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):

Chapter V

Semantic or lexical gender agreement: the effect of adnominal gender marking on pronominal agreement[*]

Abstract

This study investigates the effect of adnominal lexical gender marking on gender agreement variation in pronouns. Pronouns can show two kinds of gender agreement: lexical gender agreement, that is, agreement with the lexical gender of the antecedent noun, or semantic gender agreement, that is, agreement with the properties of the referent. The two kinds of gender agreement occur side by side in Dutch, with an increasing preference for semantic gender agreement. Dutch differs from earlier stages of the language in this respect and it differs also from German, which shows similar agreement variation, but with a strong preference for lexical gender agreement. These differences may be explained by the reduced visibility of lexical gender on adnominal elements in Dutch compared to German and earlier stages of Dutch. The experiment presented in this chapter investigates whether the variation between the two kinds of gender agreement in pronouns is influenced by lexical gender marking on the antecedent. The results show that the absence of lexical gender marking on the antecedent increases the likelihood of semantic agreement in pronouns. This finding reveals that the direct visibility of the noun's lexical gender affects the individual speaker's choice between semantic and lexical gender agreement in the pronoun. This effect can explain the synchronic variation between two the kinds of gender agreement, as well as a shift in the ratio between semantic and lexical gender agreement over time.

1. Introduction

This study is about variation between two types of pronominal gender agreement: agreement based on the lexically stored gender of the antecedent noun, henceforth called 'lexical gender agreement', or agreement based on certain properties of the referent, henceforth called 'semantic gender agreement'. In English, where nouns do not have lexically stored genders, pronominal agreement is exclusively semantic:

---

masculine and feminine pronouns are used with male and female referents respectively and neuter pronouns are used with inanimate referents and sometimes animals. In other Germanic languages, such as German, on the other hand, each noun has a lexically stored gender and pronominal agreement is generally based on this gender of the antecedent noun.

The history of English shows that gender systems can change from a predominantly lexical gender agreement system to a semantic agreement system. Old English had a lexical gender system like German, in which nouns were either masculine, feminine or neuter, and pronouns agreed with these nominal genders. However, as gender marking in the noun phrase was subject to deflection and was ultimately completely lost during the Middle English period, the lexical gender of nouns was no longer apparent to speakers and pronominal gender agreement became exclusively semantic (cf. Curzan 2003).

However, languages do not always show either lexical or semantic agreement in the pronoun. Also in languages with lexical nominal genders, semantic gender agreement exists. Semantic agreement already existed besides lexical gender agreement in Old English (Curzan 2003: 62) and this variation is also found in languages with lexical gender today. Semantic agreement can surface when the gender of the noun conflicts with the semantics of the referent. Such a conflict is possible in gender systems in which nominal gender is not systematically based on a semantic principle, which is the case in the Germanic languages. A conflict exists for instance with the German neuter noun *Mädchen* 'girl', for which pronominal agreement varies between neuter, in accordance with its lexical gender, and feminine, in accordance with the sex of the referent, as shown in (1) below (Corbett 1991: 228).

(1) Schau dir dieses Mädchen an, wie gut sie/es   Tennis
     look   you DEM.N girl(N)   at   how well 3SG.F/3SG.N tennis

     spielt.
     plays

     'Look at this girl, how well she plays tennis.'

Besides the sex of the referent, another semantic agreement principle plays a role in the Germanic languages: the degree of individuation of the referent (Siemund 2002, 2008, Audring 2009). A distinction is made between referents with a high degree of individuation, things that are bounded in nature, such as concrete objects, and referents with a low degree of individuation, things that are unbounded, such as materials and liquids. Referents with a high degree of individuation tend to receive masculine or common gender pronouns, while referents with a low degree of individuation tend to receive neuter pronouns. This semantic gender distinction has been found in Dutch (Van Haeringen 1936, 1951, Fletcher 1987, Siemund 2002, Audring 2006, 2009), Flemish (De Vos & De Vogelaer 2011, De Vogelaer & De Sutter 2011, De Vos 2013, 2014), Frisian (Wahrig-Burfeind 1989), West Jutland Danish (Ringgaard 1973, Braunmüller 2000), and in English dialects, such as West Somerset English (Siemund 2002, 2008).

