



UvA-DARE (Digital Academic Repository)

Semantic versus lexical gender

Synchronic and diachronic variation in Germanic gender agreement

Kraaikamp, M.

Publication date

2017

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Kraaikamp, M. (2017). *Semantic versus lexical gender: Synchronic and diachronic variation in Germanic gender agreement*. [Thesis, fully internal, Universiteit van Amsterdam]. LOT.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter IV

The diachrony of semantic gender agreement: findings from Middle Dutch^{*}

Abstract

This chapter presents the results of a corpus study of pronominal gender agreement in Middle Dutch. In present-day Dutch and in several other Germanic varieties, pronouns show semantic gender agreement that is based on the degree of individuation of the referent. Dutch pronouns show variation between this type of agreement and lexical gender agreement. This study investigates how old semantic agreement based on individuation is. In particular, it aims to answer the question whether semantic agreement has developed in response to the change from the Germanic three-gender system to a two-gender system or dates back to before this change. The results show that agreement based on individuation already existed in Middle Dutch, when the original three-gender system was still in place. This shows that this type of agreement did not develop in response to the change from three to two nominal genders. The semantic interpretation of the genders along the lines of individuation apparently existed already and could be an old Germanic, possibly Indo-European, feature. What seems to have changed over time is the proportion of semantic to lexical agreement, as semantic agreement appears to occur more frequently in present-day Dutch than in Middle Dutch. This shift in agreement preference may be due to the loss of adnominal gender marking and the resulting reduced visibility of lexical gender in the noun phrase.

1. Introduction

This study is about gender agreement and the variation that exists, both synchronically and diachronically, between ‘lexical gender agreement’ and ‘semantic gender agreement’. Lexical gender agreement is agreement that is based on the lexically stored gender of a noun, while semantic gender agreement is agreement that is based on certain properties of a referent, such as animacy or sex. Languages differ with respect to the type of gender agreement they display. In

^{*} A slightly different version of this chapter has been accepted for publication as: Kraaikamp, Margot. To appear. The diachrony of semantic gender agreement: findings from Middle Dutch. *Journal of Germanic Linguistics*.

CHAPTER IV

English, for instance, pronouns show exclusively semantic gender agreement: masculine and feminine pronouns are used for male and female referents respectively, and neuter pronouns are used for inanimate referents and sometimes animals. In Dutch, on the other hand, each noun has a lexically stored gender, and agreement targets typically show agreement with this lexical gender. However, even in languages such as Dutch, with a nominal gender system, semantic agreement occurs, especially with human referents. An example of this agreement variation in Dutch is shown in (1).

- (1) **Dat** meisje **dat** daar staat ken ik. **Ze** zit bij mij
DEM.N girl(N) REL.N there stands know I 3SG.F sits with me

op school.
in school

‘I know that girl standing over there. She is in my school.’

In example (1), the determiner and the relative pronoun are neuter, agreeing with the neuter gender of the noun *meisje* ‘girl’, while the personal pronoun is feminine, agreeing with the sex of the referent. With nouns such as *meisje*, the semantic gender of the referent conflicts with the lexical gender of the noun and this semantic gender is apparently able to override lexical gender in the personal pronoun.

This agreement variation is not uncommon cross-linguistically: it is often possible for personal pronouns to agree semantically while other agreement targets agree with the lexical gender of the noun. This is captured in Corbett’s (1979) Agreement Hierarchy, presented in (2) below.

- (2) The Agreement Hierarchy (adapted from Corbett 1979)

attributive > predicate > relative pronoun > personal pronoun

lexical
agreement

semantic
agreement

Cross-linguistically, personal pronouns are most inclined towards semantic agreement, while attributive (or adnominal) elements, such as determiners and attributive adjectives, tend to agree with the lexical gender of the noun. However, the two types of agreement are not necessarily categorically distinguished from each other, with every agreement target showing either lexical or semantic gender agreement. Instead, individual agreement targets can display variation between the two types of agreement, with the likelihood of semantic agreement being higher in elements at the right of the hierarchy.

This intra-target agreement variation is clearly visible in Dutch. Example (1) above has two possible variants, shown in (3a) and (3b) below.

- (3) a. **Dat** meisje **dat** daar staat ken ik. **Het** zit bij mij
 DEM.N girl(N) REL.N there stands know I 3SG.N sits with me

op school

in school

‘I know that girl standing over there. She is in my school.’

- b. **Dat** meisje **die** daar staat ken ik. **Ze** zit bij mij
 DEM.N girl(N) REL.C there stands know I 3SG.F sits with me

op school

in school

‘I know that girl standing over there. She is in my school.’

The personal pronoun can be neuter instead of feminine, as in (3a), although this is infrequent and seems more formal. Also the relative pronoun shows variation between lexical and semantic agreement. While it agrees lexically in (3a), which is the norm, it can agree semantically as well, as in (3b), particularly in informal spoken language.

CHAPTER IV

In view of the Agreement Hierarchy, Corbett (1979) proposes a diachronic relation between the two types of agreement, in which semantic agreement is first introduced in the personal pronoun and then potentially moves leftwards on the hierarchy, gradually encroaching on the domain of lexical gender: “This leads to the diachronic predication to be made from the hierarchy – namely that semantic agreement will first affect the rightmost position and then spread leftwards” (Corbett 1979: 218). This pathway suggests that lexical agreement is generally the older type of agreement, while semantic agreement is an innovation that starts in the pronoun.

However, it is not necessarily the case that semantic agreement is always new compared to lexical agreement. Semantic agreement may have always been part and parcel of gender systems that display it. It seems that, at least in the West Germanic languages, semantic agreement with animate referents has long existed alongside lexical agreement. Specifically, semantic agreement based on the sex of the referent already occurred in Old and Middle High German (Birkenes, Chroni & Fleischer 2014) and in Old English, when English still had a nominal gender system (Curzan 2003). An example from Middle Dutch was encountered in the present study with the neuter noun *wif* ‘woman’, shown in (4) below:

(4) Overmits **een-s** **heyden-s** **wijf-s** **die** hij
because ART.INDEF-GEN.N heathen-GEN.N woman-GEN.N REL.ACC.F he

soe seer mynde ende lieff hadde. Des **sij** wel merkede [...]
so very loved and dear held that 3SG.NOM.F well noticed

‘Because of a heathen woman whom he loved and held dear so much. Which she noticed clearly [...]’

(Dirc Potter, *Blome der doechden*, 1475-1495. Accessed via dbnl.nl.)

In view of the Agreement Hierarchy discussed above, example (4) is particularly remarkable, as it not only shows semantic agreement in the personal pronoun *sij*, but also in the relative pronoun *die*, which directly follows the neuter-marked antecedent.

In present-day Dutch, and in several other Germanic varieties today, another semantic agreement principle has been found to play a role in pronouns: the degree of individuation of the referent (Van Haeringen 1936; 1951, Fletcher 1987, Siemund 2002, Audring 2006; 2009). A distinction is made between referents with a high degree of individuation, that is, things that are countable and bounded in nature, such as concrete objects, and referents with a low degree of individuation, things that are uncountable and unbounded in nature, such as materials and liquids. Referents with a high degree of individuation tend to receive masculine and common gender agreement, while referents with a low degree of individuation tend to receive neuter gender agreement. Examples of this semantic agreement from the *Corpus Gesproken Nederlands* ('Corpus of Spoken Dutch') are shown in (5) and (6) below (from Audring 2006: 95-99).

- (5) Moet je nog wat informatie over dat **boek** hebben?
 need you more some information about DEM.N book(N) have

Dan moet 'k 'm nog niet gaan inleveren.
 then should I 3SG.M yet not go return

'Do you need some more information about that book? Then I shouldn't return it yet.'

- (6) 't zit toch ook bij **olijfolie** wel een beetje in
 it is in.fact also with olive.oil(C) PRT a bit about

hoe 't geconserveerd wordt.
 how 3SG.N preserved is

'In fact also with olive oil, it matters how it is preserved.'

