Fehringer C.

[Internal constraints on the use of *gaan* versus *zullen* as future markers in spoken Dutch.](#)

Nederlandse Taalkunde 2017, 22(3), 359-387.

Copyright:

This is the authors' accepted manuscript of an article that has been published in its final definitive form by Amsterdam University Press, 2017.

DOI link to article:

<https://doi.org/10.5117/NEDTAA2017.3.FEHR>

Date deposited:

04/07/2017

This work is licensed under a

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)

Internal constraints on the use of *gaan* versus *zullen* as future markers in spoken Dutch: a quantitative variationist approach.¹

Carol Fehringer

Abstract

This paper provides a quantitative variationist analysis of the future markers *gaan* and *zullen* in modern spoken Dutch in the light of current discussion on grammaticalisation. In order to determine to what extent *gaan* has been grammaticalised, internal constraints on the distribution of *gaan* and *zullen* such as verb type, animacy of subject, grammatical person, clause type, sentence type and proximity of future reference are investigated, using Netherlandic and Belgian data from the Corpus of Spoken Dutch. The purpose of investigating constraints such as animacy of subject and proximity of future reference is to establish to what extent the behaviour of *gaan* is still determined by its lexical history (i.e. as a verb denoting motion towards an immediate goal), and results demonstrate that these two constraints persist in modern Dutch. Moreover, research on grammaticalisation in other languages has shown that grammaticalised markers can become specialised to certain morphosyntactic environments such as particular grammatical persons or types of clause. Indeed, it will be shown here that the use of *gaan* is favoured in subordinate clauses and non-first person utterances, which it has in common with English *be going to*. By contrast, *zullen* is shown to occur in pragmatically more nuanced contexts, indicated by the presence of modal particles.

¹ Many thanks to the two anonymous referees for their very helpful comments.

Keywords: future, grammaticalisation, variation, Dutch, modal

1. Introduction

1.1. The andative ('going to') future

In recent decades much research has been carried out on andative future constructions, in which a verb with the original lexical meaning of 'to go' develops into a marker of future tense through the process of grammaticalisation. For instance, English *be going to*, which originally denoted movement along a path: e.g. *I am going* (somewhere) *to marry Bill*, became an indicator of future time: e.g. *I am going to marry Bill* = *I will marry Bill* (Hopper and Traugott 2003: 2-3). This newly emerged future marker then came to exist alongside the older established markers *will* and *shall*. Such a development can also be found in Dutch, in which the lexical verb *gaan* 'to go' has come to be used together with an infinitive as a future marker: e.g. *Ik ga volgend jaar een huis kopen* 'I'm going to buy a house next year' (Donaldson 1997: 140). Now, *gaan* competes with the older established auxiliary *zullen* to mark future time: e.g. *Ik ga een huis kopen* vs. *Ik zal een huis kopen* 'I'm going to buy a house' vs. 'I will buy a house'.

Bybee, Perkins and Pagliuca (1994: 5-6) point out that, during the process of grammaticalisation, the original meaning of the lexical form may persist to some extent in the new grammatical marker and that these varying degrees of persistence can indicate how far along the path of grammaticalisation a particular construction is. In other words, a future marker's lexical history may be reflected in constraints on its current distribution (Hopper 1991). For instance, if we consider Dutch *gaan* as an indicator of movement towards a

goal, we would expect it to co-occur with animate rather than inanimate subjects: e.g. *Ik ga een huis kopen*. Therefore, if in one variety of Dutch (variety A) *gaan* is favoured with animate subjects while in another (variety B) it is unconstrained by animacy (e.g. *Gaat het regenen?* 'Is it going to rain?' (Donaldson, 1997: 140)), this could indicate that the construction is at a more advanced stage of grammaticalisation in variety B: i.e. the verb has been semantically bleached and can be used more generally in a wider range of contexts.

Indeed, quantitative, variationist studies of English *going to*, such as those of Poplack and Tagliamonte (2000), Torres-Cacoullos and Walker (2009) and Tagliamonte (2013), investigate processes of grammaticalisation by examining the internal factors which contribute to the variation of forms. For instance, not only animacy of subject but also clause type, sentence type, grammatical person and proximity of future reference have all been shown to play an important role in the distribution of competing future markers. Such studies allow us to gain a deeper understanding of the process of ongoing grammaticalisation in that they reveal how variant forms can become generalised and/or conventionalised in certain contexts.

By contrast, research on Dutch *gaan* has largely focused on overall frequency effects (e.g. Kraaikamp 2009), collocational preferences (Hilpert 2008) and semantic and/or pragmatic accounts of its use (Weijers 2012; Beheydt 2005) in order to gain some insight into the grammaticalisation process. Possible internal constraints on the use of *gaan*, such as the influence of clause type, sentence type and grammatical person, for example, have not been investigated to date, at least not in a quantitative analysis. Thus, the present paper will contribute to ongoing research into the grammaticalisation of *gaan* by applying the methodology of quantitative variationist studies such as those of Poplack and Tagliamonte (2000) and Tagliamonte (2013) in order to ascertain whether this future marker is continuing to undergo grammaticalisation to the same degree as has been reported for English *going to*.

In order to do this, the frequency and distribution of *gaan* versus *zullen* will be investigated using the Corpus of Spoken Dutch (*Corpus Gesproken Nederlands*), which offers a wide range of spoken data from both The Netherlands and Belgium. It is well known that the use of future *gaan* is more widespread in Belgium than in the Netherlands, in that it is more frequent and co-occurs with a wider range of verbs (van Bree 1997, Colleman 2000, Decroos 2000). The current analysis will also allow us to compare and contrast the behaviour of each variety regarding the extent to which specific internal factors determine the choice of future marker.

1.2. *Gaan* + infinitive as a future marker in Dutch

Pedagogical grammars of Dutch for English speakers (e.g. Donaldson 1997, Shetter and Ham 2007) state that the use of *gaan* + infinitive to refer to future time in Dutch is roughly equivalent to that of *going to* in English. Like English *going to*, *gaan* is particularly common in the spoken language but less so in writing (Beheydt 2005: 258). In spoken Dutch it regularly competes with the other syntactic future marker *zullen*, which is roughly equivalent to English *will*.² Consider the following examples from the Corpus of Spoken Dutch (CGN):³

(1) a. Ik ga morgenmiddag eens bellen (vl)	I'm just going to ring tomorrow afternoon
b. Ik zal vanavond eens naar de Glenn
bellen (vl)	I'll just ring Glenn this evening

² Note that the present tense can also be used to refer to future time: e.g. *Ik ga morgen naar de kerk* 'I'm going to church tomorrow'.

³ Unless stated otherwise, all Dutch examples are taken from the CGN (*Corpus Gesproken Nederlands*), which is discussed in more detail in the data collection and analysis section. Following CGN conventions, examples from the Netherlandic Dutch sub-corpus are labelled (nl) and those from the Belgian Dutch sub-corpus are labelled (vl), the latter being an abbreviation of *Vlaams* 'Flemish'.

In most of the future tense contexts found in the CGN, *gaan* and *zullen* appear to be used interchangeably. However, there is much debate in the literature as to how semantically equivalent these two variants are. Some researchers argue that there are semantic and pragmatic differences associated with the use of these future markers. For instance, the standard reference grammar of Dutch for Dutch speakers, the *Algemene Nederlandse Spraakkunst* (ANS), states that *zullen* has modal nuances and tends to be used to express uncertainty, whereas *gaan* usually expresses intention. Consider the following two examples and their interpretations:⁴

(2) a. Wat zal er morgen gebeuren?	What'll happen tomorrow?
<i>Wat hangt ons boven het hoofd?</i>	<i>What's hanging over our head?</i>
b. Wat gaat er morgen gebeuren?	What's going to happen tomorrow?
<i>Wat staat er op het programma?</i>	<i>What's on the programme?</i>

Example (2a) implies that the speaker does not know what will happen tomorrow and has no control over the events whereas example (2b) implies that someone has planned something for the next day and the speaker is enquiring what the plan is. However, the ANS then goes on to say that this difference is only apparent in interrogative contexts. It is much less clear, and even 'practically absent' in declarative utterances (E-ANS: 18.5.4.3.iii).