Examples of semantic agreement based on individuation from spoken Dutch are shown in (2) and (3) below (from the *Corpus Gesproken Nederlands* 'Corpus of Spoken Dutch', from Audring 2006: 95-99). Standard Dutch spoken in the Netherlands (subsequently referred to as 'Dutch') has two nominal genders, common and neuter, where common gender is a conflation of former masculine and feminine gender. In example (2), a masculine pronoun is used in reference to a neuter noun denoting a concrete object, while in example (3), a neuter pronoun is used with a common gender noun denoting a mass.

(2) 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.'

(3) '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.'

The semantic agreement pattern can be described on the basis of the Individuation Hierarchy (Siemund 2002, Audring 2009). The hierarchy in (4) below, adapted from Audring (2009), shows the relevant semantic categories and the pronominal agreement pattern in Dutch.

(4)   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)

  fem./masc.   *masculine*   *neuter*
  common   *common*

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. These referents tend to receive masculine and common gender agreement, and feminine agreement if they are animate referents that are female. Entities with a low degree of individuation have unclear boundaries, are uncountable and have less specific characteristics. These referents tend to receive neuter agreement in the pronoun.

The pronominal agreement pattern in standard English can be related to this hierarchy as well (Siemund 2002: 225). The only difference between English and Dutch is that the cut-off point between non-neuter and neuter agreement is situated further to the left on the hierarchy in English: masculine and feminine pronouns are

used with animate referents on the far left end of the hierarchy, while neuter pronouns are used with all other referents.

An important difference with English, however, is that in Dutch semantic agreement exists beside lexical gender agreement. Agreement targets in the Dutch noun phrase, determiners and attributive adjectives, show agreement with the lexical gender of the noun, as in example (2) above, in which the neuter demonstrative determiner *dat* agrees with the noun's neuter gender. A common gender noun would take the determiner *die*. This difference between agreement targets is cross-linguistically common. It is captured in Corbett's (1979) Agreement Hierarchy, which shows that the likelihood of semantic gender agreement is the highest in personal pronouns and the lowest in adnominal elements. While adnominal elements always agree with the lexical gender of the noun in Dutch, pronouns show variation between lexical and semantic agreement. The noun *boek* in example (2) can alternatively receive neuter agreement in the pronoun, and the noun *olijfolie* in example (3) can alternatively receive masculine agreement, in accordance with their lexical genders.

Apparently, semantic gender is able to override lexical gender in Dutch pronouns. In the *Corpus Gesproken Nederlands* ('Corpus of Spoken Dutch'), Audring (2009:169-170) found semantic agreement in more than half of all pronominal references. Although semantic agreement is not a new phenomenon, as it already existed in Middle Dutch, its frequency seems to have increased over time (Kraaikamp to appear) and may be increasing still, as younger speakers today appear to show semantic agreement more often than older speakers (Audring 2009: 168-169).

A question that arises is what determines the variation between lexical and semantic gender agreement in Dutch pronouns. As lexical nominal gender is still consistently marked in the noun phrase and has not disappeared, as it has in English, there exists a conflict between two bases for agreement. Certain factors are known to play a role, such as the distance between the pronoun and the antecedent. Corbett (1979: 220) observes that the further a pronoun is separated from its antecedent, the higher the likelihood of semantic agreement becomes, an effect that was corroborated for Dutch by Audring (2009: 165).

Another factor that affects the likelihood of semantic agreement is the degree of conflict between the referent and the gender of noun: the more conflicting a noun's gender is, the more likely it is for semantic agreement to occur. Audring (2009: 167) found that referents that have an extremely high or low the degree of individuation, that is, referents towards the ends of the Individuation Hierarchy, are more likely to receive semantic agreement when the noun's gender conflicts with them than referents with a more moderate degree of individuation. For instance, human referents, such as *meisje* 'girl', at the left end of the hierarchy, and unspecific masses, such as *olijfolie* 'olive oil', at the right end, receive semantic agreement in the large majority of pronominal references when the noun's gender conflicts with them, 94% and 88% respectively, while nouns denoting objects, such as *boek* 'book', which take a middle position on the hierarchy, receive semantic agreement in 52% of the cases.

In a corpus study of pronominal agreement in Flemish, De Vos (2013) found another factor to play a role in agreement variation: the discourse prominence of the antecedent/referent. De Vos found a higher ratio of semantic agreement with referents that have a lower referential status, that is, referents that are less 'given' in the discourse, as visible from kind of determiner used with the noun, such as a definite or indefinite article. Also, the likelihood of semantic agreement appears to higher when the antecedent is not the subject of the sentence, but an object or oblique noun phrase.

