



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*. 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, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter II

The semantics of the Dutch gender system^{*}

Abstract

In this chapter it is argued that, although Dutch gender assignment is not systematically organized along semantic lines in the lexicon, the gender system has a semantic basis. Audring (2006, 2009) found that Dutch pronouns often show semantic gender agreement that is based on the degree of individuation of their referent. It is shown that this pronominal agreement behaviour relates to an existing semantic interpretation of the genders, as the same semantic distinction can also be found in the nominal domain. It surfaces particularly in cases where the gender of nouns is variable. The semantic interpretation of the genders possibly goes back their roots in Proto-Indo-European. It is argued that, ever since nominal gender became an invariable, lexically stored feature of nouns, the semantic basis of nominal gender assignment has become disrupted. This creates a conflict between lexical and semantic gender agreement in pronouns. It is suggested that the surfacing of semantic agreement in this conflict is connected with a reduced marking of lexical gender in the nominal domain.

1. Introduction

This study focuses on the gender system of Standard Northern Dutch (subsequently referred to as ‘Dutch’), a language with two nominal genders, common and neuter.¹ As in the other Germanic languages with two genders, common gender in Dutch is a conflation of former masculine and feminine gender. The two nominal genders are marked on determiners, attributive adjectives and pronouns. Only the personal pronoun still distinguishes three genders, masculine, feminine and neuter. Table 1 gives an overview of all the forms that mark gender in Dutch. Gender distinctions only exist in the singular. In the plural, the common gender forms are used.

^{*} A slightly different version of this chapter was published as: Kraaikamp, Margot. 2012. The semantics of the Dutch gender system. *Journal of Germanic Linguistics* 24.3. 193-232. <http://dx.doi.org/10.1017/S1470542712000074>

¹ Several Eastern and Southern Dutch dialects still distinguish three nominal genders, masculine, feminine, and neuter. In the Standard Northern Dutch spoken language, however, the distinction between masculine and feminine nouns is no longer made (see Audring 2009: 94).

CHAPTER II

Table 1. Dutch gender-marked forms

Determiners		common	neuter
Definite article ‘the’		<i>de</i>	<i>het/'t</i>
Demonstrative determiner	distal ‘that’	<i>die</i>	<i>dat</i>
	proximate ‘this’	<i>deze</i>	<i>dit</i>
Demonstrative determiner ‘such’		<i>zulke</i>	<i>zulk</i>
Interrogative determiner ‘which’		<i>welke</i>	<i>welk</i>
First plural possessive determiner ‘our’		<i>onze</i>	<i>ons</i>
Collective determiner ‘every’		<i>iedere/elke</i>	<i>ieder/elk</i>
Indefinite quantifying determiner ‘many a’		<i>menige</i>	<i>menig</i>
Adjectives		common	neuter
Attributive adjective (indefinite NPs only)		<i>-e</i> suffix	no suffix
Pronouns		common	neuter
Relative pronoun ‘that’		<i>die</i>	<i>dat</i>
Demonstrative pronoun	distal ‘that’	<i>die</i>	<i>dat</i>
	proximate ‘this’	<i>deze</i>	<i>dit</i>
Personal pronoun ‘it’ (nominative – oblique)		masc.	fem. ²
		<i>hij/'ie – hem/'m</i>	<i>zij/ze – haar/'r</i>
			<i>het/'t</i>

² As discussed later in the text, in spoken language, feminine pronouns are almost exclusively used for animate, female referents. Masculine pronouns generally agree with common gender nouns.

Nominal gender assignment is largely arbitrary in Dutch. Although there exists a limited set of semantic and formal regularities in nominal gender assignment, there is generally no motivation as to why one noun is common and another noun is neuter gender (see, for instance, Donaldson 1987, Haeseryn et al. 1997, Booij 2002).

In pronominal gender agreement, two types of agreement are found: ‘lexical gender agreement’, that is, agreement based on the lexical gender of the noun, and ‘semantic gender agreement’, that is, agreement based on the properties of the referent. An example of lexical gender agreement is shown in (1).

- (1) Ik kocht het **huis** van mijn grootouders
 I bought DET.N house(N) of my grandparents

 en wilde ’t opknappen.
 and wanted 3SG.N renovate
 ‘I bought my grandparents’ house and wanted to renovate it.’

When pronouns refer to animate entities, however, they tend to show agreement with the referent’s sex instead of with the lexical gender of the noun. An example is shown in (2), where the feminine pronoun is used, while the antecedent noun is neuter.

- (2) Het **meisje** ging snel naar school,
 DET.N girl(N) went quickly to school

 want **ze** was te laat.
 because 3SG.F was too late

 ‘The girl left for school quickly, because she was late.’

This kind of semantic agreement, with animate referents, is quite common in languages that otherwise show lexical gender agreement (see Corbett 1991).

Audring (2006, 2009) shows that another kind of semantic gender agreement occurs in Dutch, with inanimate entities. This semantic agreement involves an

CHAPTER II

opposition between masculine/common gender and neuter gender, and is based on the degree of individuation of the referent. Masculine and common gender pronouns, on the one hand, tend to be used with referents that have a high degree of individuation, that is, discrete, countable entities with clear boundaries, typically concrete objects, such as a cup or a book. Neuter pronouns, on the other hand, tend to be used with referents that have a low degree of individuation, that is, nondiscrete, uncountable entities with unclear boundaries, typically masses, such as sugar or water.

This agreement pattern was first described by Van Haeringen (1936, 1951), who observed that any object may be pronominalized by a masculine pronoun in Dutch (1936: 20), and that the neuter pronoun, instead of the masculine pronoun, may be used for non-neuter mass nouns (1951: 13). Audring (2006, 2009) has supported this early observation with spoken language data from the *Corpus Gesproken Nederlands* (CGN, ‘Corpus of Spoken Dutch’). Examples (3) - (6) below (adapted from Audring 2006: 95-99), demonstrate this semantic agreement for both personal and demonstrative pronouns. Example (3) shows the use of a masculine personal pronoun with a neuter noun that denotes a concrete object:

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

Dan moet ik **'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.’

Similarly, a common gender demonstrative pronoun is used with a neuter noun denoting a concrete object in example (4):

- (4) Heb jij een **fototoestel**?
 have you a camera(N)

Nee ik kan **die** van m'n broer wel lenen.
 no I can DEM.C of my brother PRT borrow

'Do you have a camera? No I can borrow my brother's.'

The other way around, in example (5), a neuter personal pronoun is used with a common gender noun that denotes a mass:

- (5) '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.'

Similarly, a neuter demonstrative pronoun is used with a common gender noun denoting a mass in example (6):

- (6) Wij hebben daar geen behang zitten
 we have there no wallpaper hanging

maar gewoon **verf** en **dat** is oranje.
 but just paint(C) and DEM.N is orange

'We don't have wallpaper there but just paint and that is orange.'

This type of agreement is frequent in Dutch. Audring (2009: 16) found that in 65% of the cases where there is a conflict between lexical and semantic gender (that

CHAPTER II

is, with common gender mass nouns and neuter gender count nouns), the personal or demonstrative pronoun does not agree with the noun's lexical gender.

Semantic agreement is found with the relative pronoun as well, but less frequently – in 25% of the cases where there is a conflict between lexical and semantic gender (Audring 2009: 160). It appears that the greater the distance between the antecedent and the pronoun – both structural/syntactic distance and plain word distance – the higher the likelihood of semantic agreement (Audring 2009: 165-166). Examples (7a) and (7b) (adapted from Audring 2006: 98-99), show the use of a common gender relative pronoun with a neuter noun, and a neuter relative pronoun with a common gender noun, respectively.