Example (5) shows the use of a masculine pronoun in reference to a neuter noun denoting a concrete object, while example (6) shows the use of a neuter pronoun with a common gender noun denoting a mass.

CHAPTER IV

Pronominal agreement of this kind, viz. masculine agreement with objects and neuter agreement with masses, has been found in, besides Dutch, Helgoland Frisian (Wahrig-Burfeind 1989), West Jutland Danish (Ringgaard 1973, Braunmüller 2000) and in several English dialects, in particular West Somerset English (Siemund 2002; 2008). Semantic agreement with neuter pronouns has also been found in City Frisian (Wahrig-Burfeind 1989) and in Flemish (De Vos & De Vogelaer 2011, De Vogelaer & De Sutter 2011). A similar phenomenon with respect to the use of neuter has been found in Standard Danish, Norwegian and Swedish, which show semantic neuter agreement with referents of low individuation in the predicative adjective (Braunmüller 2000, Enger 2004; 2013). An example from Norwegian with the masculine noun *vodka* ‘vodka’ is shown in (7) below (from Enger 2004: 6):

- (7) *Vodka er sun-t.*
vodka(M) is healthy-N.SG
‘Vodka is healthy.’

Also in Romance varieties, which have a system of masculine and feminine lexical gender, historically neuter forms are used with referents of low individuation. This has been observed in South-Central Italian dialects, mainly in pronouns and determiners, and in Spanish dialects, in pronouns and post-nominal adjectives (Fernández-Ordóñez 2009). An example from a Spanish dialect with the masculine noun *vinu* ‘wine’ is shown in (8) below (from Fernández-Ordóñez 2009: 58):

- (8) *El buen vinu blanc-o se toma frí-o. Pruéba-lo.*
DET.M good.M wine(M) white-N is drunk cold-N taste-3SG.N
‘Good white wine is drunk cold. Taste it.’

The question arises where this semantic agreement based on individuation comes from. It has been suggested that this type of agreement is a new development in these European varieties. Fernández-Ordóñez (2009: 56), for instance, states that “[s]ome Western Indo-European (IE) languages, namely English, Ibero-Romance, Scandinavian and South-Central Italian varieties, have developed new gender distinctions based on the count/mass interpretation of nouns” and “these mass/count

distinctions seem to be new and independent developments of older IE gender systems that are usually well known”.

More specifically, it has been suggested for Dutch that agreement based on individuation has developed as a result of changes that occurred in the nominal gender system (Audring 2006; 2009, De Vos & De Vogelaer 2011). Dutch no longer has the original three Germanic genders, since masculine and feminine nominal gender conflated into common gender in the seventeenth century. Audring (2006; 2009) proposes that the resulting mismatch between the nominal genders in Dutch, common and neuter, and the genders of the personal pronoun, masculine, feminine and neuter, instigated the development of a new, semantic agreement system in pronouns: “When regarded with unbiased eyes, the ‘wrong’ pronouns of spoken Dutch represent a useful and ingenious case of recovery from a historical problem. Speakers of Dutch have reinvented their pronoun genders by putting new semantic foundations under the gender system” (Audring 2009: 217).

This suggestion has been followed up by De Vos & De Vogelaer (2011) who studied pronominal agreement in the dialect of Moerzeke, a Flemish dialect of Dutch in which the distinction between masculine and feminine nominal gender still exists but is disappearing, since the distinction is only sporadically marked in the noun phrase. They found the use of masculine pronouns for feminine nouns and vice versa, with all kinds of referents, which can be considered agreement errors resulting from the loss of the masculine-feminine distinction in the adnominal marking. Additionally, they found the use of neuter pronouns for both masculine and feminine nouns with referents that have a low degree of individuation. In line with Audring’s (2009) suggestion, De Vos & De Vogelaer (2011: 276) interpret this finding as the rise of an innovative agreement system caused by the apparent breakdown of the three-gender system: “This process can be considered an instance of morphological regularization, in other words, the rise of an innovative rule system when the traditional system becomes too opaque to be successfully acquired”.

However, as with semantic agreement based on sex, agreement based on individuation may not be a new phenomenon in Dutch. Audring (2009: 194) also notes that “even in a thriving three-gender system, there can be some degree of semantic agreement, coupled to conceptual distinctions such as natural gender and individuation”. It is possible therefore that semantic agreement based on

CHAPTER IV

individuation has always existed beside lexical gender agreement in Dutch. The fact that it is found in several different Germanic and Romance varieties suggests that it could be an older Germanic or Indo-European feature.

There are in fact indications that individuation was part of the semantic basis of the Proto-Indo-European (PIE) genders. Ever since it was found that the PIE gender system did not originally consist of three genders, but two, with the third, feminine, gender developing later, it has been clear that sex could not have been the semantic basis of the PIE genders (cf. Luraghi 2011). Instead, several theories on the development of PIE gender involve a connection between the genders and individuation. Lehmann (1958), for instance, proposes that PIE gender was a variable feature of nouns and that gender variation was based on different interpretations of the referent as either an individual, a collective, or a resultative/mass entity. In a reconstructed lexicon of PIE nouns, Matasović (2004: 133–134) also observes semantic tendencies in gender assignment which appear to run along the lines of individuation: while nouns referring to humans, animals and most plants are consistently masculine/feminine gender, nouns referring to fluids and substances are always neuter gender. Following Lehmann 1958, Leiss (2000) observes remnants of meaning based gender variation in Old High German, where nominal gender variation appears to distinguish between a count, collective, and mass interpretation of the referent.

Kraaikamp (2012) argues for Dutch that, although the association between the genders and individuation has not been preserved in lexical gender, it is still visible in cases where nominal gender is variable. Since lexical gender is a fixed, invariable feature of nouns, it cannot easily respond to meaning changes over time, making any patterns in lexical gender assignment vulnerable to disruption. However, the semantic pattern still surfaces in cases where nominal gender is variable, for instance with double gender nouns (see also Semplicini 2012) and with nominalized adjectives. Nouns referring to materials, such as *steen* ‘stone’ or *diamant* ‘diamond’, can occur in different genders, with a concomitant shift in meaning. They are common gender when they refer to a unit of the material, ‘a stone’ or ‘a diamond’, and they are neuter when they refer to the material as a mass, ‘stone’ or ‘diamond’ (Haeseryn et al. 1997). Nominalized adjectives show a similar gender alternation associated with a meaning difference. They are assigned common gender when they

refer to an individual, for example *de mooie* ‘the pretty one’, but neuter when the referent is not individuated, as in *het mooie ervan* ‘what is nice about it’. The fact that individuation plays a role in nominal gender as well as pronominal gender suggests that the semantic association of the genders is not an innovation that started in pronouns, but something that has always been part of the Dutch gender system.

This study addresses the question of when semantic agreement based on individuation has developed in Dutch pronouns, particularly if it developed after the change from a three-gender system to a two-gender system or if it existed before this change. Data from Middle Dutch, where the three nominal genders were still in place, can shed light on this issue. If agreement based on individuation resulted from the loss of the three-gender system, it is not expected to occur before this change. This chapter presents the results of a corpus study of pronominal gender agreement in Middle Dutch and is organized as follows. The next section discusses semantic agreement in present-day Dutch in more detail and formulates the specific expectations for the corpus study. Section 3 describes the sources and methodology of the corpus study. The results are presented in Section 4, followed by a discussion of the findings in Section 5. Section 6 concludes this chapter.

2. Gender agreement in present-day Dutch

Standard Dutch spoken in The Netherlands today (subsequently referred to as ‘Dutch’) has two nominal genders, common and neuter, where common gender is a conflation of former masculine and feminine gender. Several Eastern and Southern Dutch dialects, particularly Flemish dialects, still distinguish the original three nominal genders. Also, dictionaries of Standard Dutch sometimes still report masculine or feminine gender for common gender nouns. However, the distinction between masculine and feminine nouns is no longer made in Standard Dutch spoken in The Netherlands.