While the factors mentioned above all play a role in synchronic variation, they do not explain diachronic variation in the frequency of semantic agreement, as there is no reason to assume that the average distance of pronouns to their antecedents, the degree of individuation of referents or the discourse prominence of referents and antecedents have changed over time. Another factor that could play a role, in both synchronic and diachronic variation, is the marking of lexical gender on the antecedent, that is, whether or not the gender of the antecedent noun is expressed on an accompanying adjective or determiner. This factor could be relevant in diachronic as well as synchronic variation, as adnominal gender marking has been subject to changes over time.

The aim of this study is to investigate experimentally whether gender marking on the antecedent has an effect on agreement in the pronoun. This chapter is

organized as follows. The next section discusses the changes that took place in adnominal gender marking in Dutch and what has been proposed in the literature regarding the relation between these changes and pronominal agreement. This section concludes with the specific expectation for the present study. Section 3 describes the design of the experiment. The results of the experiment are presented in Section 4, followed by a discussion in Section 5. Section 6 concludes this chapter.

2. Diachronic change in Dutch nominal gender

The Dutch gender system has changed from a system with originally three nominal genders, masculine, feminine and neuter, to a system with two nominal genders, common and neuter. This change was the result of deflection in the noun phrase. While the definite article distinguished masculine, feminine and neuter gender in Middle Dutch (ca. 1200-1600), for example accusative masculine *dien*, feminine *die* and neuter *dat*, these forms have been reduced to two forms, common *de* and neuter *het*. This loss of form distinctions for masculine and feminine gender in the noun phrase has gradually led to the conflation of these two genders around the seventeenth century (Geerts 1966). Audring (2006, 2009) notes that this change from three to two nominal genders has created a mismatch between nominal and pronominal gender in Dutch, as there are still masculine, feminine and neuter personal pronouns. Audring argues that this has created uncertainty regarding agreement with common gender nouns, about whether they should receive masculine or feminine agreement in the pronoun. This uncertainty may have led speakers to resort to semantic instead of lexical gender agreement in the pronoun.

In line with this idea, De Vogelaer & De Sutter (2011) found a relation between lexical agreement and the preservation of the masculine-feminine distinction in Flemish dialects of Dutch that still mark masculine gender to a more or lesser extent. Their questionnaire study showed more lexical gender agreement in dialects that still mark masculine gender in the noun phrase than in dialects where the two genders have become almost indistinguishable. De Vos (2014: 202-203) observed this effect even more directly in her corpus study of Flemish, which showed that masculine nouns accompanied by a distinctively masculine determiner receive more correct masculine agreement in the pronoun than masculine nouns occurring without such a determiner. These findings support the idea that

uncertainty about the gender of former masculine and feminine nouns may cause semantic agreement to increase.

However, while the conflation of masculine and feminine gender may have instigated a rise in semantic agreement, it does not immediately explain the synchronic agreement variation observed in Dutch today. Uncertainty about the gender of common gender nouns does not seem to play a role anymore. The change from three to two genders was completed a long time ago and agreement with masculine pronouns seems to have become the default for these nouns, at least in the spoken language, where feminine pronouns are now used exclusively with female referents (cf. Audring 2009: 92). Also, uncertainty about agreement with common gender nouns does not immediately explain the agreement variation that is observed with neuter nouns.

Another factor that could play a role, which is related to the conflation of masculine and feminine nominal gender, is the general visibility of lexical gender in the noun phrase. Kraaikamp (2012: 207-8) suggests that there could be a relation between how often lexical gender is marked in the noun phrase and the frequency of semantic agreement in the pronoun. As the erosion of inflectional endings led to the conflation of masculine and feminine gender marking in the noun phrase, some adnominal elements lost their gender marking properties entirely. For example, the indefinite article and most possessive determiners have become completely invariant, so that they no longer mark gender at all. Table 1 below gives an overview of the most frequent adnominal elements and whether or not they mark gender in present-day Dutch compared to Middle Dutch.

Table 1. 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

Gender used to be marked on more adnominal elements than today (cf. grammars of Middle Dutch, e.g. Van Royen 1991, Mooijaart & Van der Wal 2008). The form of the indefinite article varied with the gender of the noun in Middle Dutch, for example accusative masculine *enen*, feminine *ene* and neuter *een*, while present-day Dutch uses the same form, *een,* with all nouns. The possessive determiners showed the same gender variation in Middle Dutch, for example accusative masculine *minen*, feminine *mine* and neuter *mijn* 'my', but possessive determiners have become invariant in present-day Dutch, except for the first person plural possessive, which distinguishes common *onze* versus neuter *ons* 'our'. The attributive adjective used to mark gender in both definite and indefinite noun phrases, but it is now invariantly inflected with *-e* in definite noun phrases. These changes in adnominal gender marking mean that nouns are now less frequently accompanied by an element that expresses their gender.