(7) a. Misschien is 't ook wel handig om zo'n **toestel**
maybe is it also quite handy to such.a device(N)

te nemen **die** je d'r zo af kan halen.
to take DEM.C you there simply off can take

'Maybe it is rather handy to take the kind of device that you can simply remove.'

b. dat er geen **apparatuur** onbeheerd is achtergebleven **dat**
that there no equipment(C) unattended is stayed.behind DEM.N

aan staat.
on stands

'That no equipment has been left unattended that is switched on.'

What these data show is the existence of semantic gender agreement in the pronominal domain of a language that otherwise has a lexical gender agreement system. Cross-linguistically, this phenomenon is not uncommon. The observation that semantic agreement can occur in the personal and demonstrative pronoun, and to a lesser extent also in the relative pronoun, while determiners show lexical gender

agreement is consistent with Corbett’s (1979) Agreement Hierarchy. Corbett observes that cross-linguistically, gender agreement can be represented on a hierarchy, with the likelihood of semantic gender agreement increasing the further one moves towards the right end of the scale:

(8) The Agreement Hierarchy (adapted from Corbett 1979)

attributive > predicate > relative pronoun > personal pronoun	
lexical	semantic
agreement	agreement

According to this hierarchy, determiners and attributive adjectives (adnominal elements) are most likely to agree with the noun’s lexical gender, while personal and demonstrative pronouns are generally the first to show semantic instead of lexical gender agreement. This pattern is also observed in spoken Dutch.

The question is why this semantic agreement occurs in Dutch, and why it is based on individuation. In this chapter, it is argued that individuation not only plays a role in the kind of pronominal gender agreement illustrated above, but in fact constitutes the basis of a semantic contrast that operates in the Dutch gender system as a whole. It is shown that the semantic agreement observed in pronouns relates to an existing semantic distinction between the genders, which arguably dates back to Proto-Indo-European. It is suggested that the surfacing of semantic agreement in pronouns is related to a reduced marking of lexical gender in the noun phrase. The next section discusses previous research showing that the semantic interpretation of the Dutch pronouns can be related to a semantic hierarchy called the Individuation Hierarchy, and that this interpretation of the pronouns is not unique to Dutch, but has been found in other Germanic languages as well.

2. The Individuation Hierarchy

The semantic agreement pattern in Dutch pronouns can be described in terms of a distinction between referents on the basis of their position on the Individuation Hierarchy (Siemund 2002, Audring 2006, 2009). This hierarchy, shown in (9), is

CHAPTER II

essentially an elaboration of the Animacy Hierarchy (Silverstein 1976), in which the inanimate section of the hierarchy is further subdivided.

(9) The Individuation Hierarchy³

Human	>	Animal	>	Object	>	Bounded	>	Specific	>	Unbounded
						abstract		mass		abstract, Unspecific mass

Entities on the left of the hierarchy have a high degree of individuation; they have a clearly bounded shape and specific characteristics. Entities on the right have a low degree of individuation; they have unclear boundaries and are less specified. Animate entities have a high degree of individuation. Next to them on the hierarchy are concrete objects, such as a book or a table. Abstracts are intangible things, concepts. They can be bounded and countable, such as a problem, an idea, or a question, or unbounded and uncountable, such as pain, love, or happiness. Specific masses are specific instantiations of masses, such as my mother's soup or his tea. Lowest on the hierarchy, with unbounded abstracts, are unspecific masses, such as snow, rain, or honey (Audring 2006: 102-103).

The boundary between the domain of masculine/common pronouns and the domain of neuter pronouns is situated somewhere between bounded and unbounded referents in Dutch. Siemund (2002: 224-225) shows that the pronominal gender systems of other Germanic languages can be related to this hierarchy as well. The

³ The hierarchy presented here is an adaptation of Audring's (2006: 102, 2009: 127) rendition of the Individuation Hierarchy. Audring uses an adapted version of Sasse's (1993: 659) Continuum of Individuality to explain the Dutch pronominal gender system. In Sasse's hierarchy, abstracts form a single category, while Audring distinguishes between bounded abstracts (such as a problem or a word) and unbounded abstracts (such as love or pain). In Audring's hierarchy, bounded abstracts are grouped together with objects (such as a book or a table). This grouping was motivated by the fact that in Audring's corpus data it was often difficult to tease apart concrete and abstract readings of nouns denoting objects (Audring 2009: 70). However, as the two categories are not equal in terms of individuation, they are separated in the version of the hierarchy adopted here. It is of course still an empirical question whether or not a distinction is made between the two kinds of referents in pronominal agreement.

pronominal gender system of Standard English is based on a split between animate and inanimate entities, whereby neuter pronouns are used for all inanimate referents. This constitutes a cut-off point that is situated further to the left than in Dutch, as shown in (10).

(10) Gender distinctions on the Individuation Hierarchy

	Human > Animal > Object > Bounded > Specific > Unbounded	
	abstract mass	abstract, Unspecific mass
<i>Dutch</i>		
fem/masc common	masc common	neuter
<i>English</i>		
fem/masc	neuter	

Just as with the grammatical markings on the Animacy Hierarchy, the number of gender distinctions increases towards the left. In both English and Dutch, a further distinction between masculine and feminine gender is made on the left end of the hierarchy, with humans and animals. Here pronominal gender varies between masculine and feminine depending on the sex of the referent.⁴

Not only in Dutch, but also in several other Germanic languages the boundary between common, or masculine and feminine, gender on the one hand, and neuter gender on the other hand, lies between bounded and unbounded referents. Siemund (2002) notes that in some regional varieties of English, pronouns show a count

⁴ The boundaries between the domains of the genders on the hierarchy are not absolute, but vary within certain ranges. In English, nonhuman animates are sometimes referred to with the neuter pronoun and sometimes with a gendered pronoun, masculine or feminine. Similarly, in Dutch, animals are sometimes referred to with a default masculine pronoun, and sometimes with a pronoun that agrees with their sex. (Of course, this distinction only becomes apparent with female referents). This variation seems to depend on how familiar the speaker is with the animal, for example, whether one refers to their own pet or to some other animal, or whether one refers to a large animal or a small insect. Obviously, whether or not the speaker knows the animal's sex also plays a role here.

CHAPTER II

versus mass distinction. Particularly, in the English dialect of West Somerset, masculine and feminine pronouns are used for count nouns (that is, with bounded referents), while neuter pronouns are used for mass nouns. This phenomenon in South-Western dialects was also observed by Jespersen (1924: 240):

Here and there we find a tendency to establish a grammatical distinction between thing-words (countables) and mass-words (uncountables) [...] In the south-western dialects of England “full shapen things” are referred to as *he*, acc. *en* (from OE *hine*) and take the pronominal adjuncts *theäse*, *thik*, while “unshapen quantities” are referred to as *it* and take *this*, *that*.

In several varieties of Frisian as well, the choice between a masculine or neuter pronoun depends on the bounded versus unbounded nature of the referent (Wahrig-Burfeind 1989). The West Frisian varieties of City Frisian (Stedsk) have two nominal genders, common and neuter. With mass nouns, regardless of whether they are common or neuter, the neuter personal pronoun *it*/*'t* is systematically used. Count nouns, however, are not always pronominalized by the masculine pronoun.

In the Frisian variety spoken on the island of Helgoland, a stricter bounded-unbounded distinction in pronouns is observed. In this variety, the neuter demonstrative/personal pronoun *deät* is always used for masses and abstracts, and the masculine personal pronoun *hi/hem* is systematically used for countable objects, regardless of whether the corresponding noun is common or neuter gender.