Common and neuter gender are expressed on determiners, such as the definite article (*de* versus *het*), and on the attributive adjective (*mooie* versus *mooi* ‘pretty’),

CHAPTER IV

for instance).¹ In the pronominal domain, the two genders are expressed on the relative pronoun (*die* versus *dat*) and on the distal and proximal demonstrative pronouns (*die, deze* versus *dat, dit*). The personal pronoun expresses masculine (*hij, hem*), feminine (*zij, haar*) and neuter (*het*) gender. The feminine pronoun is used exclusively with female referents in informal spoken language (Audring 2009: 92). The masculine personal pronoun is used to agree lexically with common gender nouns.²

Lexical gender assignment is largely arbitrary in Dutch. There exists a limited set of semantic and formal regularities, such as that nouns referring to adult humans or nouns ending in *-de* or *-te* are usually common gender. However, there are no general rules behind the assignment of common or neuter gender to nouns (see, for instance, Haeseryn et al. 1997).

As discussed in Section 1, determiners and adjectives agree with the lexical gender of the noun, while pronouns show variation between lexical agreement and semantic agreement on the basis of the sex and degree of individuation of the referent. Audring (2009) captures the relevant semantic distinctions in the Individuation Hierarchy, shown in (9) below.

¹ Other determiners expressing gender are the demonstrative determiners *die/dat* ‘that’, *deze/dit* ‘this’, *zulke/zulk* ‘such’, the interrogative determiner *welke/welk* ‘which’, the collective determiners *iedere/ieder, elke/elk* ‘every’, the indefinite quantifying determiner *menige/menig* ‘many a’ and the first person plural possessive determiner *onze/ons* ‘our’.

² It is generally believed that with the disappearance of the distinction between masculine and feminine nominal gender, feminine nouns were masculinized, as former feminine nouns were increasingly referred to by masculine pronouns (Geerts 1966). However, Audring (2009: 89-91) debates whether masculine pronouns in fact agree lexically with common gender nouns in present-day Dutch and argues that when masculine pronouns are used with common gender nouns, this is always semantic agreement, not lexical agreement. A problem with this view, however, is that masculine pronouns appear to be used more easily with common gender nouns than with neuter gender nouns. This remains unexplained if masculine pronouns are not somehow considered to agree with common gender nouns.

(9) The Individuation Hierarchy and Dutch pronouns (adapted from Audring 2009: 127)

human > animal > object / bounded abstract > specific mass > unbounded abstract / unspec. mass
 (girl) (horse) (book) / (question) (my tea) (love) / (snow)

| | | |
|-------------------|------------------|---------------|
| <i>fem./masc.</i> | <i>masculine</i> | <i>neuter</i> |
| <i>common</i> | <i>common</i> | |

The Individuation Hierarchy is essentially an elaboration of the Animacy Hierarchy (Silverstein 1976), in which the inanimate part of the hierarchy is further subdivided.³ The degree of individuation of entities decreases from left to right on the hierarchy. Entities with a high degree of individuation have a clearly bounded shape, are countable and have specific characteristics, whereas entities with a low degree of individuation have an unbounded shape or unclear boundaries, are uncountable and have less specific characteristics. Animate entities have the highest degree of individuation, and within that category, humans are more individuated than animals.⁴ Next to animals there is the category of bounded objects, such as a book or a cup, and bounded abstracts, such as a question or a word. More lowly individuated are specific masses, which are specific instantiations of masses, such as my tea or this wine. Unbounded abstracts, such as love or nature, and unspecific masses, such as snow or honey, have the lowest degree of individuation.

In her study of Dutch spoken language data from the *Corpus Gesproken Nederlands* ('Corpus of Spoken Dutch'), Audring (2009) found that pronouns display a semantic agreement pattern along the lines of individuation when they do not agree with the lexical gender of the noun. Masculine and common pronouns are

³ Several slightly different versions of the Individuation Hierarchy exist in the literature. The hierarchy presented here is from Audring (2009: 127) and is based on the distinctions she found to be relevant in present-day Dutch. This is an adapted version of Sasse's (1993: 659) Continuum of Individuality. In Sasse's hierarchy, abstracts form a single category between concrete objects and masses, and specific and unspecific masses are not distinguished.

⁴ This follows from the fact that the Individuation Hierarchy is a categorization of entities from the human viewpoint on the world. From the human perspective, humans are more individuated than other animals, and they in turn are more individuated than lifeless objects (see Audring 2009: 125).

CHAPTER IV

used for referents with a high degree of individuation, including humans, animals, objects and bounded abstracts. Neuter pronouns are used for referents with a low degree of individuation, including specific masses, unbounded abstracts and unspecific masses. This type of agreement was found to be frequent. Personal and demonstrative pronouns show semantic agreement in 65% of the cases where lexical and semantic gender conflict, that is with neuter nouns referring to highly individuated entities, such as *fototoestel* ‘camera’, and common gender nouns referring to lowly individuated entities, such as *olijfolie* ‘olive oil’. The relative pronoun shows semantic agreement as well, but less frequently, in 25% of the cases (Audring 2009: 159-160). This difference is in line with Corbett’s Agreement Hierarchy presented in Section 1, which describes that relative pronouns are less inclined towards semantic agreement than personal pronouns.

Another factor that determines the frequency of semantic agreement in Dutch is the position of the referent on the Individuation Hierarchy. Semantic agreement occurs significantly more often with referents at the extreme ends of the hierarchy than with those towards the middle (Audring 2009: 167-168). This is explained by the fact that referents at the ends of the hierarchy have a more saliently high or low degree of individuation. When lexical gender conflicts with semantic gender, as with neuter for individuated referents or common gender for referents of low individuation, this conflict is most prominent with referents at the outer ends of the hierarchy. It is not unexpected therefore to find most semantic agreement there.

The present-day Dutch gender system thus shows two different principles for gender agreement: lexical gender, which is largely arbitrary, and semantic gender, which is based on the degree of individuation and sex of the referent. This leads to competition and agreement variation in those cases where the two genders do not coincide. This situation may have always existed in Dutch if the genders have always had a semantic interpretation. Semantic agreement with animate referents is known to have existed already alongside lexical agreement in Middle Dutch (cf. Section 1). The expectation is that semantic agreement with inanimate referents exists in Middle Dutch as well, with masculine pronouns being used for referents that have a high degree of individuation and neuter pronouns being used for referents that have a low degree of individuation.

3. Sources and methodology

3.1. Textual sources

The texts used for this study had to meet three important requirements: they had to date from before the conflation of masculine and feminine nominal gender in Dutch, they had to be prose texts, so that considerations of rhyme or special styles are not at play, and they had to contain a relatively high number of pronominal references to inanimate referents. It was decided to look at sixteenth century recipe books. The following two texts were used:

Een notabel boecxken van cokeryen. Printed in Brussels in ca. 1514, by Thomas van der Noot, anonymous author (Jansen-Sieben & Van der Molen-Willebrands, eds., 1994, accessed via dbnl.nl).

Dat batement van recepten. Een secreetboek uit de zestiende eeuw. Printed in Antwerp in 1549, by Hans de Laet, anonymous author (Willy L. Braekman, ed., 1990, accessed via dbnl.nl).

The *Boecxken* is a collection of food recipes. It contains 175 recipes, comprising approximately 17,000 words in total. The *Batement* is a collection of food recipes and recipes for medicinal and household purposes. It contains 258 recipes, comprising approximately 22,000 words in total. Both texts were printed in the first half of the 16th century.

The authors of the recipes are anonymous, but the names of the printers are known, as are the locations where the books were printed, Bussels and Antwerp. The choice for texts from cities in the southern part of the Dutch speaking area was practical, as no suitable sources from the north were available for this study. Most of the Middle Dutch literature is from the south, as the culturally most important cities of the Dutch speaking area were located there until the fall of Antwerp in 1585. However, the three-gender system that is observed in these southern texts can be considered representative of the system that must have preceded the gender system of both present-day Southern and Northern Dutch.