With respect to pronominal agreement with human referents, Geerts (1966: 134) already suggests that the presence of a gender-marking adnominal element favours lexical gender agreement in the pronoun. He gives an example of agreement with the neuter noun *ventje* 'boy, little guy' and observes that it is easier to refer to this noun with the masculine pronoun *hem* 'him' when the noun is combined with the indefinite article *een* (*een ventje*) than when it is combined with the definite article *het* (*het ventje*). Geerts' intuition is that it would be more common to use the neuter pronoun *het* in the latter case.

The experiment presented in this chapter investigates whether gender marking on the antecedent in fact influences the choice between lexical and semantic agreement in the pronoun. It is expected that if a noun is accompanied by a determiner that marks its gender, speakers are more likely to agree with the lexical gender of the noun than when the noun is not accompanied by such a determiner.

3. Method

3.1 Design

The experiment consists of a sentence completion task that elicits pronominal references to a neuter noun denoting an object, which is either combined with a determiner that marks its gender or with a determiner that does not mark gender. The test focuses on object referents, because the variation between lexical and

semantic agreement exists most clearly with this category: neuter nouns denoting objects receive semantic agreement in approximately half of the cases in Dutch (Audring 2009, cf. Section 1).

The design of the experiment was such that the effect of inter-speaker variation is minimized. Speakers can differ in their agreement preferences, in the sense that some speakers may be generally more inclined towards semantic agreement than others. Such inter-speaker variation could obscure an effect of the variable under investigation. Therefore, a repeated measures design was used: the same participant was tested on both conditions of the relevant test variable, the gender-marked antecedent and the unmarked antecedent (see Section 3.2 below). This way, if the two conditions yield different results, this difference cannot be attributed to inter-speaker variation. In addition to this, it was decided to test many subjects with only few test items, so that it is not the case that relatively large portions of the data come from single participants.

3.2. Test items

The nouns *boek* 'book' and *bord* 'plate' were used in the experiment, combined with either the possessive determiner *m'n* 'my', which does not mark gender, as the same form is used with common and neuter nouns, or the proximal demonstrative determiner *dit* 'this', which marks neuter gender, as this form is used only with neuter nouns, common nouns taking the form *deze* 'this'.

These two determiners were chosen, and not, for instance, the gender-neutral indefinite article *een* 'a' and the neuter definite article *het* 'the', because the demonstrative and possessive determiner have a similar effect on the specificity of the referent. This is relevant because specificity is a feature that affects the degree of individuation of the referent: a more specific referent has a higher degree of individuation than a less specific referent. Since the likelihood of semantic agreement varies with the degree of individuation of the referent (Audring 2009, cf. Section 1), it is crucial that the referents in this experiment do not differ in this respect.

In terms of discourse prominence, another factor involved in agreement variation (De Vos 2013, cf. Section 1), the possessive and demonstrative noun phrase used in the experiment can be considered to have similar discourse

prominence, as both noun phrases introduce a new referent and function as a direct object in the test sentence. Should any difference between the discourse status of the possessive and demonstrative noun phrase still play a role in the experiment, it is at least not likely to bias the results towards confirmation of the hypothesis. De Vos (2013) found a slightly higher ratio of lexical gender agreement with possessive noun phrases than with demonstrative noun phrases in her corpus. The expectation for this experiment is exactly the other way around: More lexical agreement is expected with the demonstrative noun phrase than with the possessive noun phrase, as the demonstrative determiner marks gender and the possessive determiner *m'n*, used in this experiment, does not.

3.3. Participants

A total of 117 native speakers of Dutch participated in this study, 59 men and 58 women. Their ages ranged between 18 and 31, with an average age of 22. The participants were all students at the University of Amsterdam, mostly science majors. Students of linguistics were not included.

