

Learned borrowing or contact-induced change: Verb cluster word order in Early-Modern Frisian

Abstract

In West Frisian, when there are two verbs in a cluster (an auxiliary verb and a main verb) the auxiliary-final word order is considered to be prescriptively correct. However, verb clusters display much word order variation in the West Germanic languages in general and present-day Frisian speakers have been observed using auxiliary-first two-verb clusters, such as 'hie sjoen' (had seen). In Dutch, both orders have been grammatical for at least four centuries, and recently intensified language contact between these two languages is an obvious factor in this development. However, in older Frisian texts, this 'ungrammatical' auxiliary-first order has also been attested [Hoekstra \(2012\)](#). This raises a question: is the modern use of this word order really a new development taken from Dutch, or something older? To be able to tell the difference between these options, we test whether Early-Modern Frisian auxiliary-first orders have similar usage patterns as the modern Dutch ones. We have extracted verb clusters from a corpus of Early-Modern Frisian texts and compare the factors that predict usage of the auxiliary-first orders in the Middle Frisian texts to the factors that predict this order in Dutch. We conclude that these auxiliary-first verb cluster word orders used in Early-Modern Frisian texts likely came from Dutch, but through learned borrowing, while the present-day Frisian variation arose from contact through widespread bilingualism.

Keywords: language contact; historical linguistics; West-Frisian; verb clusters; multifactorial regression modeling

1 Introduction

Verb clusters display much word order variation in the West Germanic languages. In Frisian,¹ when there are two verbs in a cluster (an auxiliary

¹ In this article, we use the term *Frisian* to refer to the West-Frisian language (*Westerlauwers Fries*), the variety of Frisian spoken in the Dutch province of Fryslân, as opposed to

verb and a main verb), the order of verbs is prescriptively considered to be fixed. However, in practice it appears that both logically possible verb cluster orders are now being used in Frisian:

- (1) Anne sei dat er my **sjoen hie**.
Anne said that he me seen had
'Anne said that he had seen me'
- (2) Anne sei dat er my **hie sjoen**.
Anne said that he me had seen
'Anne said that he had seen me'

Example 1 shows the 2-1 order, so called because the syntactically higher head verb (referred to as 1) comes after the lower lexical verb (referred to as 2). Example 2 shows the 1-2 order, where the head verb comes first. This terminology can also be extended to larger clusters — a 1-2-3-4 cluster is head-first and has the main verb last. A cluster that is completely head-first is called an ascending cluster (the numbers go up), a completely head-final cluster is descending (2-1, 3-2-1 etc.).

Prescriptively, only the order in example (1) is considered normative, but Frisian speakers now use two-verb clusters in the 1-2 order as well, as in example (2). This situation is similar to modern Dutch, where the main verb can be used in both positions. The change in Frisian verb cluster order appears to be recent, and influenced by language contact with Dutch ([de Haan 1996](#)). It has even been found that Frisian bilingual children have similar word order preferences in their Frisian as in their Dutch, producing both orders ([Meyer et al. 2015](#)). However, the non-normative 1-2 order also appears in older Frisian texts. For example, in the *Oudfriese Oorkonden* (late mediaeval charters), written between 1397 and 1545, this order is used 40% of the time, in a sample of 810 verb clusters ([Larooij 1991](#)).

This raises some questions on the nature of this change in verb cluster order preferences. Is the modern use of this word order actually a new development taken from Dutch, a consequence of hundreds of years of language contact, or a continuation of an older language-internal development? And what kinds of language contact between Dutch and Frisian may have affected this change?

Saterland Frisian or North Frisian, spoken in parts of northern Germany. The West-Frisian language should also not be confused with the West Frisian dialect of Dutch, which is a Hollandic Dutch dialect influenced by the West Frisian language.

To study this, we have to consider an intermediate stage of the language, which is Early-Modern Frisian.² Studying a 17th century essay from the time period in which Early-Modern Frisian was spoken, Hoekstra (2012) found that 10% of the verb clusters in this text are in the 1-2 order, taking this as a Dutch contact effect. But can it be a contact effect when the percentage of 'Dutch-like' orders is no different than in older texts? We would expect word order preferences between the languages to converge in the case of contact effects. The percentage of 1-2 orders should then be a quantity that is between the 20-40% of the Old Frisian charters and the 70% of modern day written Dutch. Alternatively, this finding of 10% 1-2 orders in the 17th century might be a continuation of verb cluster order variation as an older Germanic phenomenon, as found in the *Oudfriese Oorkonden*. And if it is a contact effect, why are the relative frequencies of 1-2 order clusters in these early-modern texts much lower than among young bilingual speakers of Frisian in the 21st century?