The modal nuances associated with *zullen* can be explained by investigating the verb's history. Early attestations of the verb used as an auxiliary demonstrate that it had deontic meaning ('must', 'should', 'be obliged to', 'be necessary'), as the following examples from the *Woordenboek der Nederlandsche Taal* illustrate (WNT, *ZULLEN*, 500-1200, 2):

⁴ These are taken from the electronic version of the ANS, namely the E-ANS (18·5·4·3·iii).

- (3) a. Ther kuning leydoda mich in sinan winkelnere ande wisda mich, wie ich hina minnan
sule (LW 030,03 Egmond, Holland, ca. 1100).

'The king lead me to his wine cellar and taught me how I **should** love him.'

- b. Thu **scalt** hiro saluti congaudere (LW 137,15 Egmond, Holland, ca. 1100).

'You **must** rejoice at their salvation.'

- c. Ande thaz **sculon** sie duon propter timores nocturnos (LW 051,23 Egmond, Holland,
ca. 1100).

'And they **must** do that because of the night terrors.'

In addition to this usage, *zullen* was also used as a 'hulpwerkwoord van de toekomstige tijd' (WNT, *ZULLEN*, 500-1200, 3), i.e. a future auxiliary, and also as a modal verb with an epistemic function (expressing supposition). The WNT gives examples (4a) and (4b) of *zullen* as a future auxiliary and example (4c) demonstrates its use as an epistemic modal:

- (4) a. Inde nuesan **sal** also holz that gesazt uuart bi fluzze nuassere
(WPs (hs. FA) 001,03 Zuid-Nederriijn, Nederriijn, 901-1000 (kopie 1591-1610)).

'And he **will** be like a tree that was placed by a stream.'

- b. Ic **sal** beidan sin thie behaldon mi deda Fan Luzzilheide geistis in fan geuuidere.

(WPs (hs. H) 054,09 Zuid-Nederriijn, Nederriijn, 901-1000 (kopie 1595-1605)).

'I **will** look for him who saved me from pusillanimity of spirit and against the storm.'

- c. Iz **solten** thrie himele [wesen] nach them, daz wir hauen gelesen.

(Mfr.Reimb. C, r. 006 Werden, Essen?, Noord-Oost Nederland, 1151-1200).

'There **should** be three heavens, according to what we have read.'

It should be noted, however, that it is often difficult to differentiate between the future and modal use of *zullen*, both diachronically and synchronically, which leads many researchers to believe that there cannot be an absolute distinction between the temporal and modal reading of the verb.⁵

If we consider the earliest uses of *gaan* plus infinitive, we see that it is used with its lexical meaning 'to go', specifically meaning 'to walk on foot'. In the following examples, the subject is moving towards a goal where an action will then be carried out (WNT, *GAAN*, 1200-1300, I. A. 4):

- (5) a. Tho sie ire gebet habeten getan, uor thaz kint sie **giengen** stan.

(*Mfr.Reimb. B, r. 162* Werden, Essen?, Noord-Oost Nederland, 1151-1200).

'When they had finished their worship they **went** to stand in front of the child.'

- b. Si verbiten delpendiere alsi **gaen** drinken ter riuere.

(*Nat.Bl.D p. 236, r. 9-11*, West-Vlaanderen, 1287).

'They bite the elephants to death when they **go** to drink at the river.'

- c. Die te prijmetide in den wijngart werken **gaen**, dat sijn die in harre kintscheit hen te

Gode bekiren. (*Diat. p. 152, r. 31-32*, Brabant-West, 1291-1300).

'Those who at the first hour **go** to work in the vineyard are they who in their childhood are converted to God.'

Later, in the seventeenth, eighteenth and nineteenth centuries, we find occurrences of *gaan* + infinitive with future reference. These often express the meaning 'to prepare oneself to carry

⁵ Indeed, there has been some disagreement in the literature as to the status of *zullen*. Beheydt (2005: 266) makes the point that, although *zullen* can often have modal connotations in modern Dutch, one should not treat this verb as a modal auxiliary, as it can also be used to express pure futurity. By contrast, Verkuyl and Broekhuis (2013) argue that *zullen* should be treated purely as a modal verb with no temporal meaning at all (see the discussion between Boogaart 2013 and Verkuyl and Broekhuis 2013).

out an action' (WNT: *GAAN*, 1500 onwards, 1.1), and the use of *gaan* denotes intention on the part of the subject:⁶

- (6) a. Thans **ga** ik het gewigtigste van uwen brief omstandig beantwoorden

(WOLFF en DEKEN, *Leev.* 1, 57 [1784]).

'Now I'm **going** to give a full answer to the most important part of your letter.'

- b. Nu verzoek ik, dat gy attent zyt, want ik **ga** u iets vertellen,

(WOLFF en DEKEN, *Leev.* 1, 264 [1784]).

'Now I'm asking you to pay attention because I'm **going** to tell you something.'

- c. Zie, man! wat **ga** ik je doen: ik beleende mijn' gouden boog,

(FOKKE, *Verz. W.* 12, 83 [1802]).

'Look, man, what am I **going** to do to you: I pawned my golden bow.'

Despite the existence of such constructions in the seventeenth, eighteenth and nineteenth centuries, researchers such as Hilpert (2008: 183) and Weijers (1012: 19) point out that *gaan* as a future marker did not become frequent until the twentieth century. Weijers' (2012) examination of the *Van Dale* Dutch dictionary of 1898 reveals that the only examples that are close to future constructions with *gaan* are those in which it is coupled with another verb by using *en* 'and': e.g. *Wilt gij weten, hoe het Hollandsch niet behoort te worden gesproken, zoo **ga en zet** u neder in den Amsterdamschen schouwburg*. 'If you want to know how Dutch is not supposed to be spoken, then **go and sit** down at the theatre of Amsterdam.' (Weijers, 2012: 20-21). Such combinations are no longer usual in modern standard Dutch, however, where the *gaan* + infinitive construction is preferred (Weijers, 2012: 21).

⁶ In addition to this usage there are also instances of *gaan* expressing the beginning of an action (WNT, *GAAN*, 1500 onwards, I. II): e.g. *Straks vóór beurstijd **gingt** gij het geheim mij toevertrouwen* (CONSC. 4, 52 b [ed. 1869]) 'Soon before market you **started** to entrust me with the secret.' This inchoative use of *gaan* is common today: e.g. *Ze **gaat** huilen* 'she's starting to cry.'

Thus, the meaning of intention associated with self-propelled movement gave rise to a sense of immediacy, which then developed into future meaning. This notion of immediacy has been used to account for the observation that *gaan* + infinitive is often used to refer to the imminent rather than the more distant future (see e.g. Ten Cate, 1991), an observation that has often been made for the preference of *going to* over *will* in English (e.g. Royster and Steadman 1923).

One important difference between the *gaan* future in Dutch and *going to* in English is that *gaan* is used more frequently with some verbs than with others, whereas *going to* is less lexically restricted.⁷ Much of the research on Dutch *gaan* has concentrated on establishing which types of verb can be used with this future marker and on attempting to find reasons for these preferences. For instance, the ANS identifies groups of verbs that regularly co-occur with *gaan* as a future marker: i) verbs denoting movement from one place to another: e.g. *wandelen* 'to go for a walk', *slapen* 'to go to bed'; ii) verbs denoting important life changes: e.g. *trouwen* 'to get married', *verhuizen* 'to move house'; iii) verbs denoting weather conditions: e.g. *regenen* 'to rain', *waaien* 'to be windy' (E-ANS:18·5·4·3·iii).

Beheydt (2005) argues that all of the groupings listed in the ANS can be subsumed under the general interpretation of *gaan* as expressing an action whose origin is rooted in the present (see also Ten Cate 1991 and De Rooij 1985 & 1986 for a similar interpretation). This approach accords with what Leech, Hundt, Mair and Smith (2009: 107-108) refer to as the textbook definition of *going to* in English: it is a future marker that refers to 'a future happening that in some sense is implicit in the present state of affairs — typically either an outcome of existing intentions or existing causes — often with the implication that the future event will happen soon'. For example, the reason for the use of *gaan* with *trouwen* is that one has already made one's mind up to marry and perhaps has already started the preparations.

