genuine ‘connection between consciousnesses’ (Bakhtin, 1929, p. 95) does not seem to take place between the male and the female social speech types. These qualitative observations are backed up by the novel’s algorithmic bisection into an insiders’ community and an outsiders’ community, of which the first features more male than female characters. That Elisabeth and Karen are grouped into this insiders’ community does, however, not conform smoothly with the present qualitative analysis, as these characters are not genuinely part of the community of (male) friends around Philip, and they have a rather instrumental function in the narrative.

A closer look shows that the novel deliberately seems to create this schematic opposition between groups such as fraternity and society, and men and women. This is subtly expressed through the metaphor of the casino where Philip and Matt work as a croupier, which is ironically called ‘the firm Sly & Fraud’ (p. 120). In a briefing before their work shift, their boss motivates the croupiers by saying that ‘[they] are going to unite people and bring them fun. Decorate lives’ (p. 54). This euphemism can be read as a critical reference to fraternity: claiming to unite and decorate the lives of their members, the fraternity collective possibly tricks people into avoidance behavior in the same way gambling does. Whereas Philip seems to have little problem with fraternity as such, he dabbles with the ethics of being an employee of the casino:

> We do not work here to deceive people, I say to myself in the stairway. We are here to facilitate. It is their choice to be here. (p. 111)

Repeating the logic he learned at the casino, Philip deceives himself. In truth, he knows very well that the casino fosters an environment of abuse, addiction, and avoidance. His attitude toward the casino can be interpreted as symbolizing his latent attitude toward the fraternity. Deep down, it is suggested, Philip knows that the strong opposition between fraternity life and society leads to self-deceit as well, which becomes painfully clear when his friend Jacob appears to have lived a life of deceit for years and finally commits suicide. The casino claims to help those at the margins of society (‘There is nobody in town who cares more
about those people than we do’ [p. 226]). Similarly, the fraternity community also claims to cater to young people who do not have to be part of society yet. Both the casino and the fraternity are deceitful in postulating a divide between them and society in the first place. Although it is not stated explicitly in the novel, reading between the lines shows that the narrative self-consciously plays with this schematic opposition between groups.

‘Exclusively thinking in oppositions is an indication of intellectual laziness’ (p. 134), Jacob states in one of his conversations with Philip. This can be read as hidden meta-commentary on the narrative as a whole. Statistically and narratologically, the bisection into an insiders’ community and an outsiders’ community creates a framework for analyzing the novel in oppositional terms. Such an oppositional, schematic analysis indeed shows a hierarchy in the novel’s representation of certain social groups, between fraternity and society, as well as between men and women. This very opposition, however, is also thematized in the novel through, for instance, the metaphor of the casino. As such, the narrative seems to play with these observed statistical and narratological patterns. This is not to say that these patterns are meaningless: the novel still contains an obvious opposition between fraternity and society, as well as an absence of gender polyphony. But although the novel does conform to these statistical and narratological patterns, the literary stylistic mechanisms of metaphors, symbols, and subtle metacommentary open up possibilities for deconstructing those same patterns. In this sense, the oppositions can be interpreted as a form of critical mimesis: although the opposition between fraternity and society is represented as highly schematic, such literary stylistic mechanisms call into question the moral scheme associated with it.

4.5 MODEL II: HOMOPHILY

Whereas the former section studied polyphony in character communities by detecting subgroups of characters in all of the 170 networks, the present section focuses on the similarities between individual characters. The current analysis thus reverses the order of the analysis carried out in the former section. Instead of breaking down the networks into smaller portions and then computing frequency distributions of gender, descent, education, and age of those portions, the point of focus now lies on the similarity of these demographic categories between any
two characters sharing edges. As such, this section presents an alternative model for analyzing polyphony in narrative communities to the model proposed in the previous section. Whereas the former analysis approached polyphony in terms of top-down detected communities, the present analysis studies polyphony through the concept of homophily (see section 4.2.2). A high degree of homophily on gender, descent, education, or age indicates that a novel is segregated with regard to these demographic categories. In a novel where, for instance, older characters are mostly connected to other older characters, there is no dialogic interaction between characters from different age groups. High homophily thus suggests a low degree of polyphony: there is not so much a ‘connection between consciousnesses’ (Bakhtin, 1929/2003a, p. 95) in terms of, for instance, older and younger characters. Conversely, low homophily can be framed as indicating a high degree of polyphony. When connections between older and younger characters occur relatively often, there is arguably a greater dialogic interaction between characters with different ages, and thus a higher degree of polyphony with regards to age.