We present a corpus analysis of Early-Modern Frisian verb clusters as an investigation of the nature of the variation that occurs in the texts. To account for the 1-2 orders that are used in Early-Modern Frisian texts, we are particularly interested in the contexts in which this order is used in the Early-Modern Frisian corpus, as this may help us to identify the source of this construction in Early-Modern Frisian.

We will proceed as follows. In section 2, we discuss verb cluster word order variation in Frisian more broadly, describing verb cluster order in earlier and later stages of Frisian to identify possible sources of variation. We end the section with three possible hypotheses on where the order variation found in Early-Modern Frisian texts might have come from. Next, in section 3, we discuss the nature of the variation in Dutch, which is necessary to understand any Dutch-Frisian contact effects that we might find. In section 4 we present the corpus and methodology of our study, in which we gather data on Early-Modern Frisian verb cluster orders. Results are presented in section 5, followed by a discussion in section 6 of the results in relation to the hypotheses on the source of the variation, and a conclusion in section 7.

² While 16th to 19th century Frisian is sometimes referred to as *Middle Frisian*, we will use the term *Early-Modern Frisian* in order to be consistent with the naming of related Germanic languages such as Dutch and Low German. We call the variety that was used from around the year 1400 to the 16th century *Middle Frisian*, following Versloot (2004).

2 Verb clusters in Frisian

In this section, we will identify possible sources for the verb cluster order variation in Early-Modern Frisian by examining verb clusters and language contact in Frisian more broadly. We first summarize what is known about verb cluster order in Modern Frisian, and then go back in time to verb clusters in historical varieties of Frisian in section 2.2. In section 2.3 we move to the topic of contact in historical Frisian texts by discussing the literature on texts containing mixed Middle Frisian. This literature proposes two possible contact situations that may have led to these mixed texts, which we analyse in section 2.4. Then, we move back to the topic of verb clusters and investigate how they are used in these Middle Frisian mixed texts in section 2.5. Lastly, all of this leads us to three possible sources for the use of 1-2 orders in Early-Modern Frisian texts, which are proposed in section 2.6.

2.1 Modern Frisian

Like other West-Germanic languages, Frisian expresses properties such as tense and aspect by means of auxiliary verbs. Frisian is a verb-final language (SOV), so these verbs end up clustered together at the end of the sentence in an embedded context. The auxiliary verb, which is the syntactic head of the cluster, is then typically in the final position, as in the example below (repeated from 1):

- (3) Anne sei dat er my **sjoen hie**.
 Anne said that he me seen had
 ‘Anne said that he had seen me’

If we were to base this section on verb cluster order in Frisian only on the rules listed in reference grammars, it would be a short section. Verb cluster order in Modern Frisian is typically described as fixed. According to most older speakers of Frisian, only this 2-1 verb cluster order should be used. In fact, modern reference grammars still explicitly exclude verb cluster variation in Frisian. For example, in the Frisian grammar of Popkema (2006), it is mentioned that there is order variation in the Dutch verbal end group in some cases, but that this freedom is not present in Frisian (Popkema 2006: p. 247). It is explained that only the 2-1 order should be used. In practice however, this is not the whole story. It has been shown that language contact between Frisian and Dutch has also affected the word order of verb clusters that speakers use in their Frisian: both word orders are used, as in Dutch.

The linguistic literature on modern Frisian usually describes it as a language that is affected by language contact. The variety of Frisian that has undergone contact-induced change has been called Interference Frisian (de Haan 1997). According to de Haan (1997), Frisian is undergoing *interference under full bilingualism*, as opposed to situations where a first language affects a second language, or vice versa. This is because all speakers of Frisian also speak Dutch and are thus fully bilingual.

In Interference Frisian, both 2-1 and 1-2 verb orders are used and they are 'accepted by an increasing number of Frisian speakers' (de Haan 1997: pp. 289-290). It has been noted that this is not just a case of copying the Dutch system of verb cluster ordering. Taking Frisian verb clusters of more than two verbs under consideration, Koenenman & Postma (2006) tested for the existence of hybrid word orders by means of a grammaticality judgement task. They find that hybrid orders are indeed accepted by Frisian secondary school children with two Frisian-speaking parents, yet less frequently than the orders that are typical in standard Dutch and standard Frisian. They argue that some of these hybrid orders can only exist by mixing grammatical parameters from the two standard languages. Alternatively, Hoekstra & Versloot (2016) argue that the acceptance of these hybrid orders can also be analysed as a consequence of frequency and similarity effects, without assuming grammatical parameters. In their study, the grammaticality judgement data is best modeled when Dutch and Frisian constructions are assumed to be equally frequent in the children's language input. Even though the two analyses are based on different theories of syntax, both findings suggest that modern Frisian verb cluster order preferences are shifting due to a form of language contact in which constructions from two languages are being combined.