⁷ Having said that, *going to* does have clear collocational preferences, as Hilpert's empirical study demonstrates (2008: 41).

Conversely, ANS notes that *gaan* is rarely used with the stative verbs *hebben* 'to have' and *zijn* 'to be' (E-ANS: 18·5·4·3·iii). Beheydt (2005: 252) argues that the reason for this is that these verbs semantically clash with the idea of intentional actions. In sum, Beheydt (2005) concludes that the exact interpretation of *gaan* is a pragmatic matter.

This view is not shared by Hilpert (2008), however, who argues that intention is not as central to the constructional meaning of *gaan* + infinitive as has been argued by some scholars. His study investigates the collocational preferences of *gaan* at different historical stages in order to establish how its present-day grammatical meaning developed out of its original lexical meaning of movement. His results show that *gaan* frequently attracts collexemes with a low degree of intentionality: for instance, weather verbs and verbs denoting human activities that are often involuntary eg. *twijfelen* 'to doubt', *voelen* 'to feel' and *sterven* 'to die'. Also, *gebeuren* 'to happen' is a frequent collexeme. Hilpert's (2008:188) conclusion is that *gaan* has undergone an important semantic development: early usages refer to events involving motion and intention, and then, subsequently, the constructional meaning of *gaan* broadens to occur with verbs that are not necessarily compatible with the meaning of intentionality. This suggests that intention plays less of a role in the interpretation of *gaan* as a future marker in contemporary Dutch than it did in the earlier stages of the construction.

Kraaikamp (2009) carries out a quantitative investigation of the relative frequency of *gaan* versus *zullen* in the Netherlandic Dutch sub-corpus of the CGN. She focuses on three frequently occurring verbs denoting an action: *doen* 'to do', *zoeken* 'to look for', *bellen* 'to ring up'; and two 'passive' verbs: *worden* 'to become' and *krijgen* 'to get/receive'. Her results show that the 'action' verbs favour the use of *gaan* over *zullen* by 81%, whereas the 'passive' verbs occur with a lower rate of *gaan* (53%).

As Kraaikamp's (2009) study only examines Netherlandic Dutch, the use of *gaan* appears to be fairly restricted. If, however, we were to take Belgian Dutch into account, we

might find a different picture. For instance, Coleman (2000) demonstrates how *gaan* is more widespread in Belgian Dutch than in Netherlandic Dutch. It has a higher frequency of use and can be used with verbs which would normally disallow it in Netherlandic Dutch, namely the stative verbs *hebben* and *zijn*, modal verbs and *gaan* itself (e.g. *Ik ga gaan* 'I'm going to go'). Although Belgian Dutch *gaan* has a similar history to its Netherlandic counterpart, the WNT (GAAN, 1500 onwards, I. II, footnote 1) points out that some Belgian writers were using *gaan* as a pure future auxiliary when Netherlandic use was primarily denoting the preparation for an action: e.g. *Eilaas, de Gendarmes gaan komen, vader - om u te vangen!* (CONSC. 2, 190 b [ed. 1868]) 'Alas, the gendarmes are **going** to come, father - to catch you!'; '*Gaat gij zwijgen?*' *bulderde Margriet* (CONSC. 2, 408 b [ed. 1868]) 'Are you **going** to keep quiet?' roared Margriet.' This Belgian practice was believed to have been influenced by French and, according to the WNT is ... *als een gallicisme af te keuren* ('is to be rejected as a Gallicism').

The fact that *gaan* can combine with *gaan* itself in Belgian Dutch and not in Netherlandic Dutch is explained by de Rooij (1986: 17) as having to do with the greater degree of grammaticalisation of the Belgian Dutch auxiliary. He states that the Netherlandic auxiliary *gaan* still has 'meaning' for the speaker and the co-occurrence of this with *gaan* would be considered a tautology. In Belgian Dutch, on the other hand, it is more 'auxiliary-like'. Indeed, in some regional varieties, particularly those of West Flemish, *gaan* has completely ousted *zullen* as the dominant future marker (Beheydt, 2005:256).⁸

Summing up, the future tense system of Dutch has features which appear to be characteristic of grammaticalisation in progress. There is competition between an older, more grammaticalised, and a younger, less grammaticalised, construction (*zullen* versus *gaan* respectively). There has been semantic bleaching of *gaan* so that it can be used in contexts other than those where the lexical meaning 'to go' is required. This semantic process is

⁸ In some varieties of West Flemish we can find multiple uses of *gaan*: *We gaan hier eens gaan gaan eten* 'We are going to go and go eat here', cited at <http://www.vlaamswwoordenboek.be/definities/toon/12907>.

context dependent and it can therefore be described in terms of context-induced reinterpretation (see Heine and Kuteva 2002: 2-3). When forms for concrete (lexical) meanings are used to also express more abstract (grammatical) meanings, grammatical forms emerge: in this case, an action 'go to' is used to refer also to a grammatical notion 'future'. Here we are dealing with a unidirectional process: *gaan* becomes more grammatical over time. As a result, such context-induced reinterpretation has often been described in terms of metaphorical transfer: leading from movement in space to movement in time (see e.g. van Bree 1997: 72). Once *gaan* has gained the grammatical function of a future marker, it is in a position to generalise to areas in which, historically, it was unacceptable. In the following sections we will present data from the Corpus of Spoken Dutch in order to examine the distribution of *gaan* vs. *zullen* in a wide variety of contexts. Our methods for data collection and analysis are set out below.

2. Materials and methodology

2.1. Frequency of *gaan* versus *zullen* in the Corpus of Spoken Dutch

The Corpus of Spoken Dutch (*Corpus Gesproken Nederlands* or CGN) is an electronic corpus of contemporary Dutch, as spoken by adults in Flanders and the Netherlands, spanning the years 1998 to 2004.⁹ The CGN consists of a number of sub-corpora: e.g. spontaneous conversations ('face-to-face'), telephone conversations, interviews, business speech, classroom lessons and debates. Our analysis focuses on the first sub-corpus (spontaneous face-to-face conversations), which comprises a total of 3,000,000 words, as this genre has

⁹ Full information on this corpus can be found at <http://lands.let.ru.nl/cgn/ehome.htm>.

been argued in the literature to show high frequencies of *gaan* as a future marker (Beheydt 2005: 259).

Following work by Ten Cate (1991), Kraaikamp (2009), and studies of English *going to* versus *will* (Berglund 2000, Szmrecsanyi 2003; Tagliamonte 2013), our analysis concentrates on the use of *gaan* as opposed to *zullen*, as these are the two syntactic markers of future time in Dutch. Although the present tense can also be used to refer to future events, in such cases the future reference lies in the grammatical context rather than in the verb form itself, which makes it fundamentally different from the two syntactic future markers (Berglund 2000: 29). Therefore, all of the future utterances in our data set contain either *gaan* or *zullen*.

Building on the work of Kraaikamp (2009) we investigated the use of *gaan* with action verbs and passive verbs, using the same tokens as in her analysis (action verbs: *doen* 'to do', *zoeken* 'to look for', *bellen* 'to ring up'; passive verbs: *krijgen* 'to get/receive', *worden* 'to become'). These verbs were chosen on the basis of their relatively high frequency in future utterances in the corpus. In addition, we also noted that certain other common verbs such as *zien* 'to see' and *gebeuren* 'to happen' appeared to pattern with the passive verbs *worden* and *krijgen* in their use of *gaan*. These verbs all have in common that their subjects are not agents, but patients or experiencers. For this reason we labelled our groups 'agent verbs' and 'non-agent verbs' respectively in order to include *zien* and *gebeuren* in our analysis.¹⁰

Following work on Belgian Dutch (e.g. Coleman 2000, Decroos 2000, van Bree 1997), we also added the lexical verb *gaan* as a separate category, as the co-occurrence of the future marker *gaan* with lexical *gaan*, e.g. *Ik ga nu gaan* 'I'm going to go now', has been discussed as a measure for the grammaticalisation of the future marker in Belgian Dutch (see also de Rooij 1986: 17). It has also been pointed out that the stative verbs *hebben* and *zijn* can co-

¹⁰ Note that in colloquial Belgian Dutch the verb *zien* can also be used as an agent verb: e.g. *We gaan morgen zien voor een nieuwe laptop* <http://9maand.be/forum/moeilijk-zwanger-worden/zwanger-worden-met-endometriose.147610/page/52> 'We're going to look for a new laptop tomorrow'. Such uses were not included in our data set.

occur with both *zullen* and *gaan* in Belgian Dutch, whereas they rarely occur with *gaan* in Netherlandic Dutch.¹¹ Therefore, we also decided to investigate the relative frequency of *gaan* with *hebben* and *zijn* compared to the other three verb types in Belgian Dutch. Below are some examples of the verb types included in our data set:¹²

(7) Agent Verbs

a. Ik ga mijn best doen (vl)	I'm going to do my best
b. Ze gaan een huizeke zoeken in Frankrijk
voor zes personen (vl)	They are going to look for a house
in France for six people
c. Ze gaat vrijdag bellen hè (vl)	She's going to ring on Friday, isn't she?