3.4. Test

The test was designed in such a way that participants produce anaphoric pronouns while being unaware that these are the focus of the test. The design was inspired by the pronominal agreement test used by Braun & Haig (2010). The test consists of five incomplete sentences, presented in written form, which have to be completed using content words provided between brackets. The participants are told that the test is about forming sentences in spontaneous speech. They are asked to read each sentence out loud and complete it using the provided content words. The participants' responses are audio recorded by the experimenter. The test is short and the responses are provided orally, in order to minimize the chance of participants becoming aware of the focus of the test and giving overly conscious responses. The test contains two test sentences and three filler sentences. An example of a filler sentence is shown in (5) below:

(5)   *Die gegevens zoek ik meteen wel op, want ik ... [zit, nu, toevallig, computer]*
'I will look up that information directly, because I … [sit, now, coincidentally, computer]

The typical response would be as in (5'):

(5')   *Die gegevens zoek ik meteen wel op, want ik zit nu toevallig achter de computer.*
'I will look up that information directly, because I happen to be behind the computer now.'

Each sentence consists of two clauses of which the first clause and the first two words of the second clause are provided. In the test sentences, the first clause contains the test item, which serves as an antecedent to an object pronoun in the following clause. Examples of the test sentences are shown (6) and (7) below:

(6)   *Ik heb m'n boek nog niet uit, maar ik …. [moet, vandaag, terugbrengen, bibliotheek]*
'I have not finished my book yet, but I … [have.to, today, return, library]

(7)   *Ik heb dit bord net al gebruikt, dus ik …. [ga, eerst, omspoelen, keuken]*
'I already used this plate just now, so I … [go, first, rinse, kitchen]

Typical responses would be as in (6') and (7') (pronouns in bold):

(6')   *Ik heb m'n boek nog niet uit, maar ik moet **het/hem** vandaag terugbrengen naar de bibliotheek.*
'I have not finished my book yet, but I have to return it to the library today'.

(7')   *Ik heb dit bord net al gebruikt, dus ik ga **het/hem** eerst omspoelen in de keuken.*
'I already used this plate just now, so I will rinse it in the kitchen first'.

The test sentences are constructed in such a way that a personal pronoun is most likely to be used. The position where the pronoun should be used is always fourth in the clause, directly after the auxiliary verb. The distance between the test item and the position of the pronoun is always six words. The antecedent varies in whether or not it is marked for gender. In one test sentence the antecedent is not marked, as in (6), while in the other test sentence it is marked, as in (7).

Two versions of the test are used in order to control for item and order effects. In test version 1, *bord* is marked for gender (*dit bord*), while *boek* is unmarked (*m'n boek*), as in example (6) and (7) above. This is the other way around in test version 2: *boek* is marked for gender (*dit boek*), while *bord* is unmarked (*m'n bord*). The order in which the marked and unmarked item are presented varies in the different test versions. The gender-marked antecedent comes first in test version 1, while it comes last in the other version. The test versions are distributed evenly over the participants.

The two test sentences are separated by two filler sentences and the test starts with a filler sentence for practice. The fillers do not contain any pronouns and are not completed with pronouns. In order to distract the participants from the actual focus of the test, all sentences, both the filler sentences and test sentences, are completed with a clause ending in a directional phrase. The preposition that should be used in this phrase differs per sentence and is not provided between brackets. This was to make participants think that prepositions are the focus of the test.

4. Results

A total of 212 responses from 106 participants were obtained and included in the analysis. The responses from eleven participants had to be excluded, because the participant either repeated the noun phrase instead of using a pronoun (this occurred six times), misread the test item (once) or produced an ungrammatical word order in which the pronoun did not follow directly after the auxiliary (four times). The latter happened when a participant simply read the first two words between brackets out loud without first forming the sentence in their mind. In case of such an erroneous response, both responses from the participant were excluded from the dataset. The balanced ratio of male and female participants and the ratio of test version 1 and 2 is well maintained in the final dataset. The final dataset contains responses from 52

men and 54 women. Test version 1 was used 54 times and test version 2 was used 52 times.

Personal pronouns were used in the large majority of cases, as expected. Eight participants used a demonstrative pronoun instead of a personal pronoun in one of the test sentences. This occurred with both test sentences and involved both the neuter and common demonstrative pronoun. In the analyses that follow, the responses with a neuter demonstrative are subsumed under the responses with neuter personal pronouns and those with a common demonstrative are subsumed under the responses with masculine personal pronouns.

The participants used both neuter and masculine pronouns to refer to the neuter nouns *boek* and *bord*. This confirms that there is variation with these object nouns between lexical gender agreement with neuter and semantic gender agreement with masculine pronouns. In total, there were 120 responses with a neuter pronoun (57%) and 92 responses with a masculine pronoun (43%).