While grammaticality judgement evidence may not be reliable for a construction with such strong normative biases as Frisian verb cluster order, there is more direct evidence as well. Meyer et al. (2015) conducted a sentence repetition study with Dutch and Frisian two-verb clusters among Frisian bilingual children to investigate this ongoing language change. They compare the children's performance on the two languages, and also compare their performance to that of monolingual Dutch children. They found a large amount of Dutch interference in the Frisian task, with the bilingual children having almost the same word order preferences as in the Dutch task. Sometimes, Dutch verbs were also used in the otherwise Frisian sentences. On the other hand, only marginal amounts of interference were found in the Dutch task: the children had a slightly greater preference for

2-1 orders than monolingual Dutch children. These results indicate that the contact-induced change in verb cluster orders has progressed quite far.

2.2 *Historical varieties of Frisian*

It does not seem to be the case that the 1-2 order was always ungrammatical until recent times. In Middle Frisian texts, written before the order preferences in Dutch started shifting towards the 1-2 order, this non-normative 1-2 order also appears. For instance, in the *Elder Skeltenariucht* from 1485, the 1-2 order is used about 10% of the time (Bor 1971). Furthermore, as mentioned in the introduction, in the *Oudfriesse Oorkonden*, written between 1397 and 1545, this order is used 40% of the time according to Larooij (1991), with a sample size of 810 verb clusters. However, Larooij (1991) does not take the Middle Dutch interference in these texts into account, which may account for some of the 1-2 orders. This interference often takes the form of formulaic opening and closing sentences that show borrowed lexical elements: in the Old Frisian charters from 1397-1460, 52% of the closing sequences contain Dutch lexical elements, such as *ons* instead of *uws* for ‘our’, though in later years this percentage decreases. If such borrowing also took place within the verb clusters that occur in the formulaic opening sequences, the borrowed clusters would have been in the 1-2 order, as this order was prevalent in Middle Dutch charter opening sequences (Boonen 2007).

In Old High German and Old English texts, verb cluster orders vary as well, from which Bloem et al. (2015) inferred that Proto-West-Germanic exhibited some degree of verb cluster order variation. Their reconstruction assumes particularly free word order options for verb clusters with ‘to have’ as an auxiliary verb, as this construction was just grammaticalizing at the time. The large degree of variation with this verb was also noted by Larooij’s (1991) study of the Old Frisian charters. These studies show that verb cluster order variation may be an older Germanic phenomenon that has disappeared in some varieties (English and 20th century Frisian) or in some cases only in the case of two-verb clusters (German).

Another reference point on the diachronic timeline of verb cluster variation is the study by Hoekstra (2012), who conducted a study of verb clusters in a 17th century essay by the Frisian author Gysbert Japicx written in Early-Modern Frisian: the largest Frisian prose text from the 17th and 18th century. Hoekstra found that 10% of the verb clusters in this text are clusters in the 1-2 order (or ascending, in the case of larger clusters), and that these

clusters exhibit some properties that typical Frisian 2-1 ordered clusters do not show. Most notably, they show the Infinitivus-pro-Participio (IPP) effect, where an infinitival verb is used in middle positions of a cluster, when a participial verb might be expected. This phenomenon is normally found in Dutch and German. Therefore, [Hoekstra \(2012\)](#) takes these 1-2 clusters as a Dutch contact effect. Other factors noted by [Hoekstra \(2012\)](#) to be associated with ascending orders are larger clusters, ‘te’-infinitive constructions, causative *dwaan* ‘to do’ as an auxiliary verb, and infinitival main verbs. [Hoekstra \(2012\)](#) argues that these factors are also associated with Dutch interference, particularly in the written modality. For example, larger clusters are more commonly ascending (i.e. an 1-2-3 order cluster), and larger clusters are more common in written Dutch than in spoken Frisian. Therefore, the use of large ascending clusters in written Frisian may be influenced by written Dutch. This link can be assumed because the author was bilingual, authored many pieces in Dutch, and had presumably learned to write in Dutch, as there was not much of a Frisian written tradition at this time.

If 1-2 order verb clusters are used in modern Frisian, in Early-Modern Frisian, and also in Middle Frisian, there is the possibility that the verb cluster order variation that is currently being observed in Interference Frisian was always there as a language-internal phenomenon. However, this idea that the 1-2 order has been in continuous use seems unlikely given the fact that it is now considered non-normative in reference grammars and according to many Frisians. There is indeed evidence to suggest that the 1-2 order was not used for some time after the Early-Modern Frisian texts were written and before approximately the 1970s or 1980s. [Wolf \(1996: p. 39\)](#) presents data from a series of studies in the early 90s, in which children in their final year of elementary school (around age 12) show 2-1 order preferences of about 40% to 60%, while their parents use the 2-1 order almost exclusively.