(8) Non-agent Verbs

a. Zij zei dat gaat hier veel te druk
worden (vl)	She said it's going to get far too busy
here
b. Ik zeg oh nee nu gaan we 't krijgen
hè (vl)	I say oh no we're going to get it now,
aren't we?
c. Ik ben echt benieuwd wat er volgend jaar
gaat gebeuren bij de verkiezingen (vl)	I'm really curious to know what's going
to happen next year at the elections
d. Ik ga mijn vrienden voor negen maanden
niet meer zien (vl)	I'm not going to see my friends for nine
months

(9) *Gaan*

¹¹ There were only 3 occurrences of *gaan* with *hebben* or *zijn* in our Netherlandic data set.

¹² The other verb type that has been reported to co-occur with *gaan* as well as *zullen* in Belgian Dutch is modal verbs. However, a preliminary examination of the frequency of *gaan* with modals shows that its distribution within the category is very uneven, which makes it difficult to treat modal verbs as a unified group. In this data-set, *moeten* and *kunnen* had a roughly even distribution of *gaan* and *zullen* (51.5% *gaan* for *moeten* and 53.7% *gaan* for *kunnen*), yet *willen* and *mogen* occurred almost exclusively with *zullen*. For this reason, modal verbs were not included in the investigation.

a. Maar dat gaat toch moeilijk gaan hè (vl)	But that's going to be difficult isn't it?
b. Hoe gaat dat gaan ? (vl)	How's that going to go?

(10) *Hebben/zijn*

a. Met het nieuw systeem gaan we dus
problemen hebben hè (vl)	So we are going to have problems with
the new system, aren't we?
b. Ik ga frank zijn zegt ze ik betaal u voor
niks (vl)	I'm going to be impertinent she says I'm
not paying you anything
c. Ik denk dat 't in 't Engels gaat zijn (vl)	I think it's going to be in English

Data was collected for the verb groups specified above and each group and variety were coded separately. As the study uses a variationist/Labovian research design, which is interested in 'alternate ways of saying "the same" thing' (Labov 1972: 188), only utterances in which both future markers could potentially vary were included in the analysis. Thus, occurrences of *gaan* which did not refer to future time were excluded: for instance, *gaan* used with an inchoative function meaning 'to begin to, start to' (11a)¹³, *gaan* used in the lexical sense of 'go' (11b), *gaan* used in verb forms other than the present (11c). In none of these types of utterance do we find *zullen* competing with *gaan*; therefore, they were not considered in the analysis.

(11)a. Want dan krijg ik altijd ja ... dan gaan
mijn benen pijn doen weet je wel (nl)	Because then I always get ... then my
legs start to hurt, you know
b. Omdat 'k niks verdien gaat ie één keer	Because I don't earn anything he goes

¹³ See Weijers (2012: 25-26) for discussion of this function of *gaan*.

- in de maand mee boodschappen doen (nl) shopping with me once a month
- c. En dan ben ik filosofie gaan doen (vl) And then I (went and) did philosophy

Conversely, occurrences of *zullen* that did not express future time were also excluded: for example, in its function as an epistemic modal expressing supposition: e.g. *Hij zal nu boven de veertig zijn, hè?* (nl) 'He'll be over forty now, won't he?' After filtering out the non-variable contexts we were left with a total of 1782 future tense utterances in our data set: 733 for Netherlandic Dutch and 1047 for Belgian Dutch. The total number of utterances with *gaan* is 1010 and with *zullen* 772 (see Figures 1 and 2 for a breakdown of the frequency of *gaan* versus *zullen* in each region).

In addition to classifying our data into the verb types outlined in (7-10) above, we also marked utterances for other internal constraints that might throw light onto the grammaticalisation of *gaan*. These are all constraints that have been investigated in the variationist literature on English *going to* (e.g. Poplack and Tagliamonte 2000): animacy of subject, proximity of future reference, grammatical person, clause type (main versus subordinate), sentence type (negative versus affirmative, interrogative versus declarative). In addition to these, we also investigated the use of *gaan* and *zullen* with modal particles, as these are an important feature of spoken Dutch and might provide us with further insight into the distribution of the future markers. A full discussion of these constraints can be found in section 2.2 below.

2.2. Further internal constraints on the distribution of *gaan* versus *zullen*

2.2.1. Animacy of subject

Historically, *gaan* with its lexical meaning of 'to go', deriving from the earlier specific meaning of 'to walk on foot' (Weijers 2012:19), occurred exclusively with animate subjects. With the subsequent semantic bleaching of *gaan* due to grammaticalisation, it began to spread to non-animate subjects. Indeed, our data demonstrate that *gaan* can now be used with a range of non-animate subjects, along with *zullen*:

(12)a. Gaat het morgen mooi weer worden? (vl)	Is it going to be nice weather tomorrow?
b. Maar weet ge wat ook 't probleem gaat zijn? (vl)	But do you know what the problem's going to be?
c. Ja van die boom toch die bij je gesnoeid gaat worden? (nl)	Yes, that tree that's going to be pruned at yours?
d. De bediening zal moeilijker worden (nl)	The operation will become harder
e. Dat zal een probleem zijn (vl)	That'll be a problem

Note that not all scholars define this variable in the same way. Some make a three-way distinction between 'human animate', 'non-human animate' and 'inanimate'. In this paper we follow the distinction made in Poplack and Tagliamonte (2000) of 'animate equals human' vs. 'non-animate equals non-human'. Occasionally, this causes some difficulty as, for example, company names are categorised here as non-human (therefore non-animate), but one could argue that they could be interpreted as human. In borderline cases such as this we coded the utterance as non-animate.

The argument, made by a number of researchers such as Ten Cate (1991) and Beheydt (2005), that *gaan* is used to express intention on the part of the speaker cannot be applied to utterances such as (12a-c), as non-animate subjects cannot express intention. This reservation

was also expressed by Hilpert (2008: 179) who, in his historical data set, found frequent use of *gaan* with the verb *gebeuren* 'to happen' and also with verbs denoting weather conditions.

Nevertheless, it still could be the case that, when investigating the relative frequencies of the future markers on a larger scale, there is a tendency to avoid the use of *gaan* with non-animates, despite the fact that such combinations are grammatically acceptable. Such a pattern of usage would be evidence for the persistence (in the sense of Hopper 1991) of the original lexical meaning of *gaan* and could point towards an incomplete stage of grammaticalisation.

2.2.2. Proximity of future time reference

Andative future constructions such as *gaan* and *going to* were originally associated with notions of immediacy and proximity, resulting from their sense of movement along a path (Poplack and Tagliamonte 2000: 336). Indeed, researchers such as de Rooij (1985 & 1986) and Ten Cate (1991) argue that *gaan* tends to be used to denote events in the proximate future. If actual usage supports these assertions, we can argue that *gaan* is not completely grammaticalised but there is some persistence of its original semantic function.

Following Poplack and Tagliamonte (2000) and Tagliamonte (2013) we coded future references as 'proximate' if they could be inferred to be within one month of the utterance. In many cases an adverbial of time made it clear as to when the future event would take place. However, where no time reference was present in the utterance, the whole dialogue was examined in order to ascertain a possible time frame. On the occasions where the wider

context provided no clues as to time reference, the utterances were marked as non-proximate.