An important side note is that this operationalization of polyphony via the statistical metric of homophily is obviously based on a pragmatic, but reductionist, take on the voices or social speech types present in a novel. A character may inhabit many voices that are not necessarily reducible to their demographic features: youthful characters, for instance, can speak with an older voice inherited from their ancestors. In order to gain a fuller account of the multivoicedness in a novel, the narratological evaluation of the statistical results devotes attention to such voices falling outside the model’s scope. In the following, the methodological design of the analysis is first described, after which its results are reported. In order to assess the relevance of these results, a novel from the corpus is read through the lens of homophily related to age. *Liefde heeft geen hersens* (2012) by Mensje van Keulen is selected as a case study because of its thematization of youth, old age, and death, as well as for its problematization of fixed age categories.

4.5.1 Dyad Assortativity

Homophily in networks can be studied by computing the assortativity for specific node attributes such as gender, descent, education, and age (see section 4.2.2). For each of the 170 character networks, the Python software package Networkx was used to compute four so-called assortativity coefficients related to gender, descent, education, and age of the characters. These assortativity coefficients are
the Pearson correlation coefficients for each dyad of characters sharing edges. The result is a number between -1 and 1, with positive values indicating a correlation between characters with similar gender, descent, education, or age, and negative values indicating a correlation between characters with different gender, descent, education, or age. For instance, the assortativity coefficient related to gender for *De lichtekooi van Loven* by Ineke van der Aa is -0.05. If the gender assortativity coefficient were 1, then this novel would feature exclusively same-sex relations (male-male and female-female). Conversely, if the gender assortativity coefficient was -1, then the novel would only feature opposite-sex relations (male-female, female-male). In reality, *De lichtekooi van Loven* has a gender assortativity coefficient close to 0, indicating that both same-sex and opposite-sex relations occur relatively equally.\(^{21}\) Table 6 shows the means for each of the four computed assortativity coefficients in the corpus. On average, the assortativity coefficients for gender (-0.11), age (-0.07), and education (-0.06) show negative values, indicating that the corpus contains more character pairs differing in gender, age, and education. For descent, this is the other way around: the positive mean of the assortativity coefficients (0.18) indicates that there are more pairs of characters with the same region of descent.\(^{22}\)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender assortativity</td>
<td>170</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.11</td>
<td>0.21</td>
</tr>
<tr>
<td>Age assortativity</td>
<td>170</td>
<td>-0.60</td>
<td>1.00</td>
<td>-0.07</td>
<td>0.26</td>
</tr>
<tr>
<td>Education assortativity</td>
<td>170</td>
<td>-0.50</td>
<td>1.00</td>
<td>-0.06</td>
<td>0.27</td>
</tr>
<tr>
<td>Descent assortativity</td>
<td>170</td>
<td>-0.61</td>
<td>1.00</td>
<td>0.18</td>
<td>0.48</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 6. Descriptive statistics of the assortativity coefficients of gender, age, education, and descent in the corpus (N=170).*

How to evaluate the meaning of these numbers? Section 4.2.2 of this chapter described general findings of research on homophily. Homophilous associations with regard to race and ethnicity have demonstrated to be most prominent in
contemporary societies (Marsden, 1987, 1988). The positive mean assortativity coefficient for descent suggests that segregation on this demographic factor is also apparent in present-day Dutch literary fiction. However, this segregation pattern is probably less dominant than it is in present-day Dutch society, which is demonstrated by a previous comparison of a part of the present dataset with a dataset on real-world networks of Dutch people showing that the fictional networks are significantly less segregated by means of descent than real ones (Volker & Smeets, 2019).

Divides in society are also caused by homophilous associations with regards to education (Marsden, 1987; Verbrugge, 1977). The negative mean of the assortativity coefficient for education suggests that this is less so in the fictional worlds of Dutch characters. Age homophily in society is rather high in specific networks such as marriages, but differs depending on setting (McPherson et al., 2001, pp. 424–425). In Dutch literary fiction, age homophily is more on the lower than on the higher end of the spectrum. Segregation in society is the least fueled by gender, part of this is due to opposite-sex marriages (Marsden, 1988). On average, opposite-sex relations are more present in the corpus than same-sex relations.