Some other relevant data can be found in [Pauwels’s \(1953\)](#) dialect survey, which presents the results of a questionnaire with 15 sentences containing verb clusters in subordinate clauses with various auxiliary verbs. This questionnaire was sent out for translation to speakers of local dialects throughout the Dutch language area. [Pauwels](#) concluded that the 2-1 order is nearly always used in Frisia, with exceptions mainly found in the south-east part of the province of Frisia, in which the Low Saxon Stellingwarfs dialect is spoken. Unfortunately, results were only presented by province, meaning that an exact percentage of 2-1 orders for only the Frisian-speaking part of Frisia is not available.

The *Dynamische Syntactische Atlas van de Nederlandse Dialecten* (Dynamic Syntactic Atlas of Dutch dialects, DynaSAND, [Barbiers et al. 2006](#)) is an

atlas of syntactic constructions, based on transcriptions of dialect speech collected throughout the Netherlands and Flanders, including the Frisian speaking areas. The data was collected between 2000-2002 from older speakers, mostly born before 1940. The DynaSAND does not show any 1-2 orders in Frisian-speaking areas.³ For three-verb clusters, most examples are also fully descending, though two of the four test sentences show exceptions.⁴ This provides further evidence that the 1-2 order was unlikely to be used during most of these speakers' lifetimes.

Another source during this time period before Interference Frisian appeared is a short article by [van der Meulen \(1937\)](#), who claimed that both Stedfrysk (Town Frisian) and Lânfrysk (Frisian) have 2-1 as a fixed rule, i.e. no verb cluster order variation. Stedfrysk is a set of dialects that are spoken in certain larger Frisian towns, which have a vocabulary that is more similar to Hollandic dialects and a grammar that is more similar to Frisian. Therefore, this observation is surprising — one might expect to find a Hollandic phenomenon such as 1-2 orders in these dialects. However, [van der Meulen \(1937\)](#) also notes that both orders were still used in an older Stedfrysk text, A. Jeltema's *Het vermaak der Slagterij* from 1768, with ascending orders appearing particularly in larger verb clusters.⁵ [van der Meulen \(1937\)](#) thus concludes from this that the fixed order in Stedfrysk is a fairly recent phenomenon. If 1-2 orders had disappeared even in this 'mixed language' of Stedfrysk, then it seems to us quite likely that they were also not used in Frisian at this time.

When discussing older Frisian texts, we must take the Frisian language's history of contact into account. During the time period in which the available texts were written and throughout its history, Frisian was a lesser used language, interacting with various culturally dominant languages. During the Middle Frisian period from about the 14th to the 16th century these languages were Middle Low German and Middle Dutch, and later, modern Dutch. Potential evidence for early language contact is the existence of 'mixed' Frisian texts, which contain Middle Frisian mixed with word forms from superstrate languages that were spoken in the area. Several authors have discussed these texts and what kind of language contact might have

³ In four test sentences focusing on two-verb clusters: *verteld hat* 'has told', *roepen ha* 'have called', *stoarn is* 'has died', *zien meist* 'may see'.

⁴ The test sentences *zwemme kinne moat* 'swim can must' and *zwemme gongen is* 'swim went is' show some exceptions where the auxiliary is placed first, the other test sentences are *roepen kinne hie* 'call can had' and *make ha moat* 'made have must'

⁵ Interestingly, this matches [Hoekstra's \(2012\)](#) observation that larger clusters in Gysbert Japicx's Early-Modern Frisian are also more likely to be in the ascending order.

given rise to them. In the next section 2.3 we summarize this literature, and in section 2.4 we present our view of the types of contact that these studies imply.

2.3 ‘Mixed Frisian’ texts and language contact

A particularly interesting Middle Frisian set of texts with regards to language contact are the *Basle Wedding Speeches*, initially annotated and edited by Buma (1957), who noted that the Middle Frisian in these texts is mixed with Middle Low German and Middle Dutch forms. Later, Blom (2008) also discussed some relevant sociolinguistic aspects of these texts, analysing deviant elements and comparing them to those in another mixed Frisian text from the fifteenth century, *Thet Freske Riim*. The Basle texts are interesting because they are wedding speeches, a register that is different from the legal documents that make up most of the extant Middle Frisian texts. Furthermore one of the three texts looks like it was written “in some haste” (Blom 2008: p. 1). This is an indication of the use of a less formal register.