CHAPTER IV

3.2. Data collection

All pronominal references to singular nominal antecedents were manually extracted from the texts. All types of pronouns were collected, including personal pronouns, both freestanding and cliticized forms, demonstrative pronouns, relative pronouns and possessive pronouns. Sequences of pronouns with the same antecedent were included. The clitic personal pronoun was the most frequently attested form. The gender-case paradigms of the freestanding and clitic personal pronoun are shown in Table 1 and 2 below.⁵

Table 1. Paradigm of the personal pronoun singular in Middle Dutch

| | MASC | FEM | NEUT |
|-----|--------------|--------------|-------|
| NOM | hi/hie | si/sie | het |
| GEN | sijns/sijnre | haer(s)/hare | sijns |
| DAT | hem/heme | haer/hare | hem |
| ACC | hem | haer/hare | het |

Table 2. Paradigm of the clitic personal pronoun singular in Middle Dutch

| | MASC | FEM | NEUT |
|-----|--------------|--------------|--------|
| NOM | -i | -si/-se | -et/-t |
| GEN | -es/-s | -ere/-er/-re | -es/-s |
| DAT | -em/-en | -ere/-er/-re | -em |
| ACC | -en/-ene/-ne | -se | -et/-t |

⁵ The paradigms presented in Tables 1 and 2 are largely based on Van Royen 1991. Grammars of Middle Dutch, e.g. Van Loey 1976, Van de Ketterij 1980, Van Royen 1991 and Mooijaart & Van der Wal 2008, differ with respect to the number of form variants, and thereby potential gender syncretisms, they report. Van Royen 1991 is the most comprehensive grammar in this respect. To ensure that potential gender syncretisms were taken into account maximally, the paradigms from this grammar formed the basis for this study, and syncretisms and forms not reported in Van Royen 1991 but mentioned in the other grammars were taken into account as well. For instance, Van Royen (1991) does not report a clitic form for the dative neuter, but Mooijaart & Van der Wal (2008) report the form *-em* shown in Table 2, which is syncretic with the masculine dative form *-em*. This form was therefore considered ambiguous masculine/neuter in this study.

Note that there are gender syncretisms in some case forms, such as the dative form *hem*, which is both masculine and neuter. Pronouns with ambiguous forms were recorded as such in the results.

References to plural antecedents were not included in this study, because there are no gender distinctions in the plural. Also not included were references to disjunctions, conjunctions, lists and other combinations of nominal antecedents, because in these cases, the antecedent does not have a (single) gender. Combinations of nominal antecedents were frequently found, because ingredients are combined and subsequent pronominal references then tend to refer to the resulting mixture. (10) is an example of this kind of reference from the *Boecxcken*:

(10) *Daerna neemt witt-en wijn met wat verjuys, dat*
 then take white-ACC.M wine with some verjus DEM.SG.ACC.N

minghelt wel tesamen met bloemen van rijs.
 mix well together with flour of rice

‘Then take white wine with some verjus, mix that well together with rice flour.’

It is interesting to note that with such references to mixtures, neuter pronouns are almost always used. This is in line with the expectation for mass referents, which these mixtures always are. However, as more than one nominal gender can be involved in such cases, these references were not included. Note that such methodological decisions were made with the intention to maximize the possibility of falsifying the hypothesis.

Other references that were excluded are those for which there is more than one plausible antecedent. This can be the case, for instance, with quantifier phrases such as *een bolleken eyiuns* ‘a ball of onion’, whereby it is unclear if the reference is to the quantifier, *bolleken*, or the quantified, *eyiun*. References to nouns that are not in the dictionary and whose gender and/or exact meaning remains uncertain were excluded as well. Finally, references with invariant neuter pronouns were excluded. These occur in nominal copula phrases, such as *Daerna neemt tartre, dat is*

CHAPTER IV

droessem van wijn die gedroocht is ‘Then take tarter, that is wine sedate that has been dried’. In such phrases, the pronoun does not show gender agreement but is invariantly neuter.

3.3. Categorization

For each pronominal reference, the gender of the pronoun and the type of pronoun was recorded. The gender of the antecedent was determined and the referent was categorized according to its semantic class. The methodology used to determine the gender of the antecedent and the semantic class of the referent is described below.

3.3.1. The gender of the antecedent

The gender of the antecedent was based on adnominal gender marking, viz. gender-distinctive forms of determiners, adjectives or nominal case endings, and the gender listed for the noun in the *Middelnederlandsch Woordenboek (MNW)* ‘Middle Dutch dictionary’. This dictionary, by Jacob Verdam and Eelco Verwijs, covers the Dutch vocabulary from circa 1250 to circa 1550. Its nine parts were published consecutively from 1885 to 1929.⁶

Adnominal gender marking was considered the most reliable indication of the antecedent’s gender, as this is a direct reflection of the gender attributed to the noun by the author. It was therefore considered the primary source for the antecedent’s gender. However, the additional use of dictionary information was inevitable, because gender marking is not always present, and if it is, it is sometimes ambiguous. If there was no marking, the gender of the antecedent was based on the dictionary gender of the noun. If the gender marking was ambiguous, the gender of the noun was based on this marking in combination with the dictionary gender. How these two sources of information were mapped onto each other is explained below. In some cases, when the antecedent itself has no, or ambiguous, gender marking, the same noun occurs within the same or a neighboring, related recipe, with unambiguous marking. This marking was then taken into account as adnominal gender marking.

⁶ The *Middelnederlandsch Woordenboek (MNW)* was accessed via the *INL-Geïntegreerde TaalBank (INL-GTB)*, a digital, online database, supported by the *Instituut voor Nederlandse Lexicologie* and the *Nederlandse Taalunie*, <http://gtb.inl.nl>.

The dictionary was used in combination with the gender marking on the antecedent in cases where this gender marking was ambiguous. Ambiguous gender marking exists because of syncretisms in the gender-case paradigms of adnominal elements. To illustrate this, the Middle Dutch paradigm of the singular definite article is given in Table 3 below.⁷

Table 3. Paradigm of the definite article singular in Middle Dutch

| | MASC | FEM | NEUT |
|-----|---------------------------|----------------|--------------|
| NOM | die/de/(den) ⁸ | die/de | dat |
| GEN | des | der/dier/diere | des |
| DAT | dien/den/die | der/dier/diere | dien/den/die |
| ACC | dien/den/die | die/de | dat |

The table shows that the form of the definite article is the same for masculine and feminine gender in the nominative case, and for masculine and neuter gender in the genitive and dative case. If such ambiguous gender marking was encountered, the gender of the antecedent was determined on the basis of this marking in combination with the dictionary gender of the noun. For example, if the adnominal marking is masculine/feminine gender, and the gender listed for the noun in the dictionary is masculine, the antecedent was considered to have masculine gender. When the dictionary lists more than one gender for the noun, for example masculine and neuter gender, while the adnominal gender marking is masculine/feminine, the noun was considered to have masculine gender. If the dictionary genders coincide with

⁷ The paradigm presented in Table 3 is largely based on Van Royen 1991. The neuter dative form *den* is not reported in Van Royen 1991:94, but is reported in Van de Ketterij 1980:47 and Mooijaart & Van der Wal 2008:43. See footnote 5 for the approach adopted in this study with respect to form variants and potential gender syncretisms.

⁸ The use of the originally accusative form *den* in the nominative was rarely found in the texts used in this study. It occurs twice, with the nouns *inct* 'ink' and *roost* 'roast', which are both listed with masculine and neuter gender in the dictionary. With *inct*, the noun is the subject of a subclause that functions as the object in the main clause: [...] *dan proeft oft den inct swart ghenoech zijn sal* '[...] then check if the inc will be black enough'. With *roost*, the noun is the subject of a passive sentence: *Ende als den roost ghetrocken sal sijn uut den spete* [...] 'And when the roast will have been pulled off the spit [...]'].