The following examples illustrate some proximate contexts:¹⁴

(13)a. Ze gaat vrijdag bellen, hè	She's going to ring on Friday, isn't she?
b. Je gaat morgen niet naar de winkels
gaan hè? (vl)	You aren't going to go to the shops
tomorrow, are you?
c. Wat ga je nou met die courgette
doen? (nl)	What are you going to do with that
courgette?
d. En ik zal hem dit weekend nog even
bellen (nl)	And I'll just give him a ring this
weekend
e. Ik zal zoeken of ik die kalender ergens
heb (nl)	I'll have a look to see if I have the
calendar anywhere

2.2.3. Clause type

Much research has been carried out on the effect of clause type on the *going to* future in English, mainly as a result of Royster and Steadman's original (1923: 400) assertion that the construction entered the language mainly through its use in subordinate clauses. Indeed, this syntactic constraint can be observed in many varieties of English. Poplack and Tagliamonte (2000) and Torres-Cacoullos and Walker (2009) find that subordinate clauses have a favouring effect on the selection of *be going to* in Canadian varieties and similar results for British English varieties are presented in Tagliamonte (2013).

Our data include future expressions with *gaan*, competing with *zullen*, in a wide variety of subordinate clauses. Clauses were coded as subordinate if they were introduced by a

¹⁴ Poplack and Tagliamonte (2000) also use the designation 'distal' for events taking place one year or more in the future. As the number of such utterances was relatively low in our data set (N=14, less than 1%), this measure was not used and such utterances were simply labelled 'non-proximate'.

subordinating conjunction (e.g. *dat* 'that', *omdat* 'because', *als* 'if', *wanneer/toen* 'when', *of* 'whether', *hoewel* 'although', *alsof* 'as if', *tenzij* 'unless' etc.), an interrogative used indirectly (e.g. *wie* 'who', *wat* 'what') or a relative pronoun (e.g. *die* 'who/which', *dat* 'which/that'). Some examples are given in (14).

(14) a. Ben 'k benieuwd wat je gaat doen (nl)	I'm curious about what you are going to do
b. Kunnen we over nadenken of we dat
gaan doen of niet (nl)	We can think about whether we are going
to do that or not
c. Ik denk dat we eerst naar ons
restaurantje gaan gaan (vl)	I think that we are going to go to our
little restaurant first
d. De enige die dit gaan zien is 't
management (nl)	The only one who's going to see this is
the management
e. Even kijken wat ik zal doen (nl)	Let's see what I'll do
f. Ik denk dat het bij de kunstwerken
zal getoond worden (vl)	I think that it'll be exhibited with the
works of art

2.2.4. Sentence type (interrogative vs. declarative)

It has been pointed out in the literature on English future markers that sentence type (i.e. whether an utterance is interrogative or declarative) can have an effect on the redistribution of variants over time (e.g. Nesselhauf 2010: 174-175). For instance, Torres-Cacoullous and Walker (2009) and Tagliamonte (2013) demonstrate that *going to* is preferred over *will* in questions. Therefore we felt it necessary to include this variable in our analysis of Dutch, in case it should have a significant effect on the distribution of *gaan* and *zullen*. We coded as

'interrogative' utterances containing question words (e.g. *wat* 'what', *wie* 'who', *waarom* 'why', *wanneer* 'when', *hoe* 'how') and yes/no questions such as those in (15c) and (15e):

(15) a. Wat ga je in godsnaam met die TV
doen? (nl)	What are you going to do
with that TV, for God's sake?
b. Wanneer gaat dat gebeuren dan? (nl)	When's that going to happen then?
c. Ga jij nu eigenlijk naar Werchter gaan? (vl)	Are you going to go to Werchter then?
d. Wat zullen we in de slaapkamer doen? (nl)	What will we do in the bedroom?
e. Zal ik Cindy even bellen? (nl)	Shall I phone Cindy?

2.2.5. Sentence type (polarity)

Similarly, sentence polarity (i.e. whether an utterance is affirmative or negative) has been shown to have a significant effect on the distribution of *will* versus *going to* in some varieties of English. For instance, Tagliamonte (2013) found a preference for negative *going to* in some of the British English dialects that she investigated but in others she found no significant effect. Similarly, Nesselhauf (2010) reports a favouring effect for negative *going to* in her standard British English data, whereas Szmrecsanyi (2003) found the opposite effect for the same variety. This constraint will also be tested with our Dutch data, as *gaan* regularly competes with *zullen* in negative contexts. Here, utterances were coded as 'negative' if they contained a negator such as *niet* 'not', *geen* 'no, not a', *niets* 'nothing', *nooit* 'never' or *niemand* 'no-one' (see the examples in (16)).

(16) a. Dus dat gaat toch niet gebeuren (nl)	So that's not going to happen
b. Dat ga ik ook nooit doen (nl)	I'm never going to do that

c. Ja, maar dat zal niet gebeuren (nl)	Yes, but that won't happen
d. Nou zullen we geen ander spel doen
dan (nl)	Well, we won't do any other
game then

2.2.6. Grammatical person

Poplack and Tagliamonte (2000: 335) argue that if, as Royster and Steadman (1923: 400) claim, English *going to* was originally expressive of speaker intent, one would expect it to have originally correlated most strongly with first person subjects, as intent is most commonly expressed in the first person. Therefore, any spread of *going to* to other grammatical subjects could be seen as evidence of grammaticalisation. Consider the Dutch examples below, which show *gaan* competing with *zullen* in all three grammatical persons:

(17) a. De volgende keer ga ik in ieder geval
wel een baan zoeken (nl)	Next time I am definitely going to look
for a job
b. En nu ga je een andere hobby
zoeken? (nl)	And now you're going to look for
another hobby?
c. Dus uh hij gaat waarschijnlijk toch
wel wat anders zoeken (nl)	So he is probably going to look for
something else
d. Ik zal dat zo vlug mogelijk doen (vl)	I'll do that as quickly as possible
e. Nou ja, dat zal je zien (nl)	Well yes, you'll see
f. En zij zal de was doen (vl)	And she'll do the washing

Poplack and Tagliamonte (2000) found in their Canadian study that *going to* is now favoured in non-first person contexts in many of the communities they investigated. Similarly,

Tagliamonte (2013) found that, in some British English dialects such as York, second person subjects particularly favour *going to*, while first persons favour *will*, a result that has also been reported in Torres-Cacoullos and Walker (2009) for Canadian English.

2.2.7. Co-occurrence with modal particles

Owing to the history of *zullen* as a modal verb and the observation by some researchers that utterances with *zullen* often have modal overtones (see examples (2) and (3) above), one might expect future expressions with *zullen* to show other signs of modality, for example the presence of modal particles. In Dutch, these small words are very often used in the spoken language to convey a variety of attitudes such as uncertainty, politeness and urgency: e.g. *Is uw man soms thuis?* 'Is your husband at home, **by any chance?**' (van der Wouden 2002: 21). They are very difficult to translate into English, as English does not use such words to the same extent: e.g. *Ik zal **even** kijken* 'I'll (**just**) have a look' (Shetter and Ham, 2007: 128). Modal particles have in common that they are unstressed and appear almost always in the 'middenveld': i.e. between the main verb and the final non-finite verb(s) in main clauses, or between the subject and verb in subordinate clauses. Unlike focus particles, which modify a particular part of the utterance, e.g. a noun phrase or a prepositional phrase (e.g. *Jip speelt **alleen** met autootjes* 'Jip plays **only** with cars', (van der Wouden and Caspers 2010: 53)), modal particles modify the whole utterance and can change the nuance of the message in subtle ways. For example, a common modal particle is *maar*, which is usually used to soften an utterance: e.g. *Doe je trui **maar** uit* 'Take your jumper off (**will you**)/**Just** take your jumper off.' (van der Wouden and Caspers 2010: 53).