The language used in this manuscript appears to be a clear case of ‘contact’ Middle Frisian, influenced by Middle Dutch and Middle Low German. Several authors have written about the possible nature of this contact. Bremmer (1997) argues that the writer, who calls himself Bernard of Roordahuzum (in West Frisia), is a bilingual with “a full command neither of Frisian nor Low German, certainly not in his writing, nor in all likelihood in his spoken usage” (Bremmer 1997: p. 383). He comes to this conclusion after analysing another mixed text by the author of the *Basle Wedding Speeches*, the *Life of St Hubert*, which is a Low German text with Frisian elements, rather than a Frisian text with Low German elements. Bremmer claims that the Bernard’s mixed writing reflects a language change that is ongoing at the time:

Bernard’s use of Low German is an early witness of language mixing that has resulted in the Frisian substratum underlying the present-day Low German dialects of the former Frisian speaking areas of Groningen and Ostfriesland. In Bernard’s Frisian we see a process of language shift which, especially when large numbers of speakers are concerned, marks the beginning of language loss which will eventually lead to language death (Bremmer 1997: p. 383).

This type of contact may also have resulted in the mixed language in the *Basle Wedding Speeches*, although it is unclear to us why a writer from West Frisia would be representative of language change in the more eastern areas mentioned in Bremmer's quote above.

Blom (2008) proposes an alternative explanation, supported by an older study of *Thet Freske Riim* (Campbell 1950). Campbell finds a similar language mixture there, and suggests that it is comparable to the language of the *Basle Wedding Speeches*. This suggests that "such a mixture was normal in Frisian at the time", and Campbell calls it "the curious mixed Frisian of the fifteenth century", referring to written language (Campbell 1950: p. 208). Blom (2008) finds that the two texts indeed have similar features, and takes this as evidence for the existence of a shared written register in which using borrowed forms was normal, a register that was used in late mediaeval Frisia. It is plausible that literate mediaeval Frisians were multilingual, because their works show familiarity with texts written in Middle Dutch and Middle Low German. This familiarity may have influenced their written Frisian (Blom 2008: p. 21). After all, an author or scribe was most likely to be influenced by the language in which he had been trained to write (Blom 2008: p. 14). Blom illustrates this with a quote from Bischoff (1985: p. 1263): "im allgemeinen schreibt man nicht, wie man spricht, sondern wie man zu schreiben gelernt hat" ["Generally one does not write in the way one speaks, but in the way one has learned to write"].

2.4 Language contact, acquisition and change

The proposals by Bremmer and Blom, that of extensive mixing of the languages in general and that of a mixed written register among the literate Frisians, describe two possible contact situations that might have given rise to the mixed language in the texts. In this section, we will link these contact situations to theories on language acquisition and change, in order to learn how these two types of language contact might be reflected in a text.

In the first proposal by Bremmer (1997), the contact has the form of widespread multilingualism, and in the second proposal by Blom (2008), the contact is a limited form of 'learned contact' affecting only the literate classes. These two proposals also roughly correspond to two kinds of language change that have been distinguished in the literature: change from below and change from above (Labov 1965; 1994), where 'below' and 'above' refer to the level of awareness of the phenomenon. Furthermore, they correspond to two types of language acquisition: early acquisition and

late acquisition (Weerman 2011). In a situation of widespread multilingualism, it is likely that children learn both languages when they are young, as is the case with Frisian and Dutch in modern-day Frisia — this is early acquisition. Learned contact takes place when a speaker is older and anything the speaker learns in this way can be considered late language acquisition, even if the speaker has already acquired the language early on. Late acquisition can take place in a speaker's first language (adult L1 acquisition), as well as in acquisition of a second language (L2 acquisition) (Weerman 2011). Early acquisition is characterized by (nearly) perfect acquisition of the input structures, while late acquisition may be imperfect, for example with regards to opaque aspects of paradigms or previously unknown phonemes.

In the situation described by Bremmer (1997), the Frisian spoken in the region in which the author lived would have changed 'from below': the contact phenomena might have become a part of the speakers' language systems without being marked in some way, and would have appeared in the texts as a consequence of early acquisition by the author. Bremmer (1997) notes that a Low German text by the same author in the same manuscript also contains some Frisian forms, though less than Low German forms in the Frisian *Basle Wedding Speeches*, indicating a lack of awareness of the language mixing. In the situation described by Blom (2008), the language mixture would instead come 'from above' reflecting the writing education of the author, a form of late acquisition by the author.

There are a few factors that characterize these two different forms of language change. As noted by Weerman et al. (2013), changes from above will first become apparent in formal language, while changes from below will first become apparent in informal texts. However, this factor is not very relevant when studying Middle Frisian, as all extant texts from this time period are written in a formal register. Another possible difference is that language changes from below are likely to be more internalized in the language system, as they are acquired at a younger age. If a bilingual speaker of some form of 'interference Frisian' uses both verb cluster orders, but is not aware of this because it is normal to hear both orders in their input, it is simply a part of their grammar, just as it would be for a native speaker of the source language (i.e. Dutch or Low German). The use of the two orders will be subject to the same cognitive constraints that it is subject to in the source language, and is likely to follow typical usage patterns from the source language. On the other hand, if a Frisian writer who learned to write in Dutch uses both verb cluster orders because he noticed that the 1-2 order is used in Dutch writing in addition to the 2-1 order, or uses some other construction which does not exist otherwise in the recipient language,

this will not be internalized in the same way. Instead, the 1-2 order would be used as a stylistically marked option, used only in fixed constructions that are borrowed from Dutch, or used in different contexts than in the Dutch source language. For example, an author using learned borrowings, whose language has changed ‘from above’, might not have picked up on subtle patterns such as the fact that the 1-2 order is used much more frequently with infinitival main verbs than with participial main verbs, and therefore overuse it with participial main verbs or underuse it with infinitival main verbs.