In the Danish dialect of West Jutland, the same bounded-unbounded distinction is made (Ringgaard 1973: 30-31). Countable objects are referred to by the common (uter) demonstrative pronoun *den*, while masses and abstracts are referred to by the neuter demonstrative pronoun *det*. What is particularly interesting about West Jutish is that the semantic distinction is not only made in pronouns, but also within the noun phrase. West Jutish is generally said to have lost its nominal gender distinctions, since all nouns combine with the determiner *æ* ‘the’ or *en* ‘a’ (for example, *æ/en hus* ‘the/a house’). However, the pronouns *den* and *det* are also used as demonstrative determiners, and within the noun phrase, they make the same semantic distinction as in the pronominal domain: nouns that denote countable objects combine with *den*, while nouns that denote masses or abstracts combine with

det, for example, *den hus* ‘that house’, but *det mælk* ‘that milk’, *det skrigen* ‘that shouting’. The determiner can even vary for the same noun depending on whether its referent is conceptualized as a countable item or a mass. The noun *fisk* ‘fish’, for instance, can combine with either *den* or *det* depending on whether the speaker refers to an individual fish or a type of food. It seems, therefore, that what West Jutish has lost is not gender in the nominal domain, but rather lexically stored nominal gender. Having lost that, the gender system of West Jutish is now entirely based on the semantics of individuation.

In contrast, in Dutch, lexically stored gender still plays a role, and pronominal gender can conflict with the lexical gender of a noun. The following section addresses the question of what makes it possible for semantic agreement to surface on a large-scale in a gender system that is otherwise based on lexical gender.

3. The rise of semantic agreement

Audring (2006, 2009) proposes that the semantic agreement observed in Dutch pronouns is caused by a mismatch between the numbers of nominal and pronominal genders. She argues that, since masculine and feminine gender have conflated into one common gender, there exists a mismatch between the nominal and the pronominal gender system: Only two genders exist in the nominal domain, viz. common and neuter, while three genders, viz. masculine, feminine, and neuter, exist in the domain of the personal pronoun. According to Audring, this means the following in practice:

[S]peakers have lost the knowledge about masculine and feminine nouns, while common gender nouns in Dutch are unable to govern pronoun choice in the usual way, as there is no personal pronoun corresponding directly to a common gender noun. This problem arises whenever a speaker wants to pronominalize such a noun.

(Audring 2006: 87-88)

However, this claim is not uncontroversial. Although personal pronouns distinguish three genders, while nouns distinguish only two, in practice, this does not necessarily involve a mismatch. It should be noted that in spoken language, the

CHAPTER II

feminine pronoun is, in fact, only used for semantic agreement with female referents. Audring's findings in the *CGN* confirm this: Feminine pronouns are found exclusively with female referents (Audring 2009: 92). It seems, therefore, that at least in spoken language, the feminine pronoun is reserved for semantic gender agreement and no longer plays a role in lexical gender agreement. If this is the case, then there is no mismatch between the numbers of nominal and pronominal genders with respect to the lexical gender agreement system.

Another question that arises with the mismatch analysis is how the new division of labour between the masculine and neuter pronouns is motivated by the proposed mismatch. Audring (2006, 2009) suggests that the semantic reinterpretation of these pronouns arises as a solution to the mismatch problem. However, the mismatch exists between the common gender noun on the one hand, and the masculine and feminine pronoun, on the other hand:

(11) Mismatch

common gender noun	→ masculine personal pronoun
	→ feminine personal pronoun
neuter gender noun	→ neuter personal pronoun

If this situation causes a reinterpretation of the pronominal genders, then a new division of labour between the masculine and feminine pronoun is expected. This has in fact already occurred in spoken language since the feminine pronoun became more and more reserved for semantic agreement with animate, female referents and the masculine pronoun came to replace it in other cases. This new division of labour between the masculine and feminine pronoun was already well on its way by the end of the 17th century (see Geerts 1966: 192-210, Schönfeld & Van Loey 1970: 119-120). These historical developments can be roughly schematized as in (12) below.

(12) Diachronic developments in Dutch lexical gender

Middle Dutch (ca. 1200–1500)

masculine gender noun → masculine personal pronoun
 feminine gender noun → feminine personal pronoun
 neuter gender noun → neuter personal pronoun

Early Modern Dutch (ca. 1500–1800)

→ masculine personal pronoun
 common gender noun → feminine personal pronoun
 neuter gender noun → neuter personal pronoun

Modern Dutch (ca. 1800–today)

common gender noun → masculine personal pronoun
 neuter gender noun → neuter personal pronoun

Starting in the Late Middle Dutch period, deflection gradually led to the conflation of masculine and feminine nominal gender. It was in the 17th century that a mismatch had developed between the available nominal and pronominal genders because masculine and feminine nominal gender had merged into one category. According to Geerts (1966), this mismatch was gradually resolved as the feminine pronoun was increasingly replaced by the masculine pronoun, while the feminine form became reserved for agreement with female referents.⁵

Therefore, it is not immediately clear at this point that it is the discrepancy between the numbers of nominal and pronominal genders that causes the reinterpretation of the masculine and neuter pronoun in semantic terms. Yet there does appear to exist at least an indirect relation between semantic agreement and a reduction in the number of nominal genders. Braunmüller (2000: 44), for instance,

⁵ It is possible that what Geerts (1966) describes as a general replacement of feminine pronouns by masculine pronouns is, in fact, an increase of semantic agreement with masculine pronouns. If this is the case, then the gradual replacement of feminine pronouns by masculine pronouns should be found only with individuated referents. As it is not certain whether this is the case, this question remains open for future research.

CHAPTER II

observes for the North Germanic languages that semantic gender agreement, on the basis of individuation or biological sex, only occurs in those varieties that have reduced the original three-way nominal gender system to a two-way or zero gender system, such as Standard Danish, Standard Swedish (two genders) and the West Jutish dialect (zero genders), and not in varieties that still employ a three-way gender system, such as Icelandic.

However, the distinction among the Germanic varieties is not that clear-cut. Semantic agreement is also found in Germanic varieties that have not lost a nominal gender. For example, in German – a language with three nominal genders solidly in place – semantic agreement occurs, albeit to a more limited extent than in Dutch. Pronouns can show semantic agreement with animate referents. For instance, with neuter nouns such as *Mädchen* ‘girl’, the feminine pronoun as well as the neuter pronoun is used (Corbett 1991: 228, Audring 2009: 193). It seems that semantic agreement is also possible at the other end of the Individuation Hierarchy. Audring (2009: 193) observes that neuter pronouns can be used with non-neuter nouns with referents of low individuation. She provides the following constructed examples:

(13) a. **Kartoffelsuppe?** Ja, **das** esse ich gern.
potato soup(F) yes DEM.N eat I with.pleasure
‘Potato soup? Yes, I like to eat that.’

b. Sie mag keinen **Jazz**,
she likes no.MASC.SG jazz(M)

aber ihr Freund hört **das** immer.
but her friend listens DEM.N always

‘She doesn’t like jazz, but her boyfriend listens to it all the time.’

So, at least with referents on the outer ends of the Individuation Hierarchy, where the potential conflict between semantic and lexical gender is the largest, semantic agreement appears to be possible in German. This suggests that semantic agreement

is not an all-or-nothing phenomenon in the Germanic varieties but rather a matter of degree.