CHAPTER IV

the ambiguous marking on the antecedent, in this example, if the dictionary genders are masculine and feminine gender, the ambiguity remains and the antecedent was labelled as having ambiguous masculine/feminine gender. These ambiguous-gender antecedents were not excluded from the data, but, as indicated in the results section, they were excluded from the quantitative analyses.

With respect to the dictionary gender, it can of course never be said with certainty that the gender of a noun given in the dictionary is in fact the gender attributed to that noun in the text. A problem with dictionary gender is that nominal gender is subject to regional and diachronic variation, which may not be captured by the dictionary. The *MNW* at least aims to report gender variation and uncertainty regarding nominal gender. It regularly lists more than one gender for a particular noun, pointing either to uncertainty about the noun's gender or to gender variation.⁹ The gender information in the *MNW* generally proved to be reliable for the texts used in this study. In those cases where a noun was marked for gender in the text, the dictionary gender was almost always in accordance with this marking. Only three cases were encountered where the dictionary gender and the gender marking in the text did not match. Therefore, the *MNW* was considered a fairly reliable source of information regarding nominal gender.

3.3.2. *The semantic class of the referent*

For each pronominal reference, the referent was categorized according to its degree of individuation. The large majority of referents were inanimate referents, for which the following semantic classes were distinguished:

- object
- bounded abstract
- unbounded abstract

⁹ It is not always clear what the multiple genders in the dictionary are based on. Sometimes the citations provided in the entry demonstrate the gender variation, as the noun occurs with varying adnominal gender marking, but such examples are not always provided. It is not impossible that the reported gender variation is in some cases based on pronominal agreement. It is not necessarily the case then that the noun has different lexical genders, but the variation may exist in the pronominal agreement. For this study, this means that there could be an underreporting of semantic agreement.

- mass

In the context of the texts that were studied, some animal referents such as *een visch* ‘a fish’, could in principle be interpreted as either animate or inanimate. When such referents occur in the context of cooking and are referred to as lifeless ingredients, they were classified as objects.

The distinction between objects and masses was most relevant for this study. However, referents were sometimes ambiguous between an object and a mass interpretation. An example of such an object/mass referent from the *Boecxcken* is shown in (11):

- (11) *Neemt den palinck ende siett-en wel.*
 take DET.SG.NOM.M eel and cook-3SG.ACC.M well
- Dan doet-en uut sin-en sope ende laitt-en*
 then do-3SG.ACC.M out 3SG.POSS-ACC.M juice and let-3SG.ACC.M
- coelen.*
 cool

‘Take the eel and cook it well. Then take it out of its juice and let it cool off.’

In this example, *den palinck* could in principle refer to a singular eel or a mass of eels. In some cases, the larger context disambiguates these readings. For instance, if the fish has been chopped up earlier in the recipe, the referent is clearly a mass. However, in cases where the context does not clearly point to an object or a mass reading, the referent was classified as ambiguous object/mass.

For masses, a distinction between specific masses and unspecific masses was originally made, following Audring (2009) (cf. Section 2). However, this distinction proved to be irrelevant in the present study, as no significant differences between the two categories were found (see footnote 11 in Section 4.2). The distinction is therefore not discussed further and the two types of masses are subsumed under one category.

4. Results

4.1. Data

A total number of 731 pronominal references were collected, of which 672 are references to inanimate referents.¹⁰ The majority of these referents are concrete objects and masses. Very few abstract referents were found. The number of references per semantic class is presented in Table 4 below. A complete list of the references that were found, including the exact lexical items involved, is provided in the Appendix.

¹⁰ A total of 59 references to animate referents were found. These all occur in the medicinal and household recipes in *Dat Batement van recepten*. Most of the references (44/59) are references with feminine pronouns to female persons, involving the feminine antecedent *vrouwe* ‘woman’ or *paciente* ‘(female) patient’, for example: [...] *ende dan legtse op den nauel van die vrouwe, ende terstont sal si dat kint baren sonder groote weedom* ‘[...] and then put them on the navel of the woman, and she will immediately give birth to the child without much pain’ or [...] *ende dan g(e)uet die paciente te drincken tsauons ende tsmorgens, ende si sal haer sake krijghen* ‘[...] and then give it to the patient to drink in the evening and in the morning, and she will get her period’. Only three deviations from lexical gender occur, all with the same feminine antecedent *creature* ‘creature’, referring to an infant: *Als die eerste creature van die vrouwe voort gebracht, ter werelt gecomen zijn sal, alsoo gheringhe als 't die vroevrouwe ontfangen hebben sal, besiet hoeveel knoopkens ghi sien sult aen 't peesken dat hem aen den nauel hanght, by den welcken het gheuoecht ende gheuest was*. ‘When the first creature brought forth by the woman is born, as soon as the midwife has received it, check how many buttons you see on the ligament that hangs from its navel, through which it was fed and secured’. The pronominal references are consecutively with neuter gender (*'t*), ambiguous masculine/neuter gender (*hem*) and neuter gender (*het*). This neuter agreement may be motivated by a need for it to be clear that the sex of the new-born referent is not relevant for the instructions that are given here.

Table 4. Number of data per semantic class

| semantic class of the referent | number of references |
|---------------------------------------|-----------------------------|
| object | 187 (28%) |
| bounded abstract | 5 (0.7%) |
| mass | 437 (65%) |
| unbounded abstract | 3 (0.4%) |
| ambiguous object/mass | 40 (6%) |
| <i>total</i> | <i>672 (100%)</i> |

In the analyses that follow, the references to ambiguous object/mass referents are excluded, as they cannot be used to test the hypothesis.

4.2. References to masses and unbounded abstracts

Tables 5 and 6 show the pronominal agreements with referents that have a low degree of individuation, masses and unbounded abstracts respectively. The genders of the antecedents are shown on the left and the genders of the pronouns that refer to them are shown at the top. Ambiguous genders are indicated with a slash. For the antecedents, ambiguous gender means that the noun could be either of the two or three indicated genders, according to the dictionary and/or the adnominal marking. For the pronouns, it means that the pronoun has an ambiguous form with respect to gender, due to gender syncretisms in particular case forms, such as the masculine/feminine nominative form *die* of the demonstrative and relative pronoun.

CHAPTER IV

Table 5. Gender of antecedents and pronouns referring to masses

| MASS | pronoun | | | | | |
|-------------------|----------------|--------------|----------|---------------|-----------|--------------|
| antecedent | masc | masc/ fem | fem | masc/ neut | neut | <i>total</i> |
| masc | 27 | 18 | - | 2 | 22 | 69 |
| masc/fem | - | 3 | 12 | - | 5 | 20 |
| fem | - | 7 | 29 | - | 9 | 45 |
| masc/neut | 13 | 4 | 1 | 1 | 26 | 45 |
| fem/neut | - | 3 | 4 | - | 23 | 30 |
| neut | - | 1 | 4 | 3 | 216 | 224 |
| masc/fem/neut | - | - | 2 | - | 2 | 4 |
| <i>total</i> | 40 | 36 | 53 | 6 | 303 | 437 |

Table 6. Gender of antecedents and pronouns referring to unbounded abstracts

| UNBND ABSTRACT | pronoun | | | |
|---------------------------|----------------|-----|----------|--------------|
| antecedent | masc | fem | neut | <i>total</i> |
| masc | | | 1 | 1 |
| masc/fem | 1 | | | 1 |
| fem | | 1 | | 1 |
| <i>total</i> | 1 | 1 | 1 | 3 |

Note that, with masses and unbounded abstracts, deviations from lexical gender were expected with non-neuter antecedents, that is with masculine and feminine nouns (top three rows in Table 5, all rows in Table 6). The tables show that the pronouns mostly agree with the gender of their antecedents, but not always. The cases where the gender of the pronoun clearly deviates from that of the antecedent (cases of a gender ‘switch’) are marked in bold. A total of forty-two switches occur

with masses. These switches show a clear pattern. They are predominantly switches to neuter with masculine and feminine antecedents. With neuter antecedents, only very few switches occur. References to unbounded abstracts were very rare, but also there a switch to neuter is found with a masculine antecedent.