These studies of the *Basle Wedding Speeches* and other mixed texts have not addressed verb cluster order specifically. As mentioned in the introduction, for verb clusters there is another option besides these two forms of language contact: that the use of 1-2 orders is a continuation of an older language-internal development. This is not unlikely, because Old High German and Old English have both orders as well, and therefore Proto-West-Germanic was likely to have both orders. Perhaps surprisingly, in this case we also expect to see similar usage patterns as in modern Dutch: after all, these Proto-West-Germanic verb clusters would be subject to the same cognitive constraints as their modern Dutch counterparts. However, in this case we would not observe any other contact effects, such as borrowed verb cluster constructions or an association between borrowed words with Dutch origins and the (potentially borrowed) 1-2 order.

Knowing about verb cluster order variation in Middle Frisian situation in a text with clear interference effects may help to understand the subsequent Early-Modern Frisian situation. In the case of language contact, it might be expected that the 1-2 order is used more frequently in clusters where Dutch lexical forms are used, which is something that can be observed in a mixed text. Do the 1-2 orders used in the *Wedding Speeches* show signs of contact, or could they be an option left over from Proto-West-Germanic? Before investigating the contact hypotheses, to investigate the plausibility of the language-internal hypothesis, we will now present a brief corpus analysis of verb clusters in the Middle-Frisian mixed texts of the *Basle Wedding Speeches*, and possible relationships between word order and the source languages of the words.

2.5 Verb clusters in ‘Mixed Frisian’

The *Basle Wedding Speeches* consist of three texts. In these texts we found 67 verbal clusters: 64 comprised of two verbs (being either two verbs in subor-

dinate clauses or three verbs in main clauses, with one in the V2-position), 3 comprised of three verbs. For most of the verb clusters, the language could fairly easily be identified, using phonological and morphological criteria, e.g. *laten staan* ‘let stand’ is Dutch, *leta stan* ‘idem’ is Frisian, *heeft ghesproghen* ‘has spoken’ is Dutch, *spreka mey* ‘may speak’ is Frisian. For two clusters, the language remained undefined.

Text I had 34% Dutch verb clusters ($n = 29$), the other two 20% and 12% ($n = 10, 26$). However, this contrast was not statistically significant ($\chi^2 = 4.13$; $df = 2$; $p = 0.13$). This implies that there is no point in comparing the proportion between ascending and descending word order as a reflection of the overall language profile of the texts.

We also considered a potential relation between the language of the individual verbal clusters and the word order. Among the 62 two-word clusters with an identifiable language, there was no discernible contrast, with 38% ascending word orders in ‘Frisian’ clusters, against 33% in Dutch ($n = 47, 15$; Fisher’s Exact Test, $df = 1$; $p = 0.73$).

The hypothesis that there would be a relation between the linguistic origin of the language material from the perspective of phonology and morphology and the word order in verbal clusters is not supported by this data. Furthermore, the overall percentage of 38% ascending word order in two verb clusters ($n = 64$) in this text with a lot of interferences from non-Frisian varieties is not particularly high, compared to the percentages in purely Frisian texts from the same period.⁶ This makes it less likely that the use of the ascending word order in late-Mediaeval and early-Modern Frisian should be ascribed to language contact with Dutch and Low German.

We also considered the relation between word order and construction type. We considered the following constructions: aspectual (*gaet sitta* goes sit = ‘to sit down’), modal auxiliary, verbal past participle with ‘to have’ (*habba*), past participle with ‘to be’ (*wessa*), passive with ‘to become’ (*wesda*), present participle with ‘to be’ (*libben sint* living are = ‘are alive’). A χ^2 test over these six categories showed a statistically significant difference ($\chi^2 = 12.54$; $df = 5$; $p = 0.03$). However, this result may not be reliable due to missing values: there are no passive constructions or ‘to be’ present participles in the ascending order. From other research (Versloot et al. 2012; Bloem et al. 2017) it is known that these subcategories can be grouped into larger groups: modals and aspectual verbs tend to trigger