This view also emerges from studies on dialects of Dutch. Using written questionnaires, De Paepe & De Vogelaer (2008), De Vos (2009), and De Vogelaer & De Sutter (2011) found that in East Flemish dialects with three nominal genders, the neuter pronoun is often used with masculine and feminine nouns referring to abstracts or masses.⁶ Comparing the results for different East and West Flemish dialects, De Vogelaer & De Sutter (2011) further observed that this type of agreement is more frequent in those dialects where the distinction between masculine and feminine gender is not marked on the indefinite article, but only, in a limited number of phonological contexts, on the definite article and the attributive adjective.

Considering this, it is possible that semantic agreement correlates not with the loss of a nominal gender per se, but rather with the weakened state of the lexical gender system in general. Compared to a language like German, where semantic agreement is considerably less frequent than in Dutch, Dutch not only has a reduced number of nominal genders, but it also has a reduced gender agreement system with respect to the number of elements that show gender agreement. Indefinite determiners no longer show gender agreement in Dutch, the possessive determiner shows very limited gender agreement, and attributive adjectives show gender agreement only when they appear in indefinite noun phrases (see Table 1 in Section 1). Importantly, these are all adnominal elements, which constitute agreement targets that most reliably express lexical gender (see the Agreement Hierarchy in Section 1). Considering this, it seems possible that the high frequency of semantic agreement in Dutch pronouns is connected to the reduced marking of lexical gender on adnominal elements.

⁶ The reverse agreement pattern, that is, the use of a masculine pronoun with a neuter noun referring to an individuated item, was not found by De Paepe & De Vogelaer (2008) and did not emerge clearly in De Vos 2009. De Vogelaer & De Sutter (2011), therefore, chose not to investigate this type of agreement in their study. The absence of such an agreement pattern may indicate a developmental path in a change towards more semantic gender agreement. Possibly, neuter pronouns are used with referents of low individuation sooner than masculine/common pronouns are used with referents of high individuation.

CHAPTER II

An important question that remains is why semantic agreement in pronouns occurs along the lines of individuation. Apparently, there exists a tendency in the Germanic languages to make a gender distinction between individuated and non-individuated referents. It should be explained where this semantic interpretation of the genders comes from, more specifically, why the masculine/common pronoun is connected with a high degree of individuation and the neuter pronoun with a low degree of individuation. Perhaps the answer to this question could also explain why a new division of labour between masculine/common gender and neuter gender is found, while neuter gender was never involved in a mismatch between nominal and pronominal gender. In the following sections, it is argued that the individuation distinction found in pronominal agreement relates to an existing semantic distinction between neuter and non-neuter gender in Dutch. First, section 4 shows that the neuter gender pronoun is more generally connected with the notion of low individuation.

4. The semantics of neuter gender

It can be argued that the use of the neuter pronoun with referents of low individuation is connected to a much larger semantic role of neuter gender. The neuter pronoun is used not only for unbounded entities or lexically neuter nouns. It is also used with non-nominal antecedents that denote activities, processes, states, properties, events, or propositions, henceforth called ‘nonentities’. Consider the use of neuter pronouns in examples (14) – (17) below.

(14) a. De tuin is compleet vernield.

DET.C garden is completely ruined

Wie heeft **dat** gedaan?

who has DEM.N done

‘The garden is completely ruined. Who did that?’

b. Ik ga het huis opruimen.
 I go DET.N house clean.up

Dat is hard nodig
 DEM.N is quite necessary

en **het** zal me veel tijd kosten.
 and 3SG.N will me much time cost

‘I’m going to clean up the house. That is quite necessary and it will take me a lot of time.’

In example (14), the neuter pronoun refers to activities. In (14a), the neuter demonstrative pronoun refers to the activity of ruining the garden, and in (14b), the neuter demonstrative pronoun and the neuter personal pronoun both refer to the activity of cleaning up the house.

In (15a), the neuter demonstrative pronoun refers to a state, that is, being grey, and in (15b) it refers to a property, that is, happy:

(15) a. De broek is nu grijs, maar
 DET.C pants is now grey but

dat verandert nog door de verf.
 DEM.N change PRT by the dye

‘The pants are grey now, but that will change with the dye.’

b. Frans is vrolijk vandaag. **Dat** is hij eigenlijk altijd.
 Frans is happy today DEM.N is he actually always.

‘Frans is happy today. He always is, actually.’

CHAPTER II

In (16), the neuter personal pronoun and the neuter demonstrative pronoun refer to an entire proposition:

(16) a. De tuin is compleet vernield
DET.C garden is completely ruined

en jij ziet **het** nu pas.
and you see 3SG.N now only

‘The garden is completely ruined and you only see it now.’

b. Een kwispelende hond is niet agressief.
a tail.wagging dog is not aggressive

Dat weet iedereen.
DEM.N knows everyone

‘A dog wagging its tail is not aggressive. Everyone knows that.’

The neuter personal pronoun is also used when there is no (clear) referent at all, as in (17), where it functions as an expletive:

(17) a. **Het** regent.
3SG.N rains
‘It is raining.’

b. Hoe gaat **het** met je?
how goes 3SG.N with you
‘How are you doing?’

This use of neuter gender with nonentities is quite common in the Indo-European languages that have a neuter gender form available. Corbett (1991: 204-205) discusses the use of neuter gender in several Indo-European languages with

what he refers to as non-prototypical controllers, that is, elements with no specified gender, such as clauses and infinitival phrases. In the three-gender system of Russian, for example, the neuter verb form is used to agree with infinitival subjects, as in example (18a) (adapted from Corbett 1991: 204). Similarly, in Serbo-Croat, it is the neuter pronoun that agrees with predicative adjectives, as in example (18b) (adapted from Corbett 1991: 205).

(18) a. Dozvonit'sja byl-o problemoj.

to.ring.through was-N.SG problem

'To ring through was a problem.'

b. kod njih čovek mora biti lukav,
with them person must be cunning

a ja to nisam

but I DEM.N.SG am.not

'With them a person must be cunning, but I am not (that).'

Corbett (1991: 206) notes that this use of the neuter form cannot be explained by markedness theory, since it is not immediately clear that neuter is the unmarked form in the languages in question. This is certainly also the case for Dutch, where common gender appears to be the unmarked form. Common gender is always used in the plural, where gender distinctions are neutralized, which makes it the most frequently occurring form in the determiner paradigm. Furthermore, the large majority of nouns – approximately three out of four – are common gender, while only one out of four nouns is neuter (Van Berkum 1996).⁷ During the acquisition of nominal gender, the common determiner is also largely overgeneralized before the

⁷ Van Berkum's (1996) counts are based on the CELEX lexical database of Dutch. His lemma type count rendered 72% common gender entries and 27% neuter gender entries, a lemma ratio of approximately 3:1. Taking token frequencies into account (singular forms only), the distribution of the nominal genders in a running text was estimated: 67% of encountered singular nouns are common gender versus 33% neuter gender nouns, a ratio of approximately 2:1.

CHAPTER II

neuter determiner is used (Van der Velde 2004, Blom, Poliřenská, & Weerman 2008). However, despite this, in several ways, default status of common gender in Dutch, it is neuter gender that is used in reference to nonentities.

Corbett (1991: 205-206) proposes to explain this use of neuter gender in languages such as Russian and Serbo-Croat in terms of the distribution of the genders in the lexicon. He notes that in these languages, almost all nouns that denote humans are masculine or feminine, while nouns that denote inanimates can be masculine, feminine, or neuter. Given this distribution, Corbett (1991: 206) argues, the selection of neuter gender for agreement with non-nominal antecedents could be considered the most appropriate option in semantic terms. This analysis could be used to explain the use of neuter with nonentities in Dutch as well, as a similar distribution of common and neuter gender exists there: nouns that denote humans tend to be common rather than neuter gender in Dutch (see Haeseryn et al. 1997).

