The diachrony of gender and the interface of semantic and lexical gender in the Dutch language

Abigail Lambert[1]

University College Utrecht

The application of gender in Indo-European languages is notoriously arbitrary. This is particularly typical of the Dutch language. No single rule defines the way in which gender is determined or structured. Many factors interplay to create the Dutch grammar we have today, yet even these rules are arbitrary to a certain extent, with anomalies to the rules being common.

This paper aims to address the grammatical rules behind the structure of gender in the Dutch language and the way in which this has changed over the centuries. Other underlying grammatical aspects that may be indirectly related to this diachrony are discussed. The factors are assessed and combined in order to explain the phenomena of change. These changes are applied and compared to the grammar of other languages within the Indo-European language family, particularly to the reconstructed grammar of Proto-Indo-European.

1. Introduction and Background

Languages are in a constant process of developing and changing. A natural language is never in a state of complete stasis. Thus, language change is inevitable within the confines of social interaction and time, with the degree of change increasing and decreasing according to various social factors[2]. These changes can provide further insight and understanding as to how a language's grammar functions and changes, and how much the active use of it in day to day life affects the way it changes. The focal research question for this paper is: how far has gender representation changed between middle and modern Dutch and what factors are implicit in this change? And, furthermore, can these changes be seen within other languages belonging to the same language family?

Gender is a major grammatical component of language in the Dutch language, which makes the changes in the use and representation of gender in Dutch all the more prominent. The concept that gender changes are occurring is highly interesting and brings questions as to why these changes are occurring. It highlights the interesting process of language change and tells us that the grammar of a language is not indefinitely concrete.

Diachrony is the process of change occurring over time through history and the term being using to describe the gender changes in the Dutch language within the context of this paper. Gender is manifested in Dutch through the assumed genders of the nouns themselves, but also through their determiners or antecedents[3]. This paper will be focusing on the way in which the gender of referents affects these determiners/antecedents, and what grammatical rules apply to how the gender influences them and under what circumstances.

One of the key points regarding the diachrony of gender in the Dutch language that should be taken into consideration is that the changes do not necessarily develop nor are made concrete

---

[2] Nettle, D (1999) Is the Rate of Linguistic Change Constant? p.120
[3] De Vogelaer, G. Changing Pronominal Gender in Dutch: Transmission or Diffusion? p.71

over the same span of time[4]. More specifically, concerning Dutch, these specific gender changes - the shift from three nominal genders to two nominal genders - occurred at different rates. These changes developed much faster in spontaneous spoken Dutch than in written Dutch[5]. This is due to the Dutch government actively trying to resist this shift in the nominal genders through stressing the importance of using all three of the genders used in middle Dutch. This was done through including them in public documents and publishing and releasing dictionaries that further reinforced the perceived importance of retaining the three-gender system[6]. However, this was ineffective in preventing the inevitable changing process of the spoken language; yet managed to slow down the process in written language.

The research question will be answered through addressing the relationship between lexical gender agreement and semantic gender agreement. Semantic gender agreement is a form of agreement that is based upon particular properties of a referent. Two of the most prominent properties are animacy/sex (of the referent's biological state). Lexical gender agreement is a form of agreement based upon the inherent lexically stored gender of the noun[7]. Semantic gender agreement tends to take precedence over lexical gender agreement in situations where the lexical gender agreement of the noun is ambiguous or under dispute.

Each section in this paper will address a grammatical element involved and/or affected by the gender changes within the Dutch language.

1. Dutch Gender in the Context of the Germanic Family

There are quite stark contrasts in the ways that semantic gender agreement and lexical gender agreement manifest themselves in West Germanic Languages, despite the fact that they are all so closely related in linguistic terms. English completely lacks lexical gender agreement; all agreements are done semantically[8]. German, on the other hand, makes much more consistent use of lexical gender agreement and so it employs semantic gender agreement much less frequently. Dutch falls somewhere in the middle of these two extremes as each Dutch noun has a lexically stored gender but semantic gender agreement still occurs, particularly in situation where the referent is a human being[9]. In addition, it appears as though Dutch has undergone a change from middle Dutch resembling modern German, to modern Dutch resembling modern English[10].