The most commonly used modal particles in our data set are *dan* 'then', *eens* (often pronounced 'ns) 'sometimes, just', *even* (*effe*) 'just', *maar* 'just, go ahead', *nou* 'now', *toch* 'after

all', *wel* 'just, really'.¹⁵ These translations are merely rough guidelines, however, as the meanings expressed by modal particles are often not conveyed by separate words in English but by changing stress and intonation, as the examples in (18b-c) demonstrate (with stress indicated by underlining).¹⁶

(18) a. Dat ga 'k 'ns even doen (nl)	I'm (just) going to do that
b. Morgen gaat den John u wel bellen
voor nen conference call (vl)	John <u>is</u> going to ring you tomorrow
for a conference call
c. Maar dat gaat toch moeilijk gaan hè (vl)	But it <u>is</u> going to be difficult, isn't it?
d. En wat gaat ze nou doen? (nl)	And what's she going to do (now)?
e. Ik zal effe wat komkommer d'rdoor
doen (nl)	I'll (just) mix some cucumber into it
f. Ik zal hem binnenkort eens bellen (nl)	I'll (just) ring him shortly

In the current analysis, utterances were coded as either 'plus or minus particle', for the sake of simplicity, even in cases such as (18a) where there is more than one modal particle present. Despite the particles having a range of meanings they were treated as one uniform group, as the numbers of individual particles were not high enough to separate them into different groups for analysis.

3. Results

3.1. Overall frequency analysis

¹⁵ *Zullen* plus *wel* often denotes probability: e.g. *Peter zal wel slagen* 'Peter will probably pass/Peter is bound to pass' (E-ANS: 18·5·4·4·ii·f).

¹⁶ For more information, including the historical development of modal particles, see Vismans (1995).

An initial investigation of the relative frequency of *gaan* versus *zullen* in agent versus non-agent verbs (see examples (7-8) above) reveals a pattern similar to that found by Kraaikamp (2009) for Netherlandic Dutch: *gaan* occurs more frequently with verbs whose subject is an agent and *zullen* occurs more frequently with non-agent verbs (see Figure 1). This is a highly significant difference ($\chi^2(1) = 92.97$, $p < 0.001$). In Belgian Dutch, however, *gaan* is the marker of choice in both verb types, as it is generally more frequent than in Netherlandic Dutch. (If we conflate agent and non-agent verbs we see that *gaan* accounts for 60.0% of the utterances in Netherlandic Dutch and 69.8% in Belgian Dutch, which is a significant difference: $\chi^2(1) = 11.98$, $p = 0.001$). Having said that, we can see from Figure 1 that the proportion of *gaan* over *zullen* in Belgian Dutch is also significantly higher in agent verbs than in non-agent verbs ($\chi^2(1) = 61.82$, $p < 0.001$), which indicates that this verb-type constraint is operational in both varieties, despite the higher overall frequency of *gaan* in Belgian Dutch.

Figure 1: Frequency of *gaan* versus *zullen* in agent versus non-agent verbs

It has been pointed out in the literature that Belgian Dutch allows the future marker *gaan* with the lexical verb *gaan*, unlike Netherlandic Dutch (e.g. Coleman 2000, van Bree 1997, de Rooij 1985 & 1986). On the basis of this, researchers have argued that the future marker *gaan* is at a more advanced stage of grammaticalisation in Belgian Dutch than in Netherlandic Dutch (see Decroos 2000 for an overview of the discussion). Indeed, results from our data set appear to support this argument, as the proportion of *gaan* over *zullen* with lexical *gaan* is not significantly different from the proportion of *gaan* over *zullen* with non-agent verbs ($\chi^2(1) = 0.21$, $p = 0.644$). In other words, lexical *gaan*, as a non-agent verb, is patterning with other similar verbs, and there appears to be no additional constraints on its usage (see Figure 2).

Figure 2: Frequency of *gaan* with lexical *gaan* and *hebben/zijn* in Belgian Dutch

By contrast, despite the fact that the stative verbs *hebben* and *zijn* admit *gaan* as a future marker in Belgian Dutch, they show a clear preference for *zullen* (e.g. the difference between the frequency of the future marker *gaan* with lexical *gaan*, on the one hand, and its frequency with *hebben* and *zijn*, on the other, is highly significant ($\chi^2(1) = 13.10$, $p < 0.001$)). It is not entirely clear why this should be the case. Beheydt's (2005: 252) argument that the reason for the verbs *hebben* and *zijn* not allowing *gaan* as a future marker in Netherlandic Dutch is that they semantically clash with the idea of intentional actions cannot be applied here, as *gaan* in Belgian Dutch is the (albeit slightly) preferred choice of marker in the non-agent verb group, which also contains verbs with a low degree of intentionality (e.g. *worden*, *gebeuren*, *krijgen*). Thus, despite the apparently greater degree of grammaticalisation of *gaan* in Belgian Dutch, there are still some lexical constraints in operation.

3.2. Multivariate analysis

In this section, we investigate to what extent the internal constraints outlined in 2.2 above are significant in determining the preference of *gaan* over *zullen* as a future marker using mixed effects modelling. This is a statistical method that quantifies the simultaneous effects of several contextual predictors on a response variable. As we are dealing with a binary response (i.e. the choice between *gaan* and *zullen*), we used logistic regression. Specifically, the statistical package used for this was *lme4* in R, as the mixed-effects model is particularly effective in this type of investigation, as it is designed to deal with both fixed and random effects (see Tagliamonte and Baayen 2013: 171).

The fixed effects that we entered into the model are set out below. These are the internal constraints discussed in 2.1 and 2.2 above, together with the variable 'region':

- region (nl, vl)
- verb type (action, non-action, stative, *gaan*)
- clause type (main, subordinate)
- animacy of subject (animate, non-animate)
- grammatical person (first, second, third)
- proximity of future reference (proximate, non-proximate)
- sentence negative or affirmative
- sentence interrogative or declarative
- co-occurrence with a particle

In addition to these fixed effects, we also entered the random effect 'speaker' in order to ascertain whether individual speaker choice had any impact on the choice of *gaan* over *zullen*. The results are set out in Table 1.

	Odds ratio	LL	UL	Pr (>/z/)
(Intercept)	0.585	0.422	0.811	0.00132 **
region: vl	0.371	0.251	0.548	6.21e-07 ***
verb.type: gaan	3.309	1.971	5.555	5.95e-06 ***
verb.type: non-agent	4.417	3.139	6.215	< 2e-16 ***
verb.type: hebben/zijn	8.213	5.370	12.561	< 2e-16 ***
clause: subordinate	0.254	0.180	0.359	8.53e-15 ***
animacy: non-animate	1.843	1.030	3.299	0.03942 *
person: second	0.178	0.107	0.296	3.23e-11 ***
person: third	0.609	0.433	0.858	0.00452 **
proximity: proximate	0.652	0.454	0.938	0.02113 *
negative	0.975	0.689	1.380	0.88734
interrogative	1.027	0.717	1.469	0.88626
modal particle	3.174	2.451	4.110	< 2e-16 ***

LL and UL are the lower and upper 95% confidence intervals. The rightmost column shows the significance values: ‘***’ = 0, ‘**’ = 0.001, ‘*’ = 0.01, no asterisk = 0.1

Table 1: Results from multivariate analysis using lme4. Odds in terms of the use of *zullen* compared to *gaan*

Table 1 demonstrates that almost all of the variables have some significant effect on the choice of *gaan* over *zullen* in both varieties of Dutch. The only non-significant variables are to do with sentence type: interrogative versus declarative and negative versus affirmative. (The probability values are shown in the rightmost column).

The table shows the effects as odds ratios, where numbers higher than 1 indicate an increase in the odds of *zullen* being used over *gaan* and numbers less than 1 indicate a decrease. (1 indicates neither an increase nor a decrease, as the scores for negative (0.975) and interrogative (1.027) demonstrate). Here we can see the size of each individual effect: the further away from zero, in either direction, the stronger the effect. Thus, if we examine the odds ratio column of Table 1, we can see that the lowest odds of *zullen* occurring over *gaan* can be seen in subordinate clauses and with second person subjects. These appear to be particularly strong areas for the favouring of *gaan*, followed by region (Belgian Dutch), which has a slightly smaller effect.

Other contexts which disfavour *zullen* are utterances with third person subjects - this effect is smaller than that of the second person but it is still significant - and utterances with proximate future reference, which show a similar effect to third person contexts.

At the other end of the scale, the highest odds of *zullen* occurring over *gaan* are with *hebben/zijn*, non-agent verbs and lexical *gaan*. Following that, the internal constraint showing a relatively strong effect in the favouring of *zullen* is the presence of a modal particle.