However, the facts could be approached from a different angle. Instead of departing from the distribution of the genders in the lexicon, one could consider the use of neuter gender with nonentities to reveal something about the semantics of neuter gender itself. Returning to the Individuation Hierarchy discussed above, it can be argued that neuter gender is more extensively connected with the meaning of low individuation than the present hierarchy reveals. The larger semantic domain of neuter gender becomes more apparent if the Individuation Hierarchy is extended on the right side with nonentities – the category that can be considered the least individuated in the system.

(19) The Individuation Hierarchy (extended)

Human	>	Animal	>	Object	>	Bounded	>	Specific	>	Unbounded	>	Nonentity
						abstract		mass		abstract,		
										Unspecific		
										mass		

Under this approach, the observed division of labour between masculine/common and neuter pronouns in Dutch can be said to reside within a broader system of semantic categories: While common gender is used for referents of high individuation, neuter gender is used for a wide range of referents of low

individuation, from abstracts/masses to all other possible referents that are not bounded in nature, such as events, activities, and propositions.⁸

Considering this function of neuter gender, it is interesting to note that there are several Indo-European languages which otherwise have a two-way masculine-feminine gender system, that still use a neuter form with nonentities (Corbett 1991: 214-215). Corbett's examples include Portuguese, the Surselvan dialect of Romansh, and Spanish. Below is a Spanish example from Corbett 1991: 214.

- (20) Antes me gustaba mucho ir a los partidos de futbol,
 before me was.pleasing much go to the games of football
- pero todo **ello** ya no me interesa.
 but all it any.more not me interests

'I used to be very keen on going to football matches, but all that doesn't interest me anymore.'

The neuter pronoun *ello* 'it, that' refers to the activity of going to football matches. This kind of reference is the only function of the neuter form in Spanish, for there is no class of neuter nouns in the language. It seems that, while the original three-way gender system of Romance has been reduced to two, neuter gender has been preserved for this type of agreement (Corbett 1991: 215). With the proposed extension of the Individuation Hierarchy in mind, it can be said that the Spanish gender system operates on this hierarchy, with the boundary between neuter and feminine/masculine gender located between masses and nonentities.

So far, it has been shown that the individuation distinction observed in Dutch pronominal gender agreement is connected to an existing semantic distinction between neuter and non-neuter pronouns. In the following section, it is shown that

⁸ See Roodenburg & Hulk 2009 for a comparable, but different, analysis of neuter in Dutch. Following Picallo's (2008) analysis of neuter in Spanish and Catalan, they argue that neuter is not a true gender, but rather an underspecified element, or default, used in the absence of gender and number features on the antecedent. The present analysis differs from this view in that neuter is not considered a default element, but a regular gender that covers a certain semantic domain.

CHAPTER II

this semantic gender distinction is not restricted to pronouns, but also exists in the nominal domain.

5. The individuation distinction in the nominal domain

It is generally assumed that Dutch nominal gender assignment is not based on any semantic principle, but that it is arbitrary. Although this is largely the case, as discussed further in Section 6.2, in certain cases the semantics of individuation appear to play a role in nominal gender assignment.

5.1. Nominalized adjectives

The individuation distinction is reflected in the choice of the determiner that appears with nominalized adjectives. In Dutch, adjectives can be nominalized by adding a determiner to the inflected form of the adjective. This process is fully productive, that is, any adjective can be nominalized in this way. If there is a nominal antecedent, the determiner-adjective construction is generally analysed as an elliptical phrase, in which the gender of the determiner simply corresponds to that of the elided noun, as in (21).

- (21) a. Er was een rode auto en een zwarte auto.
there was a red car(C) and a black car(C)

Ik heb **de rode** genomen.

I have DET.C red taken

‘There was a red car and a black car. I took the red one.’

- b. Er was een dik boek en een dun boek.
there was a thick book(N) and a thin book(N)

Ik heb **het dunne** gekocht.

I have DET.N thin bought

‘There was a thick book and a thin book. I bought the thin one.’

However, in spoken language, the determiner in this construction can also agree with the degree of individuation of the referent rather than with the lexical gender of the elided, or antecedent, noun. If the elliptical phrase refers to an object, as in (21) above, the determiner could also be common gender, regardless of the neuter gender of the corresponding noun.⁹

- (22) Er was een dik boek en een dun boek.
 there was a thick book(N) and a thin book(N)

Ik heb **de** **dunne** gekocht.
 I have DET.C thin bought

‘There was a thick book and a thin book. I bought the thin one.’

If the determiner-adjective construction occurs without a corresponding noun in the linguistic context, the individuation distinction becomes even more apparent. In this case, individuation alone appears to determine the form of the determiner. The neuter determiner is always used if the referent is a nonentity, as in example (23a-b), while the common determiner is always used if the referent is animate, as in example (23c-d).¹⁰

- (23) a. De lamp verandert van kleur. Dat is **het** **leuke** ervan.
 DET.C lamp changes of color that is DET.N fun of.it
 ‘The lamp changes color. That is the fun of it.’

b. **Het** **belangrijkste** is dat hij weer kan werken.
 DET.N most.important is that he again can work
 ‘What’s important is that he can work again.’

⁹ The fact that the determiner can show semantic agreement with the referent instead of agreement with the supposed elided noun suggests that these constructions are not in fact elliptical phrases but true nominalizations.

¹⁰ Nominalized adjectives with referents located between these two extremes do not usually occur without an antecedent noun or a physical referent in the context.

CHAPTER II

- c. Eén van ons zal **de gelukkige** zijn.
one of us will DET.C lucky be
'One of us will be the lucky one.'
- d. Het is altijd zo dat **de sterkste** wint.
it is always so that DET.C strongest wins
'It is always the case that the strongest wins.'

These nominalized adjectives can be considered ad-hoc, nonlexicalized noun formations. Therefore, they do not have a lexically stored gender and in that case, gender assignment appears to be based on individuation.

The following subsections show that the individuation distinction can, in some cases, be found with lexically stored gender as well.

5.2. Double gender nouns

Lexically stored nominal gender shows much arbitrariness. Compare, for example, *de trui* 'the sweater', *de jurk* 'the dress' but *het hemd* 'the shirt', and *de stoel* 'the chair', *de kast* 'the cabinet' but *het bed* 'the bed'. These objects are the same in terms of individuation, and they do not seem to differ systematically in any other way. Yet the nouns take different genders.

Nevertheless, the individuation distinction can be observed with lexical gender in some cases. It is visible with nouns that can be either common or neuter gender, with a concomitant difference in meaning. In such minimal pairs, the neuter variant often has a less individuated referent than the common gender variant. The most straightforward examples are found among nouns denoting materials. When a noun can denote both a material and an object made of that material, it is usually neuter when it refers to the material and common when it refers to the object (Haeseryn et al. 1997). In the minimal pairs in (24), the neuter variant is a mass noun referring to a material, while the common gender variant is a count noun referring to an object.

(24)	mass	count	
	<i>het steen</i>	<i>de steen</i>	‘stone’
	<i>het diamant</i>	<i>de diamant</i>	‘diamond’
	<i>het kurk</i>	<i>de kurk</i>	‘cork’

The neuter noun *zout* ‘salt’ shows a similar gender alternation. In spoken language, when a speaker refers to the salt in a saltshaker, typically at the dinner table, a common determiner is sometimes used. This gender alternation is illustrated in (25).