Information regarding other languages from the same linguistic family as Dutch is useful in terms of understanding the different ways in which linguistically similar languages are changing over time. Grammatical gender changes within other languages from the same family as Dutch will be discussed later on.

---

[4] Nettle, D (1999) Is the Rate of Linguistic Change Constant? p.120
[5] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. p.29
[6] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. p.29
[7] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.81
[8] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.82
[9] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.82
[10] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. p.28

2. Data from the Dutch Language

2.1.   'De' and 'Het'

Articles within the Dutch language contribute to the determining of the gender of a noun. Understanding the role of the articles is important in comprehending grammatical gender in the Dutch language.

'De' and 'het' are the two definite articles in Dutch, both meaning 'the'. The two determiners distinguish nouns into two genders, the common gender and neuter gender[11].

The common gender is the name of the gender that is now considered as being a combination of the masculine and feminine gender following their conflation. However, the common gender is most often manifested as masculine in terms of its representative gender. The feminine gender specifically is only used in situations where the referent is explicitly feminine and its default lexical gender is not neuter.

The common gender is also always used for the plural version of nouns. For example:

'de hond' (the dog)   *'het hondje'* (the little dog)   *'de honden'* (the dogs)

It must also be noted that 'common gender nouns always take the pronoun whose gender corresponds to the natural, that is, biological gender of their referents'[12] (except in a few anomalous circumstances which will be outlined later on).

'Het' is the determiner used for the neuter gender class. It is also used alongside every noun in its diminutive form (as seen with *'het hondje'*)[13]. 'Het' nouns only account for approximately one quarter of all Dutch nouns, however, this statistic is under dispute[14]. Some examples of its use in the diminutive are as follows:

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'

This data shows that in every instance the meaning of the noun changes slightly in the diminutive form, by focusing the subject matter in terms of specificity.

There are other instances in which gender is presented as being particularly arbitrary. In Dutch, there are double gender nouns[15]. These are nouns that can be preceded by either of the two definite articles and thus can take on either the common gender or the neuter gender.

---

[11] Donaldson, B. (2008). Dutch: A Comprehensive Grammar. Oxon: Routledge Publishers. 2nd Edition p.27

[12] Audring, J. (2006). Pronominal Gender in Spoken Dutch. Journal of Germanic Linguistics. 18.2 (2006): 85-116 p.91

[13] Donaldson, B. (2008). Dutch: A Comprehensive Grammar. Oxon: Routledge Publishers. 2nd Edition p. 27

[14] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. pp.33-35

[15] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. p.25

de/het aanrecht 'sink'

de/het schort 'apron'

de/het omslag 'envelope'

de/het kaneel 'cinnamon'

These examples[16] show just how arbitrary gender assignment is in Dutch.

Double noun compounds are another variant in the way in which gender is structured in Dutch. All noun-noun compounds take on the gender of the second noun in the compound, which is the head of the compound.

De veldslag v Het slagveld

De landwijn v Het wijnland

De toneelwereld v Het wereldtoneel

We find that Dutch has no general rule for the assignment of common or neuter gender to nouns. Lexical gender agreement is very arbitrary in modern Dutch[17] as it appears as though there is no general rule specifically applicable to this grammatical element of Dutch.

The assignment of lexical gender to nouns in Dutch is extremely arbitrary. A key example[18] of this is with the Dutch noun *'meisje'* which means 'girl' in English.

Dat	meisje	dat	daar	staat	ken	ik.	Ze	zit	bij	mij	op	school.
DEM.N	Girl(N)	REL.N	there	stands	know	I	3SG.F	Sits	With	Me	In	School.