Still, it is up for debate how the findings on gender, descent, education, and age homophily in the corpus should be assessed. Which scores on these assortativity coefficients are reasonable to expect? Research on homophily in real-world networks is one of the possible baselines. Given such real-world findings, how extraordinary are the homophily patterns in Dutch literary fiction? This baseline is used in the comparison above and is operationalized and analyzed – for a portion of only 65.4% of the present dataset – in Volker and Smeets (2019). Testing the present results against results for social networks of actual people presupposes a particular idea about the extent to which the societies portrayed in literary fiction resemble real social structures. How reasonable is it to expect literary fiction to mirror segregation patterns in societies? Leaving this question aside for now, a more neutral and formal baseline is to conduct a so-called permutation test, or randomization test,\(^{23}\) to estimate the chance that the observed mean positive degree of descent assortativity (0.18) and the observed mean negative degrees of gender (-0.11), age (-0.07), and education (-0.06) assortativity are found given the same descent, gender, age, and education ratios of the nodes and the same number of edges. The goal of this permutation test is to single out the possibility that the found assortativity values are just a random effect of the distributions of these classes. The fact, for instance, that there are more men than women in the corpus (a 60:40 ratio), increases the chance of
finding more male-female pairs than female-female pairs. The permutation test is conducted by randomly reassigning the descent, gender, education, and age labels of the characters in the corpus 1,000 times, while keeping the ratios the same. The male-female ratio, for instance, remains 60:40, but the specific characters who are male or female change with every random permutation. Then, for every of the 1,000 random permutations, the assortativity coefficients are recalculated. Finally, the assortativity values as observed in the actual dataset are compared to the values resulting from these random permutations. The question, then, is whether or not the actual assortativity values differ significantly from these random assortativity values. If it does, then this is an indication that the actual assortativity values are not just a random effect of the gender, descent, education, and age ratios.

Following the approach by Kraicer and Piper (2019), 1,000 permutation tests were conducted, and the assortativity coefficients for each of these permutations were calculated. Then, the means of the 1,000 permuted assortativity coefficients for gender, descent, education, and age were computed. In four one sample t-tests, the actual means – the gender, descent, education, and age assortativity coefficients for the actual dataset – were compared with these four permutation means as a baseline to determine whether the actual means are significantly different from these permutation means. The mean gender assortativity of the novels in the corpus \(M = -0.11, SD = 0.21\) is slightly higher than the permutations’ mean gender assortativity of \(-0.12 (SD = 0.02)\). There is, however, no statistically significant difference between the actual observed mean gender assortativity in the corpus and the permutations’ mean gender assortativity, \(t(169) = 0.55, p = 0.586, 95\% \text{ BCa CI } [-0.23 \text{ to } 0.04]\). This suggest that the extent of gender homophily in the corpus is likely to be just a random effect of the 60:40 male-female ratio of characters. In other words: the found negative gender homophily seems to be from the result of the relative amount of male and female characters rather than signaling hetero- or homonormative gender interaction patterns.

This is different for descent, education, and age. For descent, there is a strong difference between the actual mean assortativity and the assortativity of the 1,000 permutations. Whereas the actual descent assortativity in the corpus is a positive value of \(0.18 (SD = 0.48)\), indicating more interactions between characters with the same descent, the mean descent assortativity of the 1,000 permutations shows a negative value of \(-0.11 (SD = 0.018)\), which is a statistically significant difference \((t(169) = 8.010, p < 0.0001, 95\% \text{ BCa CI } [0.23 \text{ to } 0.37])\). As descent homophily does not appear to be a random effect of the ratio of non-migrants and migrant characters in the corpus \((\pm 90:10, \text{ excluding the portion of characters with})\)
descent unknown), the interactions between characters with the same descent thus seems to be a signal of segregation by descent. Characters with the same descent, in other words, tend to flock together.

The mean education assortativity \((M = -0.06, SD = 0.27)\) is higher than the mean education assortativity of the 1000 permutations of \(-0.13 (SD = 0.012)\), which appears to be significantly higher \((t(169) = 3.734, p < 0.0001, 95\% \text{ BCa CI [0.04 to 0.12]})\). Something similar holds for age: the mean age assortativity \((M = -0.07, SD = 0.26)\) is significantly higher than the mean age assortativity of the 1000 permutations \((M = 0.15, SD = 0.009), t(169) = 3.647, p < 0.0001, 95\% \text{ BCa CI [0.04 to 0.10]}\). In the narrative worlds of the 170 novels, the interactions between characters from both different age groups and different educational levels are thus not a random effect of the age and education distributions. On average, there is significantly more integration and less segregation between ages and classes in the corpus.

Another way to evaluate the extent of homophily is by looking at features related to the network structure. To what extent do network structure features such as network size (i.e., the number of characters in a novel) and density predict the extent of gender, descent, education, and age assortativity? Before this question can be answered, the Pearson correlations between all the variables related to the network structure of the novels were computed in order to see how they are interrelated (see Appendix D). Figure 7 shows a scatterplot matrix representing the correlations between the number of words of the novel (‘tokencount’), number of characters (‘nrnodes’), number of edges between these characters (‘nredges’), the density, the triadic closure, and the clustering coefficient. A more diagonal fit line between a variable on the X axis and a variable on the Y axis represents a stronger correlation – a positive correlation when the line goes upward, a negative correlation when the line goes downward.
Figure 7. Scatterplot matrix showing the correlations between the following variables relating to network structure: tokencount, number of nodes, number of edges, density, triadic closure, and clustering coefficient.