⁶ It is low in comparison with Frisian poetry and the Middle Frisian charters, but it is substantially higher than in various Middle and Early Modern prose texts. Most of the speeches consist of prose, with some scattered portions of local rhyming (without developing a real verse form).

ascending word order, past participles with ‘to have’ take an intermediate position, while past or present participles with copula may be considered of a rather adjectival nature and are more inclined towards descending word order in Dutch. The percentages of ascending word order for these three groups are: 58% ($n = 26$), 33% ($n = 18$), 15% ($n = 13$). A χ^2 test over these three categories revealed a statistically significant bias in the distribution ($\chi^2 = 8.98$; $df = 2$; $p = 0.01$).⁷

None of the three three-verb clusters found in the texts were entirely descending (*3-2-1) as in 20th century Modern West Frisian. All three clusters consisted of a finite form of a modal and *werda* ‘to become’ as the middle verb. The order was entirely ascending 1-2-3 (*moete werde vergarret* ‘has to be gathered’) one time, and in the 3-1-2 order two times. From these results, we can conclude the following:

- i. The preference for ascending word order was not correlated with the use of Dutch word forms in these texts, neither on the level of the texts, nor on the level of the individual utterances.⁸ This makes it unlikely that ascending word orders found in Middle Frisian texts appear due to language contact with Dutch;
- ii. The author of the three speeches shows an inclination towards ascending word order of 38%, which is not uncommon for Frisian texts with fewer lexical and phonological Dutch interferences of that time;
- iii. The use of the word order is strongly regulated along the same cline as in Early Modern and Present Day Dutch, with a high level of application in combination with modal verbs and aspectual verbs.

Point 1 discards specific Dutch influence in the choice for ascending word order as an expression of skilled bilingualism with different word orders for different languages (which would be a non-interference scenario). Point 2 makes it unlikely that the ascending word orders in these three texts are a specific feature of the author’s interference Frisian. Even if his language admixture was the result of an early bilingualism, his word order in verbal clusters did not differ substantially from the word order in other Frisian texts of those days, which, given their linguistic character, were written by people with much less linguistic interference. Point 3 suggests that the choice for ascending word order ran along similar lines as in Dutch,

⁷ When taking all perfect tense constructions with a past participle with either ‘to have’ or ‘to be’ together as one category, this also results in a significant skewed distribution with $p = 0.002$.

⁸ One might also consider an effect the other way around: from word order to phonological shape of the words; however, the priming effect may work two ways with the same conclusion.

where we assume that it reflects aspects of sentence processing (as proposed and discussed by [De Sutter 2009](#); [Bloem et al. 2017](#); [Bloem 2016b](#)). In case of a learned borrowing, one expects a random application or at least profoundly different rates and clines of application than in the donor language. Therefore, it is unlikely that borrowing was the decisive factor resulting in the verb cluster word orders found in this mixed text: neither full bilingualism with early acquisition, nor learned borrowing with late acquisition of the ascending verb cluster orders, but rather an older Germanic word order option that was still available at this time.

Now that we have discussed verb clusters before Early-Modern Frisian, and possible sources of the 1-2 order in Early-Modern Frisian, we can formulate several hypotheses about where the 1-2 order, and with that the variation, in Early-Modern Frisian might have come from.

2.6 Hypotheses

On the basis of the literature and data just discussed, we can sketch a timeline of verb cluster order variation in Frisian. To start, it is likely that Proto-West-Germanic had verb cluster order variation. Next, we just saw that that mediaeval Frisian texts show variation, but most likely not due to language contact at all. From [Hoekstra \(2012\)](#), we know that Early-Modern Frisian texts show variation, but not why. Lastly, from other work, we know that modern Frisian has verb cluster order variation that can be attributed to language contact with Dutch due to full bilingualism

This leaves us with a knowledge gap at Early-Modern Frisian, for which there are three possibilities:

- i. Variation in Early-Modern Frisian texts is not due to contact, but a continuation of the mediaeval situation, as in the mixed Frisian of the Basle Wedding Speeches.
- ii. Variation in Early-Modern Frisian texts is due to contact through bilingualism, with early acquisition of the optionality, similar to the modern Frisian situation and the arguments made by [Bremmer \(1997\)](#) on mixed Frisian.
- iii. Variation in Early-Modern Frisian texts is due to learned borrowing, with late acquisition of the optionality, along the lines of [Blom's \(2008\)](#) account for the mixed Frisian texts.

To find out which of these sources can account for the 1-2 orders that are used in Early-Modern Frisian texts, we are particularly interested in the contexts in which the ‘Dutch’ 1-2 cluster order is used in the Early-Modern

Frisian corpus. In our study, we test whether the Early-Modern Frisian 1-2 orders occur in the same contexts as modern Dutch 1-2 orders to see whether the use of this order is a contact effect, and what type of contact is responsible for them.