The items *boek* and *bord* received neuter and masculine pronouns in similar ratios. Figure 1 below shows the percentages of neuter and masculine pronouns used with *boek* and with *bord*.
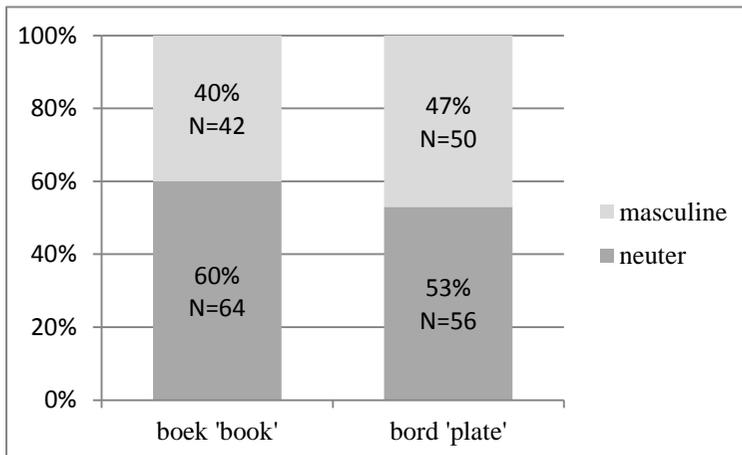

Figure 1. Percentages of neuter and masculine pronouns used
with *boek* 'book' and with *bord* 'plate'

The item *boek* receives slightly more references with neuter pronouns than *bord*, but this difference between the two nouns is not statistically significant (χ^2 (1) = 1.229, p=0.268).

Figure 2 below shows the results for the variable under investigation, the percentage of neuter and masculine pronouns used with the antecedent that is marked for neuter gender (*dit* NOUN) and with the antecedent that has no gender marking (*m'n* NOUN).
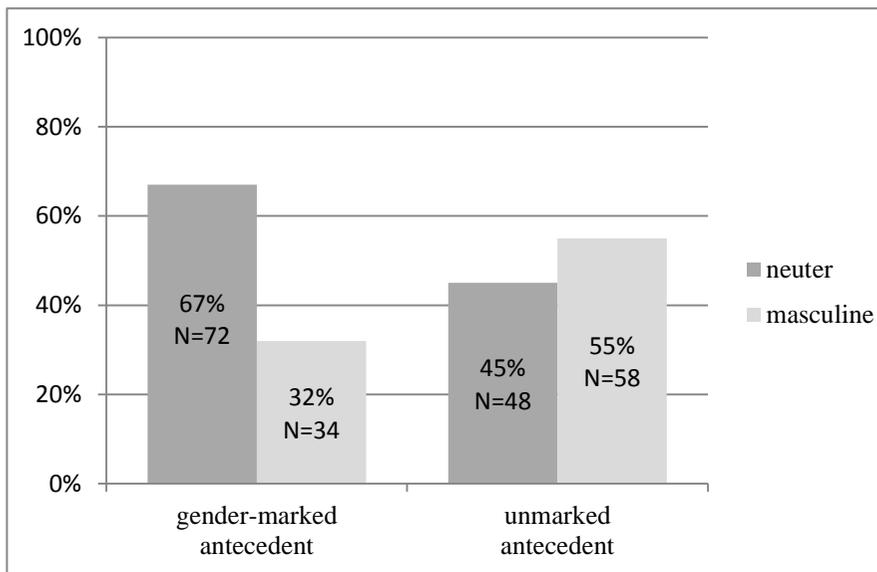

Figure 2. Percentage of neuter and masculine pronouns used with the gender-marked antecedent (*dit* NOUN) and the unmarked antecedent (*m'n* NOUN)

The responses with the gender-marked antecedent and the unmarked antecedent are different. The marked antecedent receives more references with neuter pronouns, that is, lexical gender agreement, than the unmarked antecedent. The marked antecedent receives neuter agreement in 67% of the cases, while the unmarked antecedent receives neuter agreement in 45% of the cases.

Because the experiment has a repeated measures design, in which the same participant is tested on both the gender-marked antecedent and the unmarked antecedent, the results are analysed using a McNemar test for repeated measures. This test takes the response pairs per participant as the outcome variable, that is, the

combination of the participant's response to the marked and unmarked antecedent. These response pairs are shown in Table 2 below.

	unmarked antecedent		
		neuter	masculine	*total*
gender-marked antecedent	neuter	45	27	*72*
	masculine	3	31	*34*
	total	*48*	*58*	*106*

Table 2. Frequencies of the response pairs. Contingency table for the McNemar test.

