



Semantic Versus Lexical Gender

M. Kraaikamp

Summary

Semantic versus lexical gender: synchronic and diachronic variation in Germanic gender agreement

Most Germanic languages, including Dutch, have a gender system in which each noun belongs to a certain gender. These nominal genders are reflected by gender agreement on associated words, such as determiners, adjectives and pronouns. However, pronouns do not always show gender agreement with their antecedent noun. A familiar example of this is when a masculine or feminine pronoun is used with a neuter noun that refers to a human being. For example, in Dutch the masculine pronoun *hij/hem* ‘he/him’ can be used with the neuter noun *jongetje* ‘boy’ or the feminine pronoun *zij/haar* ‘she/her’ with the neuter noun *meisje* ‘girl’. In these cases the pronoun shows what is called ‘semantic gender agreement’, agreement that is based on the properties of the referent rather than the lexical gender of the noun. Another, perhaps less familiar, kind of semantic gender agreement that occurs in Dutch involves inanimate referents. Examples of this are shown in (1) and (2) below, where in (1) a masculine pronoun (*hem*) is used with a neuter noun, *boek* ‘book’, and in (2) a neuter pronoun (*het*) is used with a common gender noun, *gember* ‘ginger’:

(1) *Ik heb m'n boek nog niet uit, maar ik moet hem terugbrengen naar de bibliotheek.*

‘I have not finished my book yet, but I have to return it to the library.’

(2) *Je kunt verse gember bijna overal krijgen, maar ik vind het vaak te duur.*

‘You can get fresh ginger almost everywhere, but I think it is often too expensive.’

The agreement pattern that these examples illustrate involves 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, referents that have a clearly bounded shape, such as the object *boek* ‘book’ in (1).

SUMMARY

Neuter gender pronouns, on the other hand, tend to be used with referents that have a low degree of individuation, that is, referents with unclear boundaries, such as the mass *gember* ‘ginger’ in (2).

The aim of this dissertation is to investigate the origin of this semantic agreement pattern in Dutch, when it has developed and what factors could be involved in its surfacing. This work consists of four separate studies that are intended to shed light on the roots of the phenomenon. The first study explores the semantic underpinnings of the Dutch gender system. The second study is an experimental study that compares the agreement behaviour in Dutch with that in German. The third study is a historical corpus study of pronominal gender agreement in Middle Dutch. The fourth and final study is an experimental study that investigates the effect of adnominal gender marking on pronominal agreement variation.

The first study shows that the association of common and neuter gender with a high and a low degree of individuation, respectively, not only exists in pronouns, but can be found in nominal gender as well. Although nominal gender assignment is largely semantically arbitrary in Dutch, a semantic pattern surfaces in certain cases. This occurs in particular when the gender of the noun is not fixed but variable. The association between neuter gender and a low degree of individuation is also visible when the antecedent is not a noun, but, for instance, a predicate. These observations indicate that the semantic interpretation of the genders is not an innovation in Dutch pronouns, but something that has long existed. There are in fact indications that the semantic interpretation of the genders goes back to their roots in Proto-Indo-European. In line with this, semantic agreement based on individuation has been found not only in Dutch, but also in other Germanic varieties, including dialects of English, Frisian and Danish, and in Romance dialects. The Indo-European gender system possibly had a semantic basis, which has become disrupted in the nominal domain ever since nominal gender became an invariable, lexically stored feature of nouns. The lexical gender of nouns is consequently no longer always in line with their meanings and as a result of this, a conflict can arise between the lexical gender of nouns and the semantic gender associated with their referents.

The second, experimental, study is aimed at investigating the possible relation between pronominal agreement based on individuation and a change that took place in the Dutch nominal gender system: the change from three to two nominal genders

as a result of the conflation of masculine and feminine gender into one common gender. Agreement based on individuation has so far only been attested in Germanic varieties in which masculine and feminine gender have conflated, as in Dutch, or that lost lexical gender altogether. The second study investigates whether agreement based on individuation exist also in German, a Germanic variety that still distinguishes the original three nominal genders, and how pronominal agreement in German compares to that in Dutch. This study involves pronoun elicitation experiments in Dutch and in German, in which native speakers of each language are prompted to refer to particular referents that differ in their degree of individuation. The results show that both speakers of German and speakers of Dutch are inclined to show semantic neuter agreement with masses and abstract referents, and semantic masculine agreement with animal referents. However, semantic masculine agreement with object referents is shown only by speakers of Dutch and also the frequency of semantic agreement is overall much higher in Dutch than in German. The findings of this study suggest that semantic agreement based on individuation is a shared Germanic feature and that the difference between Dutch and German with respect to this kind of agreement is mainly one of degree.

The third, corpus, study investigates the diachrony of agreement based on individuation in Dutch and focuses on the question whether or not this kind of agreement already existed in Dutch before the distinction between masculine and feminine nominal gender was lost. This study is a corpus study of pronominal agreement in Middle Dutch recipe books from the early 16th century. The results show that semantic agreement based on individuation existed already in Middle Dutch, in particular, semantic neuter agreement with mass referents. Semantic agreement with object referents is not attested in the corpus. It appears that in Middle Dutch, semantic agreement occurs only with referents that have either a very high degree of individuation, that is, animate referents, or that have an extremely low degree of individuation, that is, masses. A comparison of the frequency of semantic agreement in Middle Dutch to that in present-day Dutch suggests that semantic agreement has generally increased over time.

The final, experimental, study investigates a factor that could play a role in the diachronic increase of semantic gender agreement, as well as the synchronic variation in its frequency between languages such as Dutch and German: the

SUMMARY

visibility of lexical gender in the noun phrase. Dutch has changed over time in this respect and now differs from German, as deflection has led to a loss of lexical gender marking on several elements inside the noun phrase, such as the indefinite article and possessive determiners. The fourth, final, study investigates whether there is a relation between lexical gender marking in the noun phrase and the choice between semantic and lexical gender agreement in pronouns. This study involves a pronoun elicitation experiment with speakers of Dutch in which the test sentences vary in one crucial aspect: the antecedent noun is accompanied by either a determiner carrying explicit gender marking or a determiner without gender marking. The results show that the likelihood of semantic agreement is higher when there is no gender marking on the antecedent. This finding reveals that the absence of lexical gender marking in the noun phrase makes semantic agreement surface more easily in pronouns. This effect can explain both the synchronic variation between semantic and lexical gender agreement and the shift in the frequency of semantic agreement in Dutch over time.

The results of the four studies that are presented in this dissertation indicate that the semantic gender agreement that is observed in Dutch pronouns today relates to an existing semantic interpretation of the genders, which possibly goes back to the semantic roots of the gender system. Pronominal agreement based on individuation existed already in Middle Dutch and it surfaces in varying degrees in other Germanic varieties, including German. The rigid nature of the lexical gender system, that is, the fact that the lexical gender of nouns is largely fixed and invariable, creates a conflict between the lexical gender of nouns and the semantic gender that is associated with their referents. The extent to which semantic gender agreement can surface in pronouns, that is, the ratio of semantic to lexical gender agreement in pronouns, appears to depend on the visibility of lexical gender in the noun phrase. As this aspect of the lexical gender system varies both synchronically and diachronically, it can explain both the variation in the ratio of semantic to lexical gender agreement between languages and the increase of semantic gender agreement in Dutch over time.