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

In this example, both the determiner and the relative pronoun to the noun are neuter. This agrees with the noun because the gender of the noun *'meisje'* is neuter, as opposed to feminine which many would assume as being the semantic gender. However, the personal pronoun remains feminine, agreeing with the biological sex of the referent. With such nouns - like *'meisje'* - there is conflict between the semantic gender of the referent and the lexically stored gender of the noun. In these cases, semantic gender is able to override lexical gender in the personal pronoun[19], implying its grammatical superiority.

This agreement variation can often be seen cross-linguistically: it is often possible for personal pronouns to agree semantically while other agreement targets agree with the lexical gender of the noun[20].

---

[16] Donaldson, B. (2008). Dutch: A Comprehensive Grammar. Oxon: Routledge Publishers. 2nd Edition p.42

[17] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.90

[18] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.82

[19] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.82

[20] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.82

2.2.   **Individuation: The Hierarchy and its Importance**

The grammatical category of the noun can help determine whether its gender agreement is more likely to occur semantically or lexically. Individuation is the way in which a noun can be identified or differentiated. Different degrees of individuation fall within a hierarchy: the individuation hierarchy. This hierarchy consists of 5 categories of nouns. These are human, animal, object/bounded abstract, specific mass, unbounded abstract/unspecified mass. This is more easily presented on a scale[21]:

Human > Animal > Object/Bounded Abstract > Specific Mass > Unbounded > Abstract/Unspecified Mass
High Individuation   <---------------------------- >   Low Individuation

Parallel to this hierarchy, there is another scale that determines the degree of individuation for the subject depending on whether it has a high or low degree of individuation. More specifically, 'semantic agreement occurs significantly more often with referents at the extreme ends of the hierarchy than with those towards the middle.'[22] Referents at the far ends of the scale are subjected to more discussion in regard to their gender agreement than those in the center of the hierarchy.

Often in Dutch, the gender of a noun is ambiguous and so it is unclear whether the agreement should be semantic or lexical. The coherent grammatical rule that stems from the hierarchy is that 'pronouns display a semantic agreement pattern along the lines of individuation when they do not agree with the lexical gender of the noun.'[23] This rule helps to eliminate the ambiguity of a noun's gender agreement.

Lexical gender is arbitrary and semantic gender is influenced by the degree of individuation and the sex of the referent[24]. The hierarchy is necessary because the degree of individuation helps determine the gender of a noun in situations where the gender of the referent is ambiguous (so as to differentiate between common and neuter) due to the arbitrariness of gender assignment in Dutch.

3. **The Diachrony of Dutch Gender: Semantic Gender Agreement versus Lexical Gender Agreement**

Much research has previously been conducted in regard to the relationship between semantic gender agreement and lexical gender agreement. As we have already seen, semantic gender agreement and lexical gender agreement have a large amount of influence in Dutch gender and there are many rules determining the way in which these principles must be applied to the language in practice. Some data will follow regarding the way in which the balance between the use of lexical gender agreement and semantic gender agreement has shifted between middle Dutch and modern Dutch. This shift appears to be a direct result of the conflation of the masculine and feminine gender into the common gender.

---

[21] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.91

[22] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. (quoting Audring 2009:167-168) p.92

[23] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.91

[24] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.92

Table 1: This is a table showing the ratio of semantic to lexical agreement in middle Dutch and present-day Dutch[25]

Mass (Conflict)	Middle Dutch Texts (masculine and feminine antecedents)	Present-Day Dutch (common antecedents)
Non-Lexical	27%	50%
Lexical	73%	50%
Total	100%	100%

Table 1 is useful in recognising the diachrony of these two types of gender agreement between middle Dutch and Modern Dutch. Interestingly, it shows that semantic gender agreement (marked on the table as 'non-lexical') occurs almost twice as much nowadays compared to when middle Dutch was spoken, whereas the use of lexical gender agreement has reduced by almost 25%.