Many participants responded with the same pronoun to both the marked and unmarked antecedent (upper left cell and bottom right cell in Table 2). 45 out of a 106 participants (42%) consistently used neuter pronouns and 31 participants (29%) consistently used masculine pronouns. A total of 30 participants (28%) showed variation in their pronoun choice (upper right cell and bottom left cell). In these cases of variation, the large majority, 27 participants (90%), used a neuter pronoun with the marked antecedent and a masculine pronoun with the unmarked antecedent. Only three participants (10%) showed the reverse pattern. The McNemar test confirms that if speakers vary in their pronoun choice, they are significantly more likely to use a neuter pronoun with the gender-marked antecedent and a masculine pronoun with the unmarked antecedent than the other way around (Exact sign. $p<0.001$).

5. Discussion

The results presented in this chapter first of all confirm that there is gender agreement variation in Dutch pronouns. The results are in line with the earlier observation that there is a tendency to use masculine pronouns with nouns denoting concrete objects, regardless of the noun's lexical gender.

The results of the experiment reveal that pronominal agreement is affected by gender marking on the antecedent. If the antecedent noun is accompanied by a

determiner that shows the noun's gender, pronouns are more likely to agree with this gender than when there is no such gender marking on the antecedent. This effect does not have to do with a loss of knowledge or uncertainty about the gender of the noun, as may be the case with common gender nouns in Dutch, as the nouns used in the experiment, *boek* and *bord*, are neuter nouns for which no uncertainty exists. They are frequent nouns and they always receive neuter agreement in the noun phrase. The effect is therefore more likely to be what may be called a 'reminder' effect for lexical gender. The effect may work in a similar way as the distance effect. Both relate to the salience of the noun's gender to the speaker at the moment he/she chooses a pronoun. The longer ago the antecedent noun was expressed or the less explicit its lexical gender is marked, the less salient the noun's gender becomes, making it easier for semantic gender to overrule it.

It should be noted that, as all other effects on pronominal agreement variation, the observed effect of gender marking is probabilistic, not deterministic: a gender-marked noun does not always receive lexical agreement in the pronoun and an unmarked noun does not always receive semantic agreement in the pronoun. Many participants did not show variation in their pronoun choice and either agreed lexically with the both the gender-marked and unmarked antecedent or they agreed semantically with both antecedents. This confirms that there is inter-speaker variation in pronominal agreement: some speakers are more inclined towards semantic agreement than others. This variation is expected considering that there are also differences between generations of speakers, with younger speakers generally showing more semantic agreement than older speakers (Audring 2009).

The effect of adnominal marking was found with object referents, but it is likely that it exists with other referents as well, although this is open to future research. As mentioned in Section 1, the likelihood of semantic agreement is different for different referents on the Individuation Hierarchy. Object referents receive semantic agreement in approximately half the cases, but unspecific masses, such as *limonade* 'lemonade', receive a much higher ratio of semantic agreement (Audring 2009). This is explained by the extremely low position of unspecific masses on the Individuation Hierarchy. In light of the present findings an additional factor may be considered to play a role here: unspecific masses are referred to with bare nouns, possibly combined with an adjective, but never with a determiner. A

mass noun combined with a determiner, e.g. *de limonade* 'the lemonade', refers to a specific mass, which has a higher degree of individuation (cf. the Individuation Hierarchy presented in Section 2). Consequently, nouns referring to unspecific masses are infrequently marked for gender. They only carry gender marking when combined with an adjective, e.g. *koude limonade* 'cold lemonade'. This factor may contribute to the high ratio of semantic agreement with unspecific masses.

The effect of adnominal marking on pronominal agreement that is now observed in synchronic agreement variation may also play a role in diachronic change in pronominal agreement. If a speaker's choice between lexical and semantic agreement in the pronoun is influenced by gender marking on the antecedent, then an overall decrease of gender marking in the noun phrase is likely to cause a shift in agreement preference. As discussed in Section 2, Dutch has lost gender markers in the noun phrase, as some determiners that used to show a gender distinction have become invariant. The determiner *mijn* 'mine', used in the experiment as a gender-neutral determiner, is an example of this. The present findings show that this loss of adnominal gender marking could have played a role in the increase of semantic agreement in Dutch.