Unsurprisingly, the variables that were used to explain the problem of detecting communities in section 4.4.1 show a significant correlation. There is a negative correlation between density and number of characters \((r = -0.352, N = 170, p < 0.01)\), meaning that the density of the character networks decrease when there are more characters present. This is understandable: for example, a Dostoevsky novel with a great number of characters might be less fully interconnected than a novel with just a few characters. There are positive correlations between density and triadic closure \((r = 0.713, N = 170, p < 0.01)\), as well as between density and clustering coefficient \((r = 0.713, N = 170, p < 0.01)\). More dense networks are thus more fully interconnected in terms of triadic relations, and
show more clustering. Whereas the positive correlation between density and triadic closure is imaginable, it would have been more logical if density and clustering coefficient would have correlated negatively instead of positively, as a combination of high density and low clustering was given as an explanation for the problem of community detection in section 4.4.1. The variables number of nodes and number of edges show multicollinearity ($r = 0.862$, $N = 170$, $p < 0.01$), meaning that they are so strongly correlated that it does not make a difference whether one or the other is used as a predictor in a multiple regression analysis. Therefore, only the number of nodes is used in the subsequent analysis.

Does the likelihood of gender, descent, education, and age homophily increase when there are more characters in a novel, when a novel has more words, or when a novel has a higher density, clustering coefficient, or triadic closure? Hypothesizing that this is the case, four multiple linear regressions were conducted to predict 1) gender assortativity, 2) descent assortativity, 3) education assortativity, and 4) age assortativity based on the tokencount, number of nodes, density, triadic closure, and clustering coefficient. No significant results were found for descent assortativity, education assortativity, and age assortativity. As features related to the network structure, tokencount, density, triadic closure, and clustering coefficient are thus no predictors for the extent of descent, education, and age homophily. It is not the case that, for instance, a denser character network leads to, for instance, more same-age character pairs.

A significant regression equation was only found for gender assortativity ($F(5, 7135) = 2.361$, $p = 0.040$), with an $R^2$ of 0.067 and number of nodes as the only significant predictor. The predicted gender assortativity is equal to a $B$ value of $-0.107 + 0.008$ (number of nodes) (see Table 7), indicating that gender homophily increases for each additional character in a novel. The number of characters in a novel is thus a predictor of the degree of gender homophily. More characters result in more same-sex pairs. The chance of gender segregation, in other words, increases when the fictional population of a novel is larger. This pattern is not easily interpretable. In nonfictional, real-world social networks, gender homophily – or gender segregation – is relatively low, mostly because of the dominance of heterosexual romantic relations (Marsden, 1988). Compared to the descent, education, and age homophily, the extent of gender homophily is also relatively low in the fictional networks in the corpus. Why does it increase when more characters are staged?
Table 7. Linear model of predictors for gender assortivity. Only number of nodes produces a statistically significant effect.

<table>
<thead>
<tr>
<th></th>
<th>Unstandardized B</th>
<th>Coefficients std. error</th>
<th>Standardized coefficients Beta</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>-0.107</td>
<td>0.016</td>
<td>-6.673</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>C Token Count</td>
<td>4.803E-8</td>
<td>0.000</td>
<td>0.039</td>
<td>0.427</td>
<td>0.670</td>
</tr>
<tr>
<td>C Number of nodes</td>
<td>0.008</td>
<td>0.004</td>
<td>0.205</td>
<td>2.067</td>
<td>0.040</td>
</tr>
<tr>
<td>C Density</td>
<td>-0.145</td>
<td>0.129</td>
<td>-0.129</td>
<td>-1.124</td>
<td>0.263</td>
</tr>
<tr>
<td>C Triadic closure</td>
<td>0.068</td>
<td>0.126</td>
<td>0.061</td>
<td>0.541</td>
<td>0.589</td>
</tr>
<tr>
<td>C Clustering coefficient</td>
<td>0.140</td>
<td>0.146</td>
<td>0.087</td>
<td>0.958</td>
<td>0.340</td>
</tr>
</tbody>
</table>

For example, *Kunstroof* (2012) by Raymond Rombout has 29 identified characters, which is the highest number of characters in any of the novels in the corpus. Its gender assortativity coefficient is 0.007, which is relatively high in light of the mean gender assortativity coefficient of -0.11. In this novel, conforming to the general pattern, a larger population of characters goes hand in hand with more gender segregation. Why? Any theoretical explanation of this pattern seems bound to speculation as it does not relate to intuition or any literary, cultural, sociological, or other kind of theory. It seems that only a qualitative reading of this particular novel could inform the observed gender homophily. Possibly, certain thematic or stylistic dimensions of this novel might clarify why it is likely that more same-sex interactions emerge when more characters are staged.