The following table of data shows the way in which adnominal gender marking has also changed in the period between middle Dutch and modern Dutch, and more specifically shows which markers have remained the same, those that have been lost and those that have changed over the years. Adnominal gender markers tend to indicate lexical gender agreement. The data shown in table 2 suggests that the loss of some of these markers has resulted in a corresponding decrease in the use of lexical gender agreement in favour of semantic gender agreement. This is supported by the data presented in the previous table showing the frequency with which semantic gender agreement and lexical gender agreement are used.

---

[25] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.113

Table 2: Table showing the Adnominal Gender Marking Comparison[26]

Adnominal element	Middle Dutch	Present-Day Dutch
Definite article	yes	yes
Indefinite article	yes	no
Demonstrative determiners	yes	yes
Possessive determiners	yes	Only 1SG.PL
Attributive adjectives	yes	Only indefinite NPs

4. Analysis and Discussion

The conflation of the masculine and feminine genders into the common gender most likely play a role in the shifts that have occurred in semantic and lexical gender agreement. Indeed, any uncertainty in regard to the gender of a referent with common gender nouns may result in a heavy reliance on semantic gender agreement, which would not have been an issue back before the conflation of masculine and feminine into the common gender.

Furthermore, these changes to semantic and lexical gender agreement have been affected by how frequently lexical gender is marked (as shown in the previous table of data[27]). This would have been influenced by conflation but is not directly caused by it, rather the changes have occurred as an extension of the reduction in the use of lexical gender agreement, thus meaning the demand for a wider variety of adnominal gender markers is no longer as relevant. In addition, given that the number of nominal genders has reduced from 3 to 2, it would be expected that we would also see a reduction in the elements that mark gender.

Semantic gender agreement is not a new development that coincided with the conflation of the Dutch genders. Indeed, semantic gender agreement existed before the change from 3 nominal gender to 2 nominal genders[28]. The frequency with which semantic gender agreement is used is what changed as a result of the conflation of masculine and feminine into the common gender.

The concept that semantic interpretation has been enabled by individuation can possibly be attributed to being an old West Germanic feature. Indeed, 'it seems that, at least in the West Germanic languages, semantic agreement with animate referents has long existed alongside lexical agreement'[29]. This feature could also possibly even be traced back to Proto-Indo-European.

---

[26] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.115

[27] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.115

[28] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. pp.87-88

[29] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.84

5. The Broader Perspective: Indo-European Languages

The Indo-European languages are predominantly spoken and native to the European continent and some areas of Asia and India. Indeed, 'more than 95 per cent of Europeans now speak an Indo-European language'[30].   The Indo-European language family stems from one original language referred to as Proto-Indo-European, which will be furthermore referred to as PIE. In present times, the language most similar to PIE in terms of grammar and lexicon is Lithuanian.

There is much frustration to be had for the learner of a new Indo-European language concerning mastering the gender systems. Each language has such drastically different ways of expressing gender and more often than not, these methods are mostly arbitrary. Indeed, 'the gender categories of most Indo-European languages[...]do not agree with anything in the practical world[...]There seems to be no practical criterion by which the gender of a noun in German, French, or Latin could be determined.'[31]

West Germanic semantic gender agreement has long existed alongside lexical gender agreement, as outlined in the previous section. However, this is no longer the case in English, which now only retains its semantic gender agreement, having completely lost any lexical gender agreement back around the middle ages.

In Romance varieties of the Indo-European language family referents of low individuation are classified as being neuter as per the pattern followed in the individuation hierarchy. This is a fairly new development within the linguistic structure of this language family and is an example of how many languages are starting to follow similar grammatical rules that they didn't previously.

In addition, regarding the similarities in the use of the individuation hierarchy, semantic gender agreement with low individuation in neuter nouns can also be found in Danish, Norwegian and Swedish[32].