With the antecedents that have ambiguous lexical gender, it is in many cases uncertain whether the pronoun agrees with the lexical gender of the antecedent or not. However, it is interesting to note that with masculine/neuter antecedents and feminine/neuter antecedents, the majority of the references are with neuter pronouns. These could be either cases of lexical agreement with neuter nouns, or they could be switches to neuter gender with masculine and feminine nouns. It is not unlikely that some of these cases are switches to neuter, considering that such switches were found with unambiguously masculine and feminine nouns.

The same uncertainty exists with masculine/feminine antecedents (20 in total). Here, the use of masculine and feminine pronouns could be cases of lexical agreement with masculine and feminine nouns respectively or they could be switches to masculine and feminine gender. The latter seems unlikely however, considering that no switches to masculine gender were found with unambiguously feminine nouns and no switches to feminine gender were found with unambiguously masculine nouns. Because of the uncertainties with antecedents that have ambiguous gender, references to these antecedents are left out in the quantitative analyses that follow.

Note that the issue is different for pronouns with ambiguous gender. Pronouns can be in a case form that is ambiguous between two genders, for instance the masculine/feminine nominative form *die*. If this form does not conflict with the gender of the antecedent, for instance, if the antecedent is masculine in this example, the pronoun agrees with the gender of the antecedent, or the other way around, the pronoun cannot be considered to deviate from it. Such cases are counted as lexical agreements in the quantitative analyses below.

In Table 7, the total number of lexical and non-lexical agreements are shown for the different antecedents. The results for masses and unbounded abstracts are conflated here. The results are split up for antecedents that have a gender that conflicts with the low degree of individuation of the referent, that is masculine and feminine nouns, and antecedents that have a gender that does not conflict with this,

CHAPTER IV

that is neuter nouns. Antecedents with ambiguous gender have been left out here. Non-lexical agreements are those cases where the pronoun evidently does not agree with the gender of the antecedent. Lexical agreements are those cases where the pronoun agrees with the gender of the antecedent.

Table 7. Lexical and non-lexical agreement with masses and unbounded abstracts

| MASS/UNBOUNDED ABSTRACT | CONFLICT masculine and feminine antecedents | NO CONFLICT neuter antecedents |
|------------------------------------|--|---|
| non-lexical | 32 (28%) | 5 (2%) |
| lexical | 84 (72%) | 219 (98%) |
| <i>total</i> | <i>116 (100%)</i> | <i>224 (100%)</i> |

It is clear from the table that the agreement behavior is not the same for the different antecedents. With masculine and feminine antecedents, pronouns agree in lexical gender in only 72% of the cases, while they do so in 98% of the cases with neuter antecedents. The association between the gender of the antecedent, that is conflicting or not conflicting, and the type of agreement, that is lexical or non-lexical, is statistically significant ($\chi^2(1) = 51, p < 0.001$).¹¹

The 28% of non-lexical agreements with masculine and feminine nouns are all switches to neuter gender. They can therefore be called semantic agreements. They involve thirteen different nouns and twenty-seven unique antecedents.¹² Some examples are shown in (12) – (14) below.

¹¹ As mentioned in Section 3 above, the category ‘mass’ includes both specific and unspecific masses, as there was no statistically significant difference between the two categories of masses. For specific and unspecific masses separately, the ratios of non-lexical to lexical agreement in case of conflict are as follows: with specific masses, 32% (N=21) non-lexical agreement versus 68% (N=45) lexical agreement. For unspecific masses, 23% (N=11) non-lexical agreement versus 77% (N=36) lexical agreement. A Chi square test reveals that the difference between the two categories is not statistically significant: $\chi^2(1) = 0.957, p = 0.328$.

¹² Sequences of pronouns with the same antecedent were included. Out of the thirty-three cases of semantic agreement with neuter, twenty-seven are pronominal references to a unique antecedent and six are repetitions of neuter pronouns with the same antecedent.

In (12), a neuter clitic personal pronoun is used with the noun *zeem* ‘honey’, which is marked for masculine gender on the determiner and adjectives.

(12) *Ende dien clar-en ghesuiverd-en zeem doet*
 and DET.SG.ACC.M clear-SG.ACC.M purified- SG.ACC.M honey put

in eenen pot. Aldus sal men-t orboren in den ipocras
 in a pot thus shall one-3SG.ACC.N use in the hippocras

ende in den clareyt voerseyt.
 and in the claret aforementioned

‘And put the clear, purified honey in a pot. As such, it will be used in the hippocras and in the aforementioned claret.’

(*Een notabel boecxken van cokeryen*, recipe no. 172)

Example (13) below shows a sequence of pronouns in which a gender switch occurs. The antecedent is *medecijne* ‘medicine’, which is feminine according to the dictionary. The pronominal sequence shows a gender switch. First, a masculine/feminine relative pronoun is used, *de welke* ‘that which’, which agrees with the lexical gender of the antecedent, but later on a neuter clitic pronoun is used.

CHAPTER IV

- (13) *Een* *wonderlijke medecijne de* *welke*
DET.SG.NOM.M/F/N miraculous medicine DET.NOM.M/F which

dat fleercijn ende leemten geneest ter seluer tijt als
the rheumatism and paralyses cures at.the same time when

't *genut is.*
3SG.NOM.N taken is

‘A miraculous medicine which cures rheumatism and paralyses the moment it is taken.’

(*Dat batement van recepten*, recipe no. 155)

Example (14) shows non-lexical neuter agreement with a demonstrative pronoun. The pronoun is used in reference to *ghymbere* ‘ginger’, which is marked for masculine gender on the adjective.

- (14) *Dan neemt witt-en* *ghymbere ende tempert dat*
then take white- SG.ACC.M ginger and mix DEM.SG.ACC.N

met verjuys.
with verjus

‘Then take white ginger and mix that with verjus.’

(*Een notabel boecxken van cokeryen*, recipe no. 4)

Example (15) shows switches to neuter in reference to a mass and in reference to an unbounded abstract. The first one occurs with the antecedent *wijn* ‘wine’, which is accompanied by a masculine/neuter determiner and which is masculine according to the dictionary. The second switch to neuter occurs with *die weedom* ‘the pain’,

which is accompanied by a masculine/feminine determiner and is feminine according to the dictionary.

(15) *Dan neemt eenen swelch van **den** voorscreuen **wijn***
 then take a gulp of DET.SG.DAT.M/N aforementioned wine

*matelijcken werm, dat ghij 't lijden moecht, ende houw-**et***
 moderately warm that you it bear can and hold-3SG.ACC.N

*aen die side daer ghi **die** weedom geuoelt, ende 't*
 on the side where you DET.SG.ACC.F pain feel and 3SG.NOM.N

sal terstont vergaen.
 will instantly disappear

‘Then take a gulp of the aforementioned wine, moderately warm, as you can bear it, and hold it to the side where you feel the pain, and it will disappear instantly.’

(*Dat batement van recepten*, recipe no. 189)

As the examples show, semantic agreement was found with different types of pronouns. It was found with personal pronouns, both clitics and full forms, demonstrative pronouns and relative pronouns. The tendency towards semantic agreement may differ per pronoun type. In particular, the Agreement Hierarchy described in Section 1 predicts that relative pronouns are least inclined towards semantic agreement. Table 8 below shows the amount of lexical and semantic agreement per pronoun type.

CHAPTER IV

Table 8. Lexical and semantic agreement per pronoun type for masses/unbounded abstracts with masculine and feminine antecedents

| | clitic personal | full personal | demonstrative | relative |
|--------------|------------------|-----------------|------------------|------------------|
| semantic | 22 (35%) | 2 (22%) | 4 (18%) | 4 (17%) |
| lexical | 40 (65%) | 7 (78%) | 18 (82%) | 19 (83%) |
| <i>total</i> | <i>62 (100%)</i> | <i>9 (100%)</i> | <i>22 (100%)</i> | <i>23 (100%)</i> |

It appears that clitic personal pronouns have the highest tendency towards semantic agreement (35%). Full personal pronouns, demonstrative pronouns and relative pronouns all show a lower tendency towards semantic agreement (22%, 18% and 17% respectively). However, multiple chi square tests between all pronoun types reveal no statistically significant associations between the type of pronoun and the type of agreement. Even the difference between clitic personal pronouns and relative pronouns, which respectively show the highest and the lowest proportion of semantic agreement, is not statistically significant (Fisher's Exact, $p=0.122$).