More generally, the regression analyses carried out in this section invoke the image that it is challenging to find a sound, data-driven explanation for the degree of gender, descent, education, and age homophily in literary fiction. This is best exemplified by the lack of significant results for the regression analyses in which it was tested whether or not tokencount, density, triadic closure, and clustering coefficient are predictors of descent, education, and age homophily.
Furthermore, the observed effect of the number of characters on the extent of gender homophily is statistically significant, but is hard to explain from a literary-theoretical point of view. This suggests that features related to the structure of the character networks do not seem to offer sound explanations for the extent of homophily in the corpus.

In order to evaluate the statistical patterns narratologically, the close reading below offers a qualitative assessment of homophily, and subsequently of polyphony, in terms of form and content. In the following section, the extent of age homophily in Mensje van Keulen’s *Liefde heeft geen hersens* (2012) is assessed in light of the observed mean age homophily in the corpus as a whole.

### 4.5.2 Close Reading: Homophily in Mensje van Keulen’s *Liefde heeft geen hersens* (2012)

In which ways can love bridge the gap between the young, the old, and the deceased? This question is raised by Van Keulen in *Liefde heeft geen hersens* without giving a final answer. The question can be reframed in Bakhtinian terms as: how can love create a dialogic interaction between, or a polyphony of, characters from different age groups? As it is a central demographic category in the narrative world portrayed by Van Keulen, the primary focus of the current close reading is on age. How does the detected age homophily in the novel relate to its thematization of youth, old age, and death? Are there thematic or stylistic dimension that might explain or inform the extent to which the characters in the novel are segregated, or integrated, by means of their age?

The novel alternates between the first-person narrations of Romy, a (probably) middle-aged employee at a funeral home, and Harro, the forty-three year-old concierge of Romy’s apartment block who is secretly obsessed with her. A simple detective-like plot kick-starts the narrative when Romy finds her eighty-year-old neighbor Irma dead at her home. Romy suspects her son Christian, Harro suspects Romy. At the end of the novel, it is still unclear whether Irma was a murder victim or died a natural death. It does not really matter who did it, as the plot line primarily serves to illustrate the interrelations between a variety of characters. Irma’s death invokes in Romy a reflection on her problematic relation with her grown-up son Christian and daughter Blanca, as well as on her relationship with her deceased abusive husband, Louis. For Harro, Irma’s death fuels his obsession with Romy and makes him reflect on his cohabitation with his old-aged mother.
What does it mean to be either young or old in this narrative world? At first sight, the novel foregrounds stereotypes related to people of a certain age. Neighbor Irma is focalized by Romy as a stereotypical old lady who is physically impaired (‘Her eyes and ears, her memory, it seems as if she succumbed to old age’ [p. 7]), querulous (‘She complains and growls’ [ibid.]), and suspicious (‘It is annoying that she became suspicious, she already accuses me of theft’ [ibid.]). Something similar holds for Harro’s mother, who is described in Harro’s chapters as a stereotypical old lady who is xenophobic (‘They invade our country, they are impolite, they are mostly analphabetic, but they feel superior because of their joyless religion that brainwashes them, because of that they disregard the original inhabitants and harass them’ [p. 67]) and longs for death (‘She says she is currently always cold, that this is a precursor of death, which she says to welcome’ [p. 128]). Conversely, the young Christian and Blanca are portrayed by Romy as immature and in need of parental protection. Romy feels that she must talk her daughter Blanca out of her romantic relation with a man who is twenty years older than she is. Given her son Christian’s robbery of Irma when he was young – ‘a youthful indiscretion’ (p. 8), according to her – she fears that her son Christian is the one who robbed and murdered Irma.