Similarly, the effect of adnominal marking may explain differences in pronominal agreement between related languages, such as German and Dutch. Both varieties show semantic agreement of pronouns within a gender system that has lexical nominal gender. However, semantic agreement is much less frequent in German than in Dutch (Kraaikamp 2016). This difference between the two languages may be explained by differences in gender marking. Lexical gender is more visible in German compared to Dutch. German has preserved more gender marking on determiners and is quite similar to Middle Dutch (see Section 2) in this respect.

Related to the preservation of gender markers, German, as Middle Dutch, also still distinguishes three nominal genders. This means that a mismatch between nominal and pronominal gender and uncertainty about agreement with former masculine and feminine nouns does not play a role in German as it may in Dutch, in the way proposed by Audring (2009). It is possible that the mismatch situation and the loss of gender markers both contributed to the increase of semantic agreement in Dutch. However, it should be noted that the loss of the three-gender system has

followed from the loss of adnominal gender markers. Therefore, it is possible that the loss of gender markers alone is responsible for the high frequency of semantic agreement in Dutch. Also, because the present experiment focused on semantic agreement with neuter nouns, for which no uncertainty regarding their gender exists, the effect of adnominal gender marking is clearly not dependent on uncertainty about nominal gender, but is an effect on its own.

A question that arises with the present findings is whether Dutch could be moving towards a system where pronominal agreement is exclusively semantic, as in present-day English. However, it seems likely that as long as lexical gender is marked in the noun phrase there will be a competition, and consequently variation, between lexical and semantic agreement in the pronoun. There are two plausible developments that could change this situation. One is that further deflection could eventually lead to the complete loss of lexical gender from the noun phrase, as happened in the history of English (Curzan 2003). However, there are no signs of such a radical deflection process taking place in present-day Dutch. Another possible development is that while gender distinctions are preserved on adnominal elements, semantic agreement moves up the Agreement Hierarchy (cf. Section 2) and starts involving agreement within the noun phrase. A combination of this process and deflection appears to have taken place in in the Danish dialect of West Jutland. Lexical gender distinctions have completely disappeared from the noun phrase in this dialect. The demonstrative pronouns, common *den* and neuter *det*, shows semantic gender agreement based on individuation. The demonstrative pronouns are also used as demonstrative determiners, and within the noun phrase, they show the same semantic agreement as in the pronominal domain (Ringgaard 1973, Braunmüller 2000).

In Dutch, gender agreement in the noun phrase is based on the lexically stored gender of the noun and is mostly invariant. However, there is some limited agreement in the noun phrase that could be called semantic agreement. It is found with a particular set of nouns that have variable gender, for example common *de steen* 'the stone', denoting the object, versus neuter *het steen* 'the stone', denoting the material, a mass (see Kraaikamp 2012, Semplicini 2012). However, there are no indications that nominal gender variation is becoming more widespread. It seems rather that the ties between nouns and their lexical genders are quite strong, despite

the abundant and long-existing agreement variation in the pronoun. Notably, the pronominal agreement variation does not seem to affect the acquisition of nominal gender by children. Children acquiring Dutch exclusively overgeneralize common gender, which is the most frequent and default nominal gender, while neuter gender is gradually acquired for each neuter noun (Blom, Poliśenská & Weerman 2008). Interestingly, neuter gender does seem to be acquired faster for neuter nouns denoting masses than for neuter nouns denoting objects, but still, children never overgeneralize neuter gender with common gender nouns (Roodenburg & Hulk 2009). An increase of semantic agreement in the noun phrase may originate with adults instead, but it seems that further deflection of lexical gender markers, as in West Jutish, is a prerequisite for such a process to occur.

6. Conclusion

The findings presented in this chapter demonstrate that the variation between semantic and lexical gender agreement in Dutch pronouns is affected by the presence or absence of gender marking on the antecedent. Semantic agreement is more likely when the antecedent carries no gender marking than when the noun's gender is explicitly marked on the accompanying determiner. This result shows that gender marking in the noun phrase plays a role in the synchronic variation between lexical and semantic gender agreement in pronouns. The effect likely also plays a role in diachronic variation between the two kinds of gender agreement. The effect can explain the observed increase of semantic agreement in Dutch over time, since deflection has caused many determiners to lose their gender marking properties in Dutch, leading to a reduced visibility of lexical gender in the Dutch noun phrase.