In addition to the fixed effects discussed above, the random effect of individual speaker was also tested. Results show that, out of a total of 854 speakers, over two thirds produced two sentences or fewer (71.9%), which suggests that the effect of the individual speaker is likely to be negligible. Indeed, the variance of the random factor approaches zero in the analysis, meaning there is not much variability between the speakers.

In order to evaluate the model discussed above, we tested the index of concordance *C*, following Tagliamonte and Baayen (2012). This index is a generalization of the area under the receiver-operating characteristic curve in signal detection theory (Tagliamonte and Baayen 2012: 156). It measures how well the model discriminates between

the *gaan* and *zullen* responses. When C is equal to 0.5, classification performance is at chance, whereas values equal to or higher than 0.8 indicate a good performance. Figure 3 illustrates this using a ROC curve. On the plot, the model would be uninformative if the curve lay on the 45-degree line. The actual result, however, is 0.8071, which indicates a good performance.

Figure 3: Index of concordance C

4. Discussion

The quantitative variationist approach that we have taken here provides results which confirm two assertions that have been made by a number of researchers (mainly working on Netherlandic Dutch) using different methodologies: i) i.e. that *gaan* is used primarily with animate subjects; ii) that *gaan* expresses the proximate future (see e.g. de Rooij 1985 & 1986). Both of these constraints on the use of *gaan* as a future marker can be seen in terms of persistence, in the sense of Hopper (1991). *Gaan* started out as a lexical verb meaning 'to go' and could therefore only co-occur with animate subjects in its earliest history. It expressed movement along a path in order to carry out an action, which necessarily would take place within a relatively short period of time. The fact that *gaan* can now occur as a future marker with non-animate subjects and refer to future events that are non-proximate demonstrates that it has come some way down the path of grammaticalisation. However, its lexical history is still reflected in the preferences of speakers: the favouring of *zullen* in these contexts. This suggests that Dutch *gaan* is at a less advanced stage of grammaticalisation than English *going to* which, according to recent quantitative studies, is largely unconstrained by animacy (e.g. Poplack and Tagliamonte 2000), and does not significantly correlate with proximate time reference either (e.g. Torres-Cacoullous and Walker 2009, Tagliamonte 2013).

Moreover, investigations into the interaction between animacy and region, and proximity and region show that the animacy and proximity constraints are operational in both Netherlandic and Belgian Dutch, despite the higher frequency of *gaan* in the latter. It has often been argued that Belgian Dutch *gaan* is at a more advanced stage of grammaticalisation than its Netherlandic counterpart (see the discussion in Decroos 2000). Indeed, it appeared as a future marker earlier than in Netherlandic Dutch, perhaps as a result of French influence (*WNT: GAAN*, 1500 onwards, I. II, footnote 1) and it now occurs with a wider range of verbs than in

Netherlandic Dutch, including lexical *gaan*.¹⁷ Therefore, one might expect Belgian Dutch *gaan* to be less constrained by persistence effects than Netherlandic *gaan*. However, our results show no significant difference: $p = 0.13$ for the interaction between proximity and region, and $p = 0.82$ for the interaction between animacy and region.

Perhaps what is more interesting in the current analysis are the results for the constraints not previously discussed in any of the literature on Dutch *gaan*: in particular, the greater preference of *gaan* over *zullen* in subordinate clauses as opposed to main clauses and the greater preference of *zullen* over *gaan* in first person subjects as opposed to third and (especially) second person subjects. Interestingly, both of these findings have also been reported for the distribution of *going to* and *will* in many varieties of modern English (Poplack and Tagliamonte 2000, Torres-Cacoulllos and Walker 2009, Tagliamonte 2013).

There have been a number of different suggestions in the literature on English future markers as to why subordinate clauses have a favouring effect on the selection of *be going to*. Poplack and Tagliamonte (2000), following Royster and Steadman (1923: 400) state that volitional colouring seems to be reflected in the syntactic structure of the phrase, with volition stronger in main clauses and weaker in subordinate clauses. As English *will* was originally a modal verb meaning 'to want', any influence of volitional meaning on the distribution of *will* would be a case of persistence. However, this explanation does not tie in with the observation, made by a number of researchers carrying out either diachronic or apparent-time studies, (e.g. Tagliamonte 2013, Fehringer and Corrigan 2015), that this syntactic constraint is emerging over time - a development which may be somewhat surprising, as it goes against the general trend of obsolescent variants lingering on in subordinate clauses (see e.g. Bybee 2002). By contrast, in his account of *going to* versus *will*

¹⁷ There is disagreement as to how much of an influence French has played in the grammaticalisation of Belgian Dutch *gaan*. For instance, de Rooij (1986: 29) assumes that contact with French is responsible for the greater degree of grammaticalisation of *gaan* in Belgium, whereas Coleman (2000: 54-55) argues that it is more likely to be a language-internal development.

in British and American English, Szmrecsanyi (2003) prefers an explanation based on syntactic complexity and ease of processing. He suggests that because this marker is phonologically longer than the variant *will* it might buy speakers more time to process the more complex subordinate clause.

These arguments cannot be applied to our Dutch data, however, as *zullen* does not have the volitional meaning of *will* and there is no obvious processing advantage to using *gaan* + infinitive over *zullen* + infinitive.¹⁸ Torres-Cacoullos and Walker (2009: 349) state that when more than one construction develops within the same grammatical domain, there remain syntactic niches in which one or other of these is more likely to be conventionalised. Thus, it could be the case that subordinate clauses simply form a 'syntactic niche' in which *gaan* is favoured in Dutch, and that the similarity to English is coincidental. One cannot assume that patterns found in English would necessarily generalise to Dutch.

With respect to grammatical person, our results showing *gaan* being favoured in non first-person contexts mirror those found for many varieties of modern English. For instance, in six out of the eight British English dialects investigated by Tagliamonte (2013: 127) there is a strong effect of *going to* in non first-person contexts, and in her apparent-time investigation of York English she shows that there is a significant difference between the distribution of future markers in first-person versus non first-person subjects in the speech of her younger informants, as opposed to no difference in the older speakers, suggesting that *going to* is becoming specialised for non first-person contexts and *will* for first persons (Tagliamonte 2013: 132).

One reason put forward for the specialisation of *will* with first person subjects is the frequency of the collocation *I'll*, which combines the first person subject with the clitic form of *will* (e.g. Torres-Cacoullos and Walker 2009, Tagliamonte 2013). Such an explanation cannot be

¹⁸ There is also the historical argument that *going to* originated in subordinate clauses in English (Royster and Steadman 1923). However, there appears to be no evidence for this with Dutch *gaan*.

applied to *zullen*, however, as it has no clitic form. More enlightening, perhaps, is the observation made by Poplack and Tagliamonte (2000: 335) that 'attitude is most commonly expressed in the first person.' If this is the case we would expect *zullen*, which has modal overtones (and has been shown in the current analysis to significantly correlate with the use of modal particles) to be preferred over *gaan* in first person contexts. Indeed, if we take *gaan* out of the equation and investigate the frequency of modal particles with *zullen* alone, there are only two areas in which there are significant differences: grammatical person and clause type (see Table 2).¹⁹

	<i>particle</i>	<i>no part.</i>	<i>total</i>		<i>particle</i>	<i>no part.</i>	<i>total</i>
main clause	317	393	710	1st person	155	159	314
subord. clause	10	52	62	2/3 person	172	286	458

Clause: $\chi^2(1) = 17.84$, $p < 0.001$

Person: $\chi^2(1) = 10.16$, $p = 0.001$

Table 2: Co-occurrence of modal particles with *zullen*

In first-person contexts and in main clauses, the frequency of modal particles is significantly higher than in non first-person contexts and subordinate clauses respectively. This suggests that first person and main clause contexts may provide particularly favouring environments

¹⁹ The other variables show no significant differences in the frequency of modal particles: verb type (tested against agent verbs): non-agent ($\chi^2(1) = 1.11$, $p = 0.291$), *gaan* ($\chi^2(1) = 0.48$, $p = 0.491$), *hebben/zijn* ($\chi^2(1) = 1.39$, $p = 0.238$); proximity ($\chi^2(1) = 1.21$, $p = 0.272$); animacy ($\chi^2(1) = 3.46$, $p = 0.063$).

for modal expression and, as a result, *zullen* occurs with a relatively high frequency in these areas.