A closer look at the age representation in both Romy’s and Harro’s first-person narrations shows that such stereotypes do not cover the full extent of what it means to be either young or old in this novel. In light of Greimas and Courtés’s concept of the hierarchically ordered, paradigmatic collective (see section 4.3.1), the position of the characters in the age hierarchy keep shifting depending on the perspective taken. On a scale from birth to death, Romy and Harro are probably positioned somewhere in the middle. But Romy is old in the eyes of her son (‘This not appropriate for older people’ [p. 51]), whereas she sees him as young (‘For young people such lines are both an opinion and an invocation’ [p. 48]). However, she still views herself as a ‘young widow’ (p. 71) as she lost her husband at a relatively young age. Conversely, she characterizes herself as old when she utters the wish to go to ‘a grand cafe where not only young people come’ (p. 8). Furthermore, she is still among the living and is thus further to the left on the birth-death spectrum than the deceased Irma and Louis. Harro is young seen from the perspective of his elder mother but is then relatively old to still live in the same house with her. Being young or old is not a fixed label but a matter of perspective in this narrative world. Subsequently, the position of the characters in the narrative’s paradigmatic collective in terms of age is not set in stone but is fluid. In order to understand the ways in which characters from different age groups connect with one another, the statistically detected age
homophily of this novel, i.e. the assortativity coefficient for age, can be compared with the mean age assortativity coefficient of the corpus (see Figure 11 in section 4.5.1). Interpreting the novel’s age assortativity coefficient with the mean age assortativity coefficient for the corpus as a baseline, it is possible to describe the extent to which this particular case deviates from or conforms to the overall pattern; to determine, in other words, how unique or peculiar the narrative is in this respect. The mean age assortativity coefficient is -0.07, indicating a weak negative age homophily for the corpus as a whole (with 1 indicating same age group relations only, and -1 different age group relations only). On average, relations between characters from same age groups and different age groups occur relatively equally in the corpus, as the mean age assortativity coefficient is close to 0.26. The age assortativity coefficient for *Liefde heeft geen hersens* is -0.23, indicating that there are more connections between characters with different ages than is the case in the corpus as a whole. Numerically, there is thus more integration and more dialogic interaction between younger and older characters than the overall trend indicates, suggesting that there is a relatively strong polyphony of different age groups.

These statistical findings are compatible with the novel’s thematization of bridging the divides between the young, the old, and the deceased. ‘Love has no brains’: the repeatedly mentioned key sentence, and the title of the novel, illustrates this perfectly. Although quarrels, dissatisfactions, and traumas have resulted in divides between certain characters, the story demonstrates how love has the potential to create a genuine ‘connection between consciousnesses’ (Bakhtin, 1929/2003a, p. 95). The middle-aged Romy has a hard time reaching out to both younger, older, and deceased characters. Her daughter Blanca left the house to live with an older man in the same apartment block, who Romy despises because of his tattoos and his unhealthy lifestyle, focalized by her as a ‘giant of meat and fat, who lets her [Blanca] take care of him, and commands her’ (p. 88). Romy cannot seem to get in contact either with Blanca or with her son, Christian, who does not want to talk with her about the deceased Louis, her abusive husband and his abusive father. In response to one of Christian’s rants about his father, Romy suggests that his fierce attitude is due to his age: ‘If I am I right, you are out of puberty now’ (p. 60). ‘Age has nothing to do with this’, Christian answers, ‘there are plenty of old men who still get crazy when they think about their father’ (p. 60). But although Romy is not able to connect with her children through her words and actions, a strong, biological urge ties her to them: ‘Awful, most awful, is that I cannot see my children anymore when I die’ (p. 74).
In a similar vein, Romy’s relation with both the older (and then deceased) Irma and the deceased Louis is problematic. As described above, Romy focalizes Irma as querulous and suspicious and tries to cover up her death, which she thinks was caused by Christian. But in covering up her death, she is respectful and loving to Irma’s dead body, promising her to take care of her cat (p. 76). The relation between Romy and her deceased husband Louis is even more problematic, as he physically and psychologically abused her for years. Despite this, however, love keeps connecting her with him: ‘I did love him’ (p. 60).

Harro primarily interacts with his old, aged mother and Romy. Living together with his mother gets on his nerves to such an extent that he fantasizes about killing her: ‘The thought gives me visions. Whistling painters in the house, the smell of paint, a terrace in the garden, a benevolent silence, but also music, a smile’ (p. 118). The fantasy almost becomes reality at the end of the novel when he pulls her out of her chair and hits her in response to her asking him about his whereabouts. But when he sees her lying on the ground, a feeling of affection washes over him: ‘Mother...” I caress her cheek. “Oh, what a soft skin you have’ (p. 189).