- (25) a. **Het zout** van de zee prikt in mijn ogen.
 DET.N salt of the sea stings in my eyes
 ‘The salt of the sea stings my eyes.’
- b. Mag ik **de zout**?
 may I DET.C salt
 ‘Can I have the salt please?’

(Example (25b) is from <http://www.onzetaal.nl/advies/zout.php>)

In a context like (25a), where *zout* has an unbounded interpretation, the noun is always neuter, while in a context such as (25b) the noun can take a common determiner. This use of the common determiner is not readily explained by some kind of ellipsis from a compound, as a likely candidate would in that case be *zoutvaatje* ‘saltshaker’, which is neuter. It appears that in this context, the preference for common gender is due to the individuated nature of salt in a saltshaker.¹¹

¹¹ See Semplicini 2012 for a discussion of double gender nouns for which gender variation is not generally connected to a semantic difference, but to geographical/inter-speaker variation. Semplicini argues that, in some cases, the gender variation for these nouns depends on subtle differences in the degree of individuation of the referent in a particular context.

CHAPTER II

5.3. Nouns from the same semantic domain

Besides gender alternations for the same noun, the gender of nouns that belong to the same semantic domain can also show the individuation distinction. Zubin & Köpcke (1986) provide an elaborate description of nominal gender in German, in which they examine the semantic differences between neuter and non-neuter nouns in several semantic fields. They show that nouns referring to superordinate categories, that is, overarching terms that subsume more specific categories, are usually neuter, while nouns referring to more specific categories, that is, more basic level terms, tend to be masculine or feminine.

This gender distinction between neuter and non-neuter can be observed in Dutch as well. Some examples are listed in (26).

- (26) *het instrument* ‘the instrument’ – *de gitaar* ‘the guitar’
 – *de piano* ‘the piano’
 – *de trompet* ‘the trumpet’
 – *de fluit* ‘the flute’
- *de appel* ‘the apple’
 – *de peer* ‘the pear’
 – *de kers* ‘the cherry’
 – *de banaan* ‘the banana’
- *de tafel* ‘the table’
 – *de stoel* ‘the chair’
 – *de bank* ‘the couch’
 – *de kast* ‘the cabinet’

Zubin & Köpcke (1986: 144) observe that neuter gender is associated with a lesser differentiation in the lexical hierarchy and a corresponding lack of precise characteristics, while masculine and feminine gender (common gender in Dutch) are linked to a greater differentiation in the lexical hierarchy and hence more precisely defined characteristics. This distinction relates to the Individuation Hierarchy: The referents of superordinate terms are less specific than those of basic level terms.

Therefore, superordinate terms are associated with a lower degree of individuation than their corresponding basic level terms. Thus, the gender distinction between the two categories follows the individuation distinction.

The examples of the individuation distinction in lexical gender make up only a limited proportion of the lexicon, however. Therefore, they cannot be said to represent a consistent pattern in lexical gender assignment.¹² Nevertheless, their existence, together with the gender assignment pattern observed in productive nominalizations, indicates that the semantic agreement pattern found in pronouns is not an isolated phenomenon within the Dutch gender system. Rather, the behaviour of pronouns reveals a semantic principle that operates in the gender system as a whole.

Possibly, the semantics of individuation have always played a role in the gender system. The individuation distinction in the nominal domain may have been inherited from an earlier stage of the language when nominal gender assignment had a more transparent semantic basis. In general, it is likely that the assignment of nouns to different genders has not always been arbitrary, but was originally based on some semantic distinction between their referents. The following section discusses the possible semantic basis of the Indo-European genders, and how this basis may have been disrupted.

6. The semantic basis of nominal gender and its disruption

6.1. The original meanings of the Indo-European genders

Although the Indo-European genders are traditionally labeled masculine, feminine, and neuter, it is unlikely that gender assignment in the Indo-European languages originally involved a distinction on the basis of biological sex. Fodor (1959) notes that these gender labels were adopted in the heroic age of grammar writing for the

¹² Considering the fact that lexical gender otherwise shows much arbitrariness, it is interesting to note that the results of a pilot study by Reijers (2008) suggest that in the acquisition of neuter lexical gender, individuation plays a role. The acquisition of lexical gender in Dutch starts with an almost complete overgeneralization of common gender, followed by a gradual acquisition of neuter gender for neuter nouns (see Van der Velde 2004). Reijers's study shows that Dutch children acquire neuter gender for mass nouns, such as *gras* 'grass', earlier than for count nouns, such as *boek* 'book'. This suggests that, at least during acquisition, an association exists between neuter lexical gender and low individuation.

CHAPTER II

modern European languages. The 19th century grammarians were proud to discover that the same three genders existed in the modern European languages as in the honored classical languages. Driven by the idea of anthropomorphism, they thought it “reasonable and consequently superb that natural gender [that is, biological sex, MK] was reflected in language through such subtle grammatical devices” (Fodor 1959: 3).

One notable proponent of the natural gender basis was Jacob Grimm. In Grimm 1831, he claimed that the gender system was originally a reflection of biological sex, which was subsequently extended to nouns with inanimate referents through the human imagination. This would have involved the personification of inanimate objects, which were classified as having either typically masculine (such as big, firm, active), feminine (such as small, weak, passive) or sexually indifferent (such as general, underdeveloped) characteristics.

Brugmann (1889, 1897), however, opposed this view and claimed that the association of the genders with biological sex was only a secondary phenomenon. He argued that the class of nouns labeled “feminine”, that is, the nouns ending in $-\bar{a}$ and $-\bar{i}$, was not made up of nouns that referred to female animate beings, but of collective and abstract nouns. This view still stands today. Furthermore, evidence has now converged on the view that Proto-Indo-European originally had only two genders, common and neuter, and that the third, feminine, gender developed later (Matasović 2004: 33, 165, Schwink 2004: 12).¹³ Therefore, it is unlikely that biological sex formed the semantic basis of the Indo-European genders.

At the same time, there are indications that the genders were connected with different degrees of individuation. Lehmann (1958: 188-192) reconstructs a paradigm of nominal suffixes in Proto-Indo-European that gave rise to the genders. Making use of insights from phonological theory, Lehmann argues that the long vowel endings $-\bar{a}$, $-\bar{i}$, and $-\bar{u}$ resulted from a merge of different short vowels with the

¹³ Much research has been done on how the third, feminine, gender has developed. For an overview, see, for instance, Ledo-Lemos 2003. It seems clear that there is a connection between feminine gender and the meanings “abstract” and “collective”. Also, a connection appears to exist between the feminine gender suffix and the neuter nominative-accusative plural ending. However, the exact nature of the connection between feminine gender on the one hand, and the neuter plural and the abstract/collective meaning on the other hand is debated. See, for instance, Luraghi 2009 for a recent proposal regarding this issue.

same ending, the laryngeal *-h*. Thus, three nominal endings could be distinguished that gave rise to the genders: *-s* (“masculine”), *-h* (“feminine”) and *-m* (“neuter”). According to Lehmann, these suffixes were systematically connected with certain meanings. They corresponded to the respective meanings “individual”, “collective”, and “resultative”, later including “mass”.¹⁴

Every noun could be inflected with each of these suffixes, which modified the noun’s meaning in a specific way. For example, the “feminine” form *himā* (<*himah*) had the collective meaning ‘season of cold and frost’ or ‘winter’, the “masculine” form, *himás*, referred to a singular occurrence of cold or frost, while the “neuter” form, *himam*, meant “snow”, which could be paraphrased as “the result of cold”. Thus, the same noun could have different forms associated with systematic differences in meaning.