As to the quantity of genders across language within the Indo-European family: German, Icelandic, Serbo-Croat and Russian have three genders. However Dutch, Italian and French have reduced their gender systems from 3 genders to 2 genders in various ways[33] and German is beginning to show signs of the relaxation of more grammatical elements of its own. The fact that these changes have occurred in more languages than just in Dutch implies that many languages are evolving in ways that almost appear as if to mimic one another in terms of the simplification of grammatical gender.

A point that is particularly interesting as to these widespread changes is that recent research has shown that the PIE gender system was originally constructed with only 2 genders as opposed to 3[34]. This could be supported by the fact that Lithuanian - the language that is currently the closest to PIE in terms of lexicon and grammar these days[35] - also only have two genders: masculine and feminine. If PIE did indeed have only 2 genders then the main inserting point here is that all Indo-European languages stemmed from PIE and it appears as though many are currently going through the process of reverting back to the original language form of their common mother language. However, these conclusions regarding the gender structure

---

[30] Dorren, G. (2015) Lingo: A Language Spotter's Guide to Europe. London: Profile Books p.16

[31] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. p.40 quoting Bloomfield, 1933; as quoted in Zubin and Kopcke, 1981, p.439

[32] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.86

[33] Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. p.18

[34] Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. p.88

[35] Dorren, G. (2015) Lingo: A Language Spotter's Guide to Europe. London: Profile Books pp15-19

of PIE are still under discussion and so this conclusion is purely hypothetical as this issue is still being contested[36].

6. Conclusion

Many elements in Dutch grammar have contributed to the drastic changes in terms of the diachrony of gender in the Dutch language.

The diachrony of gender has occurred as a natural process of language change; all of the components discussed in this paper have changed due to the conflation of masculine and feminine into the common gender, whether directly or indirectly. Semantic gender agreement and lexical gender agreement have both been affected by the diachrony of gender, particularly in regard to the frequency with which they are used and the balance of their use thereof. In addition, the use of adnominal gender markers has changed drastically between middle and modern Dutch as a result of both the reduction in the number of nominal genders and the reduced use of lexical gender agreement.

Furthermore, the reduction in genders in other languages and the way in which semantic gender is applied using the same principles from the individuation hierarchy and the scale of low to high individuation shows that some of these changes are also applicable to other languages within the Indo-European region.

References

Audring, J. (2006). Pronominal Gender in Spoken Dutch. Journal of Germanic Linguistics. 18.2 (2006): 85-116

De Vogelaer, Gunther. 2009. Changing pronominal gender in Dutch: Transmission or diffusion? *Language variation, European perspectives*, vol. 2, ed. by Tsiplakou, Stavroula, Karyolemou, Marilena, & Pavlou, Paul, 71-80.

Donaldson, B. (2008). Dutch: A Comprehensive Grammar. Oxon: Routledge Publishers. 2nd Edition

Dorren, G. (2015) Lingo: A Language Spotter's Guide to Europe. London: Profile Books

Kraaikamp, M. (2017). Semantic versus Lexical Gender: Synchronic and Diachronic Variation in Germanic Gender Agreement. Digital Academic Repository. Retrieved from URL: https://www.lotpublications.nl/Documents/445_fulltext.pdf

Nettle, D. (1999). Is the Rate of Linguistic Change Constant? Merton College, Oxford and King's College Research Centre, Lingua 108 (1999) 119-136ere

Van Berkum, J.J.A. (1996). The Psycholinguistics of Grammatical Gender: Studies in Language Comprehension and Production. Chapter 2 pp.15-48. Retrieved from URL: repository.ubn.ru.nl/bitstream/handle/2066/146162/mmubn000001_217863914.pdf

---

[36] Proto-Indo-European is a language that has been reconstructed by historical linguists in an attempt to trace modern languages back to their common ancestor. This fact should be taken into consideration when applying theories such as these to the grammatical rules that apply to PIE.