4.3. References to objects and bounded abstracts

Tables 9 and 10 show the pronominal agreements with referents that have a high degree of individuation, objects and bounded abstracts respectively. With these referents, deviations from lexical gender were expected with non-masculine antecedents, so with feminine and neuter nouns (third and fifth row from the top in Table 9, second and third row in Table 10). Cases where the gender of the pronoun clearly deviates from the gender of the antecedent are marked in bold.

Table 9. Gender of antecedents and pronouns referring to objects

| OBJECT | pronoun | | | | | |
|-------------------|----------------|--------------|-----------|---------------|-----------|--------------|
| antecedent | masc | masc/ fem | fem | masc/ neut | neut | <i>total</i> |
| masc | 40 | 4 | 1 | 5 | 1 | <i>51</i> |
| masc/fem | 1 | 1 | 4 | 0 | 0 | <i>6</i> |
| fem | 0 | 13 | 55 | 0 | 2 | <i>70</i> |
| masc/neut | 1 | 1 | 0 | 0 | 2 | <i>4</i> |
| neut | 0 | 1 | 0 | 4 | 49 | <i>54</i> |
| masc/fem/neut | 0 | 1 | 1 | 0 | 0 | <i>2</i> |
| <i>total</i> | <i>42</i> | <i>21</i> | <i>61</i> | <i>9</i> | <i>54</i> | <i>187</i> |

Table 10. Gender of antecedents and pronouns referring to bounded abstracts

| BND ABSTRACT | pronoun | | | |
|-------------------------|----------------|---------------|----------|--------------|
| antecedent | masc/ fem | masc/ neut | neut | <i>total</i> |
| masc | 1 | 2 | 0 | <i>3</i> |
| fem | 1 | 0 | 0 | <i>1</i> |
| neut | 0 | 0 | 1 | <i>1</i> |
| <i>total</i> | <i>2</i> | <i>2</i> | <i>1</i> | <i>5</i> |

The results for objects and bounded abstracts are very different from those for masses and unbounded abstracts. Hardly any deviations occur, and the few that occur go in different directions. Switches to masculine gender were expected, but only one switch to masculine/feminine gender is found.

The total number of lexical and non-lexical agreements with objects and bounded abstracts are shown in Table 11 below. The results are split up for antecedents with a gender that conflicts with the high degree of individuation of the

CHAPTER IV

referent, that is feminine and neuter nouns, and antecedents whose gender does not conflict with the referent, that is masculine nouns. Antecedents with ambiguous gender have been left out.

Table 11. Lexical and non-lexical agreement with objects and bounded abstracts

| OBJECT/BOUNDED ABSTRACT | CONFLICT feminine and neuter antecedents | NO CONFLICT masculine antecedents |
|------------------------------------|---|--|
| non-lexical | 3 (2%) | 2 (4%) |
| lexical | 123 (98%) | 52 (96%) |
| <i>total</i> | <i>126 (100%)</i> | <i>54 (100%)</i> |

The table shows that, with objects and bounded abstracts, the pronouns almost always agree lexically, regardless of the gender of the antecedent (conflicting or not conflicting). The association between the condition, conflicting or no conflicting gender of the antecedent, and the type of agreement is not statistically significant (Fisher's Exact, $p=0.637$).

Taking the results for all referents together, shown in Table 12 below, the percentage of non-lexical agreements in cases of a gender conflict, that is, with masculine and feminine nouns referring to masses/unbounded abstracts and neuter nouns referring to objects/bounded abstracts, comes to a total of 14%.¹³ In the cases of no conflict, that is, with neuter nouns referring to masses/unbounded abstracts, and masculine nouns referring to objects/bounded abstracts, the percentage of non-lexical agreement comes down to a total of 3%.

¹³ Most non-lexical agreements in cases of conflict are semantic agreements. As shown in Tables 5, 6, 9 and 10, the gender switches that do not follow the semantic pattern are very few. The exact ratio is as follows: 13% (N=33) semantic agreement, 1% (N=2) non-lexical/non-semantic agreement and 86% (N=207) lexical agreement.

Table 12. Total lexical and non-lexical agreement (all referents)

| ALL REFERENTS | GENDER CONFLICT | NO GENDER CONFLICT |
|---------------|-------------------|--------------------|
| non-lexical | 35 (14%) | 7 (3%) |
| lexical | 207 (86%) | 271 (97%) |
| <i>total</i> | <i>242 (100%)</i> | <i>278 (100%)</i> |

5. Discussion

The results of this study show that Middle Dutch pronouns do not always agree with the lexical gender of the noun, but also show semantic gender agreement, based on the degree of individuation of the referent. As in present-day Dutch, neuter pronouns tend to be used for referents with a low degree of individuation, regardless of the gender of the antecedent noun. This finding indicates that this type of agreement did not develop in response to the loss of the Germanic three-gender system, as this was still in place in Middle Dutch, but already existed before this change. This finding is in line with the idea that the association between neuter gender and a low degree of individuation is an old Germanic feature, which possibly dates back to Proto-Indo-European (Lehmann 1958, Leiss 2000, Matasović 2004, Luraghi 2011).

What has not been found in the present study is semantic agreement of masculine pronouns with inanimate referents that have a high degree of individuation, that is, objects and bounded abstracts, which is found in present-day Dutch. This could mean that this type of agreement either does not exist at all in Middle Dutch or that it is very infrequent and could not be attested in this study. Based on the current data, it appears that semantic agreement takes place only with referents at the ends of the Individuation Hierarchy in Middle Dutch, with masses and unbounded abstracts at the right end and with human referents at the left end (recall the example of semantic agreement with *wif* ‘woman’ in Section 1), as illustrated in (16).

CHAPTER IV

(16) The Individuation Hierarchy and semantic agreement in Middle Dutch

human – animal – object/bounded abstract – **unbounded abstract/mass**

This distribution of semantic agreement is more limited than what is seen in present-day Dutch, where semantic agreement occurs with referents from all categories on the hierarchy. However, the distribution mirrors the frequency distribution of semantic agreement observed in Dutch today. As mentioned in Section 2, the frequency of semantic agreement in present-day Dutch is not the same with all referents: it occurs mostly with referents at the ends of the hierarchy, which have a most saliently high or low degree of individuation. These referents appear to be the only ones that receive semantic agreement in Middle Dutch.

Semantic agreement does not only occur with fewer referents in Middle Dutch, but it seems to be generally less frequent compared to semantic agreement in present-day Dutch. Focusing on semantic agreement with masses, which is found in both time periods, a quantitative comparison can be made with Audring's (2009) findings for present-day Dutch. Table 13 below compares the ratio of semantic to lexical agreement with nouns referring to masses in the Middle Dutch texts with the ratio found for masses in the *Corpus Gesproken Nederlands* 'Corpus of Spoken Dutch' by Audring (2009). As the comparison involves written and spoken language, it should be viewed with some caution. However, it is generally believed that the difference between spoken and written language used to be much smaller than today. Therefore it is likely that the Middle Dutch texts closely resemble the spoken language of the time.

Table 13. Ratio of semantic to lexical agreement in Middle Dutch and present-day Dutch¹⁴

| MASS (CONFLICT) | Middle Dutch texts masculine and feminine antecedents | Present-day Dutch CGN common antecedents |
|----------------------------|--|---|
| non-lexical | 27% (N=31) | 50% (N=97) |
| lexical | 73% (N=82) | 50% (N=97) |
| <i>total</i> | <i>100% (N=113)</i> | <i>100% (N=194)</i> |

The data in Table 13 indicate that semantic agreement with masses occurs almost twice as often today (50%) as in Middle Dutch (27%). This difference is statistically significant ($\chi^2(1) = 14.958, p < 0.001$).