Following Lehmann (1958), Leiss (2000) argues that remnants of this meaningful gender alternation can still be found in Old High German. In Old High German, many nouns occur in two or three different genders, which is usually interpreted as arbitrary variation leading up to a gender change. Leiss argues, however, that these different genders can often be related to the categorical meanings count (masculine), collective (feminine), and mass (neuter). An example of such gender variation is the Old High German noun that means “wood” or “string of wood”. The mass meaning was expressed by the neuter form *witu* ‘wood’, while the collective meaning was expressed by the feminine form *witta* ‘string of wood, headband’.

Moreover, in a reconstructed lexicon of Proto-Indo-European nouns, Matasović (2004: 133-134) observes semantic tendencies in gender assignment that operate along the lines of individuation. Nouns with referents at the leftmost end of the Animacy Hierarchy (humans, animals, and most plants) are consistently common

¹⁴ Lehmann (1958) argues that as the case and gender congruence system developed in Proto-Indo-European, mass nouns (such as the Sanskrit nouns *ápas* ‘work’, *áyas* ‘metal’, and *pásu* ‘stock’) were aligned with the class of resultative nouns ending in *-m*. This alignment happened because the two groups of nouns became formally similar. In addition to indicating resultatives, the ending *-m* began to mark accusative case. Consequently, resultative nouns, ending in *-m* in the nominative, had the same form for both subject and object. This aligned them with mass nouns, which did not distinguish subject and object forms either (Lehmann 1958: 197).

CHAPTER II

(masculine/feminine) gender, while nouns with referents at the rightmost end (fluids and masses) are consistently neuter. Nouns with referents located between these two extremes show a less consistent assignment pattern. Interestingly, however, nouns denoting abstracts are always assigned either feminine or neuter gender.

Considering the above, the semantic basis of gender classification in Indo-European likely involved a distinction between the meanings countable/individual, mass, and collective/abstract. The gender distinction on the basis of biological sex would have developed later, in addition to these original meanings.

(27) The semantic basis of Indo-European gender

count/individual	collective/abstract	mass
male individual	female individual	sexless
↓	↓	↓
'masculine' gender	'feminine' gender	'neuter' gender

Thus, each of the Indo-European genders corresponded to a different degree of individuation. Masculine gender (countable entities) was associated with the highest degree of individuation, while neuter gender (masses) was associated with the lowest. Feminine gender (collective/abstract entities) was associated with a degree of individuation that lies between the two extremes.

These semantic distinctions still appear to play a role in the Dutch gender system today. Particularly in the choice of pronouns, the individuation distinction, as well as the biological sex distinction, are involved. As discussed in the previous section, the semantics of individuation can be observed in lexical gender as well, although to a limited extent. The gender of many nouns does not follow the individuation distinction and appears to be arbitrary in terms of semantics. In the following section, it is argued that this situation may have resulted from a gradual, though large-scale, disruption of the semantic basis of gender classification in the lexicon.

6.2. *Disruption*

The fact that nowadays the gender of most Dutch nouns is arbitrary shows that there is no longer a semantic principle underlying lexical gender. The original semantic classification has somehow become largely disrupted in the nominal domain. One factor that must have been particularly important in this process is that nominal gender has become an invariable, lexically stored feature of the noun rather than a variable feature that modifies the noun's meaning.¹⁵ It is possible that once nominal gender becomes lexicalized, the semantic principle underlying gender classification becomes susceptible to disruption. In time, a noun tends to change its meaning, and if its gender remains the same, the connection between the meaning of the noun and its lexical gender can be lost.

Importantly, meaning shifts along the Individuation Hierarchy are not uncommon. One common type of meaning change involves the development of a more concrete, or specific meaning from a more abstract, or general one (narrowing). Another common change proceeds in the opposite direction and involves the development of a more abstract, or general meaning from a more concrete, or specific one (widening) (Williams 1986: 170-177). These semantic changes are particularly disruptive in a gender system based on individuation, as they involve shifts along the Individuation Hierarchy, from a lower to a higher degree of individuation in the case of narrowing, and from a higher to a lower degree of individuation in the case of widening.

An example of the shift towards a higher degree of individuation is the development of the neuter noun *brood* 'bread, loaf of bread'. Nowadays, it can either be a mass noun, denoting the substance bread, or a count noun, denoting a loaf of bread. This noun dates back to Old Germanic, possibly Proto-Germanic (Philippa

¹⁵ Why gender has become a fixed, invariable nominal feature is still an open question. One factor that may have played an important role is the emergence of the definite article. The grammaticalization of the gender-marked demonstrative into the definite article in Old Germanic entailed that the determiner became obligatory in a wider range of contexts. Consequently, nouns would appear with a gender-marked determiner more and more frequently. It is likely that even with variable nominal gender, most nouns would appear more frequently in one gender than in another, as certain meanings would likely be more common than others. If a noun frequently appeared with a particular gender form of the article, this article-noun combination could have become entrenched as a single unit, thus contributing to the lexicalization of nominal gender.

CHAPTER II

et al. 2007). It is likely that it has always been neuter, for according to the dictionary of Middle Dutch (*Cd-rom Middelnederlands* 1998), it was neuter in Middle Dutch and in Middle High German. According to Philippa et al. (2007), the noun originally meant “food” or “bread in general”. Only later did it develop the meaning “a piece of bread”. Thus, the neuter gender of the noun originally corresponded with the low degree of individuation of its referent. However, since the noun has also come to denote a unit of bread, while retaining its gender, there is now a discrepancy between the individuated meaning and the neuter gender of the noun. It is possible that over time, many more nouns have changed their meanings or have developed additional meanings along similar lines, while maintaining their lexically stored gender.

A more drastic disruption of the connection between semantics and gender occurs when a productive suffix that is systematically connected with a specific gender changes its meaning. A salient example is the diminutive suffix *-(t)je*, which goes back to Old Dutch *-kī* and Proto-Germanic *-kīn-* (Philippa et al. 2007). This suffix is systematically connected with neuter gender. It is likely that the diminutive originally denoted a young child (see Jurafsky 1996). With this denotation, the neuter gender corresponds with the absence, or irrelevance, of a specified sex for the referent. However, gradual metaphorical extension of this original meaning has led to several additional meanings, including partition. The suffix has often been used to derive count nouns from mass nouns, and it is still moderately productive in this function. Some examples of the partitive meaning of the Dutch diminutive suffix are given in (28).

- (28) *bier* ‘beer’ – *bier-tje* ‘glass/unit of beer’
snoep ‘candy’ – *snoep-je* ‘piece of candy’
krijt ‘chalk (mass)’ – *krijt-je* ‘chalk (writing utensil)’
vilt ‘felt’ – *vilt-je* ‘coaster’
vuil ‘dirt’ – *vuil-tje* ‘speck of dust’

Although the partitive meaning of the suffix follows quite naturally from its original semantics, shifting from ‘child’ to ‘smallness’ to ‘partition’, it creates a structural violation of the connection between meaning and gender. The concrete, countable

meaning of the partitive nouns on the right in (28) in no way corresponds with the semantics of neuter and yet the diminutive suffix systematically connects them with neuter gender.

The example of the diminutive suffix shows that the connection between gender and form is very strong. This connection is maintained even when the meaning of a morpheme changes and no longer motivates its gender or even contradicts it. This persistence of gender-form pairings can be an important factor in the gradual disruption of the semantic basis of nominal gender classification.

At the same time, this strong connection between gender and form can be responsible for the development of more formal, that is, phonological and/or morphological, gender systems. For example, although the Romance languages, like the Germanic languages, no longer have a transparent semantic basis for nominal gender assignment, many of them do show a systematic connection between gender and form. In Spanish, for instance, nouns ending in *-a* are usually feminine, while nouns ending in *-o* are usually masculine (Mel'čuk 1974). Thus, Spanish can be considered to have a formal gender assignment system.