Love thus indeed bridges divides between characters from different age groups. It bridges the divide between Romy and the younger Christian and Blanca, and between Romy and the older, or deceased, Irma and Louis. It also bridges the divide between Harro and his older mother. However, it does not bridge every divide in the novel. In light of the relatively negative detected age homophily of -0.23, it is remarkable that the only unbridged divide is between Harro and Romy, who are close in age (Harro is forty-three, Romy’s age is not made explicit but is probably somewhere between 45–55). As the tie between Romy and Harro is quite homophilous in terms of age, their separation conveniently supports the argument that the novel primarily integrates people with different ages and not people who are of a relatively equal age. The narrative structure with the alternating perspective of Romy and Harro caters to the separation between these two characters. It becomes gradually apparent through Harro’s focalization that he is obsessed with Romy. In trying to get close to her, he steals objects belonging to her, such as her glove and a drinking glass she used (p. 111). Conversely, Romy’s perspective shows that she does notice Harro positively (‘Harro takes off his jacket, it strikes me how muscled he is’ [p. 107]), but she does not notice or answer his love interest in her at all. The novel climaxes when Harro secretly observes Romy through a security camera in the elevator and catches her making out with a strange man (p. 179). In his despair, his obsession transforms into a revengeful attitude when he thinks about turning her in for covering up Irma’s
death: ‘An autopsy. If there is nothing to prove, then seeds of doubt at least have been sown’ (p. 185).

This qualitative assessment of the divides between characters in Van Keulen’s novel can explain this extent of segregation better than the multiple linear regressions conducted on the whole corpus (see section 4.5.1). Statistically, the non-homophilous ties between the younger and older characters result more in an integration than a segregation of different age groups. The middle-aged Romy differs in age from her children, as well as from the older and deceased Irma and Louis. The same holds for the forty-three-year-old Harro, who is differentiated from his mother by means of his age. However, the novel’s negative homophily of -0.23 suggests that there are more relations between older and younger characters than the overall trend in the corpus indicates. Thematically, one of the driving forces of this integration is the love motif through which the non-homophilous associations between the above-mentioned characters are bridged. In light of real-world studies arguing that homophily tends to lead to integration rather than segregation, the relation between Romy and Harro serves as a counterexample. As Romy and Harro are close in age, their interaction in the novel’s character network subsequently leads to a rather homophilous relation. However, there is a strong thematic divide between them, as Romy does not answer Harro’s love interest. The extent to which love plays a role in either the segregation or integration of people with different demographic profiles can hardly be taken into account in real-world studies on homophily. In the narrative world of Liefde heeft geen hersens, however, love appears to connect characters who are non-homophilous in terms of their age, while the absence of love poses a divide between characters who are demographically similar.

Parallel to this love motif, a death motif furthers the divide between Romy and Harro. Both Romy and Harro are confronted with death, in the first place through Irma’s passing away, but they have a diametrically opposed stance toward it. Romy is confronted daily with death at her work at the graveyard but is extremely avoidant and fearful of it ever since she was little:

As a child I had to bite in my blanket in order to not scream because of my death agony. I could not endure words such as death, dying, end, and when I read those words today, something in my eyes pushes them away. I turned away from funeral cars, let alone that I dared to look at it. (p. 16)

Her work at the graveyard does not normalize her feelings toward death at all. Quite the contrary: ‘I don’t know if I can keep working here, I don’t want to be
smothered by death’ (p. 42). Conversely, Harro’s ultimate fantasy is to be joined in death with Romy: ‘To be lying with her, finally, on that graveyard, not far from that arcade and the luminous grass …’(p. 188). Furthermore, he more than once fantasizes about killing his mother (p. 118). Death, in other words, has a totally different connotation for Harro than for Romy. Whereas Romy cannot bear or accept the idea of dying, Harro would be glad to die alongside Romy and to send his mother to the afterworld.

How do these thematic cues relate to the observed causes of homophily in the real-world? Geographic space and family structure are two important causes of real-world homophilous associations (see section 4.2.2). For age, the spatial environment of the classroom most obviously fosters homophily, but families are usually non-homophilous in terms of age as both children and parents live together. In *Liefde heeft geen hersens*, geographic space seems to foster the non-homophilous age relation between neighbors Romy and Irma. Quite obviously, living in the same neighborhood is not a guarantee of having contacts with people from the same age, as both older and younger people can live in an apartment block. Family structure is a cause of the non-homophilous age relation between Romy and her children, as well as between Harro and his mother: families usually consist of both older and younger people. The love motif ties in nicely with this: it can be argued that geographic space and family structure are a perfect breeding ground for the love that bridges the initial divides between Romy and her children, Romy and Irma, as well as between Harro and his mother.