The comparison between Middle Dutch and present-day Dutch suggests that the ratio of semantic to lexical gender agreement has shifted over time. The fact that semantic agreement also occurs with objects in present-day Dutch may relate to this change. When semantic agreement is infrequent, it may only occur with referents that have an extremely high or low degree of individuation, such as humans and masses. When it increases, semantic agreement may occur with referents that have a more moderate degree of individuation as well, including objects.

Interestingly, De Vos & De Vogelaer (2011) found the same semantic agreement pattern as found for Middle Dutch in the Flemish dialect of Moerzeke. In their questionnaire study, participants showed semantic agreement with animate referents and with mass nouns, but not with objects. Although a quantitative comparison of the data is a bit difficult, as the Flemish data were elicited, it is interesting that semantic agreement with mass nouns occurred only in 15%

¹⁴ The numbers for present-day Dutch are recalculated from Audring (2009: 167). In Audring's study, references with masculine pronouns to common gender nouns were not counted as cases of lexical agreement, but were excluded from the quantitative analyses, because, strictly speaking, masculine pronouns do not agree with common gender. As this makes a quantitative comparison with the present data problematic, Audring's (2009: 92, 167) data were recalculated by counting these references as cases of lexical agreement. Also, Audring's results for specific and unspecific masses have been added up here.

CHAPTER IV

(N=117/651) of the cases.¹⁵ This indicates that, as in Middle Dutch, the frequency of semantic agreement in this dialect is generally lower than in present-day Standard Dutch. This difference suggests that there could indeed be a connection between the frequency of semantic agreement and whether or not semantic agreement also occurs with referents that take a middle position on the Individuation Hierarchy.

If the ratio of semantic to lexical agreement has shifted over time, this raises the question of why such a shift has occurred. Lexical nominal gender has not disappeared in Dutch, as it has in English for instance, because nominal gender is still clearly marked on adnominal elements in Dutch and speakers are well aware of whether a noun is common or neuter. Nevertheless, it does appear to have become easier to overrule the lexical gender of the noun in pronominal agreement.

There are two likely causes for this increase of semantic agreement. Firstly, the conflation of masculine and feminine nominal gender could play a role, in the way suggested by Audring (2006; 2009). Uncertainty about agreement with common gender nouns, whether they should be referred to with a masculine or a feminine pronoun, may cause speakers to rely on semantic agreement rather than lexical agreement. However, this uncertainty only exists with common gender nouns, not with neuter nouns. Therefore, the conflation of masculine and feminine gender does not directly explain the rise of semantic agreement with neuter nouns referring to objects, which is precisely the type of agreement that exists in present-day Dutch but was not found for Middle Dutch.

Another possible cause for the increase of semantic agreement is a change in the general visibility of lexical gender, as proposed by Kraaikamp (2012: 207-208). There could be a relation between how frequently lexical gender is marked in the noun phrase and how likely pronouns are to agree semantically instead of lexically. This factor is related to the loss of the masculine-feminine distinction in Dutch, but it is not the same. As Dutch lost the distinctive marking for masculine and feminine

¹⁵ This is a recalculated number from De Vos & De Vogelaer (2011: 253). The percentage of 15% (N=117/651) is the ratio of semantic to (attempted) lexical agreement with mass nouns. In this recalculated number, the use of neuter pronouns is contrasted with the use of masculine and feminine pronouns for masculine and feminine nouns. Masculine pronouns are sometimes used for feminine nouns and vice versa, but as these agreements are not considered to be semantically motivated but errors due to the incipient conflation of masculine and feminine nominal gender in the Moerzeke dialect, they are counted as (attempted) lexical agreements.

gender in adnominal elements, leading to the conflation of these two genders, Dutch also lost gender marking entirely in some adnominal elements. Consequently, a notable difference between the Middle Dutch and present-day Dutch gender system is not only the number of nominal genders, but also the number of elements that mark gender. Table 14 below gives an overview of the most frequent adnominal elements and whether or not they mark gender in the two stages of the language.

Table 14. Adnominal gender marking in Middle Dutch and present-day Dutch

| adnominal element | Middle Dutch | present-day Dutch |
|---------------------------|--------------|---------------------|
| definite article | √ | √ |
| indefinite article | √ | - |
| demonstrative determiners | √ | √ |
| possessive determiners | √ | only 1SG.PL |
| attributive adjectives | √ | only indefinite NPs |

In Middle Dutch, the indefinite article marks gender, but no longer in present-day Dutch, where it is invariantly *een*. All possessive determiners mark gender in Middle Dutch, whereas only the first person plural possessive marks gender today (common *onze* and neuter *ons* ‘our’, but invariant *mijn* ‘my’, *jouw* ‘your’, etc.). Although the attributive adjective still marks gender in indefinite noun phrases (common *-e* and neuter *-Ø* suffix), it is now invariantly inflected with *-e* in definite contexts. As adnominal elements generally show lexical and not semantic gender agreement, the loss of these gender markers has led to a decrease of lexical gender marking in the Dutch noun phrase. It is possible that the resulting lower visibility of lexical gender has reduced the weight of lexical gender in the existing competition between lexical and semantic agreement in the pronoun.

The results from a questionnaire study by De Vogelaer & De Sutter (2011) on semantic agreement in Flemish dialects support the idea that there is a relation between adnominal gender marking and semantic agreement. This study compared semantic neuter agreement with masses in West and East Flemish dialects that differ in the extent to which gender is marked in the noun phrase, in particular the distinction between masculine and feminine gender, which was the focus of the

CHAPTER IV

study. The distinction is sometimes still marked through an inflectional *-n* on the definite article and attributive adjective, but it is often deleted. The crucial difference between the dialects is whether the indefinite article still marks the distinction by a separate masculine form *ne(n)*, besides feminine *een* and neuter *e(en)*, or is invariantly *een*. Semantic agreement was found to be more frequent in those dialects where the indefinite article is invariant. De Vogelaer & De Sutter (2011) interpret this finding as an indication that semantic agreement is caused by the disappearance of the distinction between masculine and feminine nominal gender. However, it could be relevant that in the dialects without the masculine indefinite form, the indefinite article has become invariant and therefore no longer marks lexical gender at all. Considering this, the higher frequency of semantic agreement in these dialects may not be due to the loss of distinctive masculine and feminine forms, but rather to the complete loss of a gender marking element.

The picture that emerges is that changes in adnominal gender marking have allowed an existing semantic agreement pattern to gain ground in Dutch. A similar shift in agreement preference may have occurred in other Germanic varieties where agreement based on individuation has been observed. The relevant question that needs to be asked now is not what caused this type of agreement to develop, but rather what makes it surface more frequently.

6. Conclusion

This chapter reported on a corpus study of pronominal gender agreement in Middle Dutch. In present-day Dutch, pronouns show variation between lexical gender agreement and semantic gender agreement that is based on the degree of individuation of the referent. The same kind of semantic agreement is found in several other Germanic varieties. This study showed that this agreement variation existed already in Middle Dutch. In particular, neuter pronouns can agree semantically with referents that have a low degree of individuation, regardless of the gender of the antecedent noun. The fact that this occurs in Middle Dutch, where the original three-gender system is still in place, shows that this type of agreement did not develop in response to the change from three to two nominal genders, but existed already before this change. This finding supports the idea that the semantic interpretation of the genders along the lines of individuation is an old Germanic

feature which possibly dates back to Proto-Indo-European. What seems to have changed over time in Dutch is the frequency of semantic agreement, which has increased in proportion to lexical agreement. This shift in agreement preference is likely due to changes in adnominal lexical gender marking. The extent to which lexical gender is marked in the noun phrase and the resulting visibility of lexical gender may be a determining factor in the variation between the two kinds of agreement.