In this respect it is interesting to note that even in gender assignment systems that are predominately based on form, there always remains a semantic core to nominal gender assignment (Corbett 1991: 8, Corbett & Fraser 2000: 294). In their typology of gender systems across the world, Corbett (1991) and Corbett & Fraser (2000) broadly distinguish between two types: (predominantly) semantic systems, in which nominal gender is assigned completely or largely according to meaning, and predominantly formal systems, in which gender is assigned largely according to the morphological or phonological properties of a noun. Interestingly, while there exist purely semantic gender systems, there appear to be no systems based solely on form.¹⁶ Semantics always play a role in gender assignment, and importantly, the

¹⁶ The question arises as to why some languages have been able to maintain their purely semantic gender systems, while Indo-European has not been able to do so. If the account of the developments in Indo-European proposed here is correct, purely semantic gender systems are expected to have at least one of the following characteristics: the gender system is relatively young, so that large-scale disruption through meaning change has not occurred yet; nominal gender is variable, so that if the meaning of a noun changes (diachronically or in context), its gender changes accordingly; the semantic basis for gender assignment is not particularly susceptible to disruption because meaning changes tend not to cross the semantic boundaries of the genders. Although this issue deserves a more thorough investigation, it is

CHAPTER II

formal and semantic assignment criteria always overlap to some extent (Corbett 1991: 63). This observation supports the idea that gender systems are primarily based on semantics, while formal regularities may develop in tandem. If the semantic basis for gender assignment becomes opaque, the nominal form itself may eventually become the basis for gender assignment, resulting in a formal gender system

However, although some formal regularities exist in nominal gender assignment, the Germanic languages have not developed transparent formal gender systems. The development of a formal system could have been prevented by the Germanic nominal stress pattern. Since in Proto-Germanic nominal stress became fixed on the initial syllable, there has always been erosion of nominal endings (Barber 2000: 92). This has led to the loss of many transparent formal gender distinctions on nouns. Consequently, lexical gender is now largely arbitrary in the Germanic languages, with respect to form as well as semantics.¹⁷

noteworthy that the gender systems described by Corbett (1991: 8-12) as strictly semantic all, in fact, show at least one of these characteristics. The semantic gender systems are all based on a distinction between either: (i) male/female humans versus all nonhumans (Tamil, Defaka, English), (ii) female animates versus all others (Diyari), or (iii) female animates and diminutives versus all others (Dizi, Halkomelem). It is possible that these semantic bases for gender classification are indeed less susceptible to disruption, as meaning changes from 'male' to 'female', 'human' to 'nonhuman', 'animate' to 'inanimate', or vice versa, are not very common. Furthermore, in English and Defaka, gender is only marked on pronouns, and in Tamil, it is only marked on the predicate. Thus, one can say that in English, and perhaps also in Tamil and Defaka, there is, in fact, no lexical gender – or at least, the noun's lexical gender is variable – so that the gender system is not subject to disruption in the lexicon.

¹⁷ The Germanic *-ing* suffix provides an interesting example of a nominal ending that used to show both a transparent formal as well as semantic gender distinction, which has been lost in the Germanic languages. The suffix dates back to Old Germanic, possibly Proto-Indo-European (Munske 1964). Munske (1964: 21) notes that originally, *-ing* had two distinct gender forms: masculine [†]*-inga* and feminine [†]*-ingo*. The masculine form derived concrete nouns, denoting persons or objects, from adjectives, nouns, and verbs. The feminine form derived abstract nouns from verbs. The two gender endings have disappeared. Now, the remaining suffix *-ing* conveys both types of meaning that are quite distinct in terms of individuation. The process of deriving abstract nouns with this suffix is still highly productive in Dutch, and although it no longer derives new concrete nouns, many of the derived concretes still exist in the lexicon, for example, *woesteling* 'ruffian', *koning* 'king', *drenkeling* 'drowned/drowning person'.

7. Semantic versus lexical gender agreement

Since the semantic basis of nominal gender assignment became disrupted, the lexical gender of nouns no longer always reflects the properties of their referents. This situation causes a conflict between lexical and semantic gender agreement in pronouns. Such a conflict of course only exists if the genders are, in fact, associated with particular referent properties. It seems that the Dutch genders have retained their semantic interpretation, that is, masculine/common gender being associated with high individuation and neuter gender with low individuation, despite the fact that these meanings have become opaque in the lexical domain. What could have played a role in the retention of the semantic interpretation of the genders is that lexical gender is not always involved in pronominal reference. This is the case in deictic reference, that is, in the absence of a linguistic antecedent, or when the antecedent is not expressed nominally, as in reference to nonentities. When no nominal antecedent is present, lexical gender does not have to interfere and pronouns can agree semantically.

In the presence of a nominal antecedent, however, lexical and semantic gender are in competition for agreement if the gender of the noun is not in accordance with the properties of the referent. Which of these two genders wins must somehow depend on the amount of pull that semantic and lexical gender each exert on the pronoun. Several factors appear to play a role here: the place of the referent on the Individuation Hierarchy, the distance between the pronoun and the antecedent, and, more generally, the amount of lexical gender marking in the nominal domain.

The pull from semantic gender appears to become greater as the degree of individuation of the referent is more saliently high or low. As discussed in Section 3, semantic agreement is less common in German than in Dutch, but it can occur with referents on the extreme ends of the Individuation Hierarchy. Audring (2009: 167-168) also observes for Dutch that semantic agreement is more likely to occur with animates and unspecific masses than with referents that take a position in between these two extremes.

The pull from lexical gender, on the other hand, seems to depend, first of all, on its degree of overt expression in the language in general. As discussed in Section 3, lexical gender agreement takes precedence over semantic gender agreement in languages that have relatively many adnominal elements marking lexical gender.

CHAPTER II

Second, the pull from lexical gender appears to depend on how far the pronoun is separated from its nominal antecedent. As noted earlier, the likelihood of semantic agreement increases the further a pronoun is separated from its antecedent noun (Audring 2009: 165-166). Apparently, the longer ago a noun was expressed, the less weight its lexical gender has in the competition for gender agreement. It could be said that it is generally the saliency of lexical gender, in the agreement system in general and in the actual context of the agreeing pronoun, that determines the pull from lexical gender in the competition between lexical and semantic gender agreement.

8. Conclusion

This study showed that the semantic agreement behaviour that has been observed in Dutch pronouns reveals a semantic principle that operates within the Dutch gender system as a whole. The gender distinction that is made between referents based on their degree of individuation was shown to relate to an existing semantic interpretation of the Dutch genders that arguably goes back to Proto-Indo-European. These semantics come to the fore most clearly in cases nominal gender is variable, such as with double gender nouns, nominalized adjectives or in pronominal reference to non-nominal antecedents. It was suggested that ever since nominal gender became an invariable, lexically stored feature of nouns, the semantic basis of nominal gender assignment has been susceptible to disruption over time. This analysis explains why lexical nominal gender is now largely arbitrary. This situation creates a conflict between lexical and semantic gender, which currently surfaces in Dutch and several other Germanic varieties where pronouns show variation between lexical and semantic gender agreement. Considering the distribution of semantic gender agreement in the Germanic languages, it was suggested that there is a relation between semantic gender agreement in pronouns on the one hand, and the number of adnominal elements that mark lexical gender, on the other hand. Semantic agreement in pronouns may only gain ground if the marking of lexical gender within the noun phrase has become sufficiently reduced.