4. Conclusions

By examining individual internal constraints on the distribution of *gaan* and *zullen* in spoken Dutch we have been able to gain some insight into the grammaticalisation of *gaan* + infinitive as a future marker. Certain constraints have been shown to exist which could be explained in terms of the persistence of some features of the lexical history of *gaan*, namely the favouring of *gaan* with animate subjects and in contexts of proximate time reference. These persistence effects are present in both Netherlandic and Belgian Dutch, despite the higher overall frequency of *gaan* in Belgian Dutch.

Furthermore, our results reveal that, in addition to verb type, grammatical person and clause type have a highly significant effect on determining the choice of future markers in both varieties of Dutch. Interestingly, these are the same environments that have been shown to be significant in many varieties of English: the andative future variant being favoured in subordinate clauses and with non first-person subjects. However, there is no evidence to suggest that English has had any influence on Dutch in this respect. It is more likely that these distributions have to do with the functions of the Dutch future markers themselves. *Zullen* has been shown to be preferred over *gaan* when the future utterance contains a modal particle, and first-person subjects and main clauses appear to be favouring contexts for such particles. Thus, it could be that these internal constraints are linked to the increased use of *zullen* in pragmatically more nuanced contexts.

References

- Beheydt, Griet (2005). Future time reference: English and Dutch compared. In: N. Delbecque, J. van der Auwera & D. Geeraerts (eds.), *Perspectives on variation: sociolinguistic, historical, comparative*. Berlin: Mouton de Gruyter, 251-274.
- Berglund, Ylva (2000). Utilising present-day English corpora: A case study concerning expressions of future time. *ICAME Journal* 24, 25–63.
- Boogaart, Ronny (2013). De modaliteit van temporaliteit. *Nederlandse Taalkunde* 18(3), 324-338.
- Bree, Cor van (1997). 'Wat is er met *gaan* aan de hand?' In: J. A. Nijdam, M. L. Gerla, K. H. van Dalen-Oskam (eds.), *Leven in de oudgermanistiek. Jubileumnummer van het mededelingenblad van de vereniging van oudgermanisten, uitgegeven ter gelegenheid van het tienjarig bestaan van de vereniging*. Leiden: Vereniging van Oudgermanisten, 72-77.
- Bybee, Joan (2002). Main clauses are innovative, subordinate clauses are conservative: consequences for the nature of construction. In: J. Bybee & M. Noonan (eds.), *Complex sentences in grammar and discourse: essays in honor of Sandra A. Thompson*. Amsterdam: John Benjamins, 1-17.
- Bybee, Joan L., Revere D. Perkins & William Pagliuca (1994). *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Colleman, Timothy (2000). Zullen, gaan of presens? Een verkennend corpusonderzoek naar de toekomstduiders in het Belgische Nederlands. In: M. Devos, V. De Tier & J. van Keymeulen (eds.), *Nochtans was scherp van zin. Een bundel artikelen aangeboden aan Hugo Ryckeboer voor zijn 65ste verjaardag*. Gent: Vakgroep Nederlandse taalkunde Universiteit Gent, 51-64.

Corpus Gesproken Nederlands (<http://lands.let.ru.nl/cgn/ehome.htm>).

Donaldson, Bruce (1997). *Dutch: a comprehensive grammar*. London: Routledge.

E-ANS. Electronic version of Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij & M.C. van den Toorn (1997). *Algemene Nederlandse Spraakkunst*. Groningen: Martinus Nijhoff. [<http://ans.ruhosting.nl/>].

Fehringer, Carol & Karen P. Corrigan, Karen (2015). The rise of the 'going-to' future in Tyneside English: evidence for further grammaticalisation. *English World-Wide* 36(2), 198-227.

Heine, Bernd & Tania Kuteva. (2002). *World Lexicon of Grammaticalization*. Cambridge: CUP.

Hilpert, Martin (2008). *Germanic future constructions. A usage-based approach to language change*. Amsterdam: Benjamins.

Hopper, Paul J. (1991). On Some Principles of Grammaticization. In: E. C. Traugott (ed.), *Approaches to grammaticalization*. Vol. 1: *Theoretical and methodological issues*. Amsterdam: Benjamins, 17–36.

Hopper, Paul J. & Elizabeth C. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Labov, William (1972). *Sociolinguistic patterns*. Philadelphia: University of Philadelphia Press.

Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith (2009). *Change in contemporary English: a grammatical study*. Cambridge: Cambridge University Press.

Kraaikamp, Margot (2009). 'Dat gaat anders worden!' Gaan als hulpwerkwoord van de toekomstige tijd. *Onze Taal* 78. [Retrieved from: <http://www.kennislink.nl/publicaties/dat-gaat-anders-worden>].

- Nesselhauf, Nadja (2010). The development of future time expressions in Late Modern English: redistribution of forms or change in discourse? *English Language and Linguistics* 14, 163–186.
- Poplack, Shana & Sali Tagliamonte, Sali (2000). The grammaticalization of *going to* in (African American) English. *Language Variation and Change* 11, 315–342.
- Rooij, Jaap de (1985 & 1986). De toekomst in het Nederlands. Over het uitdrukken van de toekomstige tijd in standaardtaal en dialect. *Taal en Tongval, Tijdschrift voor Dialectologie*, 37, 96-123 & 38, 5-32.
- Royster, Jane F. & John M. Steadman (1923) [1968]. The *going-to* future. In *Manly Anniversary Studies in Languages and Literature*. Freeport: Books for Libraries Press, 394–403.
- Shetter, William Z. & Esther Ham (2007). *Dutch. An essential grammar*. Ninth edition. New York: Routledge.
- Szmrecsanyi, Benedikt (2003). BE GOING TO Versus WILL/SHALL. Does Syntax Matter? *Journal of English Linguistics* 31, 295–323.
- Tagliamonte, Sali (2013). *Roots of English. Exploring the history of dialects*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali & Harald C. Baayen (2012). Models, forests and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2), 135-178.
- Ten Cate, Abraham P. (1991). Bemerkungen zum deutschen und niederländischen Futur. In: E. Klein (ed.), *Betriebslinguistik und Linguistikbetrieb, Linguistisches Kolloquium* 24. Tübingen: Niemeyer, 23-31.
- Torres-Cacoullos, Rena & James A. (2009). The present of the English future: grammatical variation and collocations in discourse. *Language* 85, 321–354.

Verkuyt, Henk & Hans Broekhuis (2013). Temporaliteit en modaliteit. *Nederlandse Taalkunde* 18(3), 306-323.

Vismans, Roel (1995). Dutch modal particles. Historical evidence and functional grammar. In: B. Devriendt, L. Goossens and J. van der Auwera (eds.), *Complex Structures – A functionalist perspective*. Mouton de Gruyter. The Hague, 303-342.

Weijers, Martje. (2012). 'Gå och pensionera dig!' A comparative, corpus-based study of the Swedish verb 'att gå' and the Dutch verb 'gaan' with special emphasis on ingressive and future marking. In: H. van der Liet and M. Norde (eds.), *Language for its own sake: essays on language and literature offered to Harry Peridon*. Amsterdam: University of Amsterdam, 17-32.

Woordenboek der Nederlandsche Taal. Historisch wetenschappelijk beschrijvend woordenboek van het Nederlands van 1500 tot 1976. <http://ivdnt.org/onderzoek-a-onderwijs/lexicologie-a-lexicografie/WNT>.

Wouden, Ton van der (2002). Partikels: naar een partikelwoordenboek voor het Nederlands. *Nederlandse Taalkunde* 7, 20-43.

Wouden, Ton van der & Johanneke Caspers (2010). Nederlandse partikelbeschrijving in internationaal perspectief: waar zijn we eigenlijk en waar moet het toch naar toe? *Internationale Neerlandistiek* 48, 52-62.

Dr Carol Fehringer

Newcastle University

Newcastle upon Tyne, NE17RU, UK

carol.fehringer@ncl.ac.uk