Whereas geographic space seems to fuel the non-homophilous age associations between Romy and her children, and between Romy and Irma, it simultaneously creates a fertile environment for the homophilous age relation between Harro and Romy. As Harro is the concierge of Romy’s apartment block, they meet frequently. But geographic space falls short of explaining the thematic divide between Romy and Harro. Their narrative perspectives do not converge smoothly just because they see each other on a regular basis in the apartment block. This divergence can very clearly be traced back to the absence of mutual love. More subtly, their conflicting stances toward death – Romy avoids its, Harro is open to it – reinforces their divergence of souls. Putting this love motif and death motif in dialogue with the found statistical pattern of age homophily shows that the thematic dimension of the novel is an important force behind the extent of segregation or integration. Statistically, the age assortativity coefficient of -0.23 suggest a relatively high polyphony of characters from different age groups in this novel. But this is only partly backed up by the present narratological analysis of the novel. The thematic structure is obviously not taken into account in the data-
4.6 CONCLUSION TO THIS CHAPTER

In this chapter, an attempt was made to gain a closer insight into the ways in which community co-shapes the representation of social groups in present-day Dutch literary fiction. Building on classical social theory by Tönnies, Durkheim, and Simmel on what it means for people to be united in a community, it explored both network theoretical and narratological concepts and techniques to study community in narrative fiction. For the analysis of the literary representation of social groups, community proved to be particularly useful for assessing how integrated or segregated these groups are. Building on Bakhtin’s concept of polyphony, two models were presented through which the extent of integration or segregation in the corpus could be analyzed statistically, and close readings of two novels followed to evaluate the meaning of the statistical patterns. Based on the output of these models and their narratological evaluations, at least two conclusions can be drawn as to how community co-shapes the literary representation of social groups in contemporary Dutch literature.

First, segregation by descent and age was found through use of the first model. Based on a community detection algorithm, the model divided each of the 170 novels in the corpus in two distinct groups and computed the gender, descent, education, and age distribution for each of these groups. Running a range of statistical tests revealed that these communities are not segregated by gender or education, but rather by descent and age: characters with the same descent (either migrant or nonmigrant) and from the same age group (either between age 0–45 or from age 46 and upward) tend to flock together. These patterns of segregation suggest that characters from different descent and age groups tend to be represented in oppositional terms. Statistically speaking, a line can be drawn between migrant and nonmigrant characters and between older and younger characters. Based on these results, it can be argued that these groups of
characters are depicted as distinct, separate entities in present-day Dutch literary fiction.

A close reading of communities in Philip Huff’s *Niemand in de stad* was conducted to assess the meaning of these results at the level of the individual text. As this novel presents a strong opposition between an insiders’ community of fraternity members and an outsiders’ community, this case study was selected to assess which patterns of segregation the model detected in a novel thematizing divides between groups. Because of its absence of lower educated and migrant characters, no segregation by education or descent takes place between the characters populating the novel. Conforming to the overall pattern for the entire corpus, the novel’s communities are segregated by age, but contrary to this overall pattern, these communities are most notably segregated by gender. Although an opposition between male and female characters is highlighted by the thematic structure, the novel seems to self-consciously reflect on its staged oppositions through literary stylistic mechanisms such as metaphors, symbols, and metacommentary.

Secondly, an alternative take on community was presented in the second model, which computed for every two characters how similar they are in terms of gender, descent, age, and education. This so-called homophily, or assortativity, score can be seen as an indication of segregation on the level of any two individual characters. Based on a range of statistical tests, it was argued that segregation by descent, education, and age is apparent in the (non-)homophilous associations between characters. These segregation patterns are partly in line with the first model’s finding that descent and age cause divides between characters, but this second model suggests that divides are also caused by education. Gender is in neither of the two models put forward as cause for divides. According to the second model, characters with different educational levels are also represented as distinct, separate entities, just as characters from different descent and age groups (as the first model also suggested). This finding, indeed, only holds true in a statistical sense, and it remains open as to what the meaning of this overall trend is in the context of an individual novel. A qualitative, narratological assessment of segregation by age in Mensje van Keulen’s *Liefde heeft geen hersens* demonstrated how a particular narrative relates to such a statistical pattern. As one of the novel’s main themes is age and everything associated with it, it proved to be a useful case study to analyze against the general backdrop of the statistical trends in the corpus as a whole. In general, Van Keulen’s novel is an outlier: while on average segregation by age in the corpus is high, it is relatively low in this novel. At first glance, the novel indeed thematizes integration between people
from different age groups. A closer look at the love motif in the novel shows that while divides between the young and the old are bridged for specific characters, a divide is installed between two characters from the same age group. This close reading highlights the importance of love in bridging or posing divides between characters from different social groups.

Emphasizing the interrelations between the macro and the micro levels of literary representation, this chapter has shown how the notion of community plays a part in the depiction of social groups in present-day Dutch literary fiction. Confronting statistical trends for the corpus as a whole with close readings of individual novels, it demonstrated how segregation of communities affects the representation of social groups. Most notably, the extent to which groups are segregated by a certain demographic category reflects the extent to which such groups are presented as either distinct or connected.