It has been argued that verb cluster order variation in Dutch has the function of facilitating sentence processing: the verb cluster order that is ‘easier’ or more economical in a particular context is used (De Sutter 2009; Bloem et al. 2017). By studying whether the variation in Early-Modern Frisian texts correlates with the same factors as the variation in modern Dutch, we can infer whether Early-Modern Frisian verb cluster order variation has the same functions as modern Dutch verb cluster order variation. This will allow us to find evidence for or against the hypothesis of language contact effects on verb cluster orders in older Frisian texts.

If Early-Modern Frisian 1-2 order clusters occur in similar contexts as modern Dutch clusters in the 1-2 order this would indicate that this order has the same function in both varieties, and is part of the grammar of the writer of the Early-Modern Frisian text. This can mean two things: Firstly, it could be the case that the order was not borrowed, but instead already existed in the language from Proto-West-Germanic on, subject to the same cognitive constraints as its modern Dutch counterpart. If this is the case, and we find no other evidence of borrowing verb cluster constructions, an older stage of the language would be the source of the variation: this would support hypothesis 1. Secondly, it could be the case that ‘contact through bilingualism’ is the source of the variation: hypothesis 2.

If the usage contexts are not similar between Early-Modern Frisian and modern Dutch, this means it is likely that the 1-2 order has been borrowed in some way, but with a different function than the function it has in modern Dutch. In this case, learned borrowing would be the source of the variation: this would support hypothesis 3.

The reason why we compare Early-Modern Frisian to Modern Dutch here, and not Early-Modern Dutch or Modern Frisian, is that most existing research on verb cluster order variation was conducted on modern Dutch. Results of corpus studies on Modern Dutch verb clusters will give us the most detailed picture of the contexts in which 1-2 orders are preferred. As many of these preferences are related to ease of processing and cognitive constraints, which are unlikely to differ per language, we will work under the assumption that the same processing mechanisms that are involved in writing modern Dutch text also apply to Early-Modern Frisian writers. Furthermore, previous work has shown that word order preferences differed depending on the type of auxiliary verb both in Middle Dutch (Coussé 2008)

and in Early-Modern Dutch (Versloot et al. 2012), much like in modern Dutch. We summarize the findings for modern Dutch in the next section.

3 Verb clusters in Dutch

3.1 Factors of variation

Dutch verb cluster order variation has been studied more thoroughly than the variation in Frisian. As noted in the introduction, for two-verb clusters, both the 1-2 and 2-1 orders may be used in modern Dutch. Diachronically, there have been changes in the prevalence of the word orders, much as in Frisian. These changes have been studied and charted by Coussé (2008) and Coupé (2015). As for Frisian, both orders can be found in the oldest Dutch texts (De Schutter 2012). Then, as in Frisian, the 2-1 order slowly became the dominant order in Dutch texts from around the years 1400 to 1600. But subsequently, unlike in Frisian, a development in the opposite direction began, and the use of the 1-2 order in Dutch texts started to increase (Coupé 2015). This increase continued during the following centuries (Coussé 2008) and appears to be ongoing — Olthof et al. (2017) showed that younger speakers in the Corpus Gesproken Nederlands (CGN, corpus of spoken Dutch) have a stronger preference for the 1-2 order than older speakers.

Synchronically, various corpus studies have addressed the choice that speakers of modern Dutch have between the two verb orders. With most auxiliary verbs, the orders are in free variation, and speakers would not make a conscious decision to use one over the other. The difference is generally assumed not to correspond to a meaning difference. However, corpus studies show that the orders are not used randomly (De Sutter 2005; Bloem et al. 2014). The use of particular orders has been shown to correlate with a range of factors, and various generalizations have been proposed to account for the relation between these factors and verb cluster word order:

3.1.1 Minimizing processing complexity

To account for some observations from his corpus study, such as the correlation between main verb frequency and word order, De Sutter (2005) proposed the idea that speakers will use the verb cluster word order that they stylistically prefer when they have the spare cognitive capacity to do so, and will use the default order when they do not have spare cognitive

capacity. Bloem et al. (2014) take this idea and suggest that it can be tested by studying whether factors indicating ease of processing correlate with one particular verb cluster word order, which would be the ‘default’ order. Subsequently, Bloem et al. (2017) found various patterns that support this assumption, arguing that one of the word orders, the 1-2 order, might be easier to process. When both orders are grammatical, the 1-2 order appears to be preferred by speakers in contexts that are complex to process for the purpose of minimizing the processing complexity of the utterance.

3.1.2 Structural priming

In a psycholinguistic experiment, Hartsuiker & Westenberg (2000) showed that verbal cluster orders can undergo structural priming — a 1-2 order was more likely to be produced when participants had been presented with this order in a priming sentence, even though different words were used. The effect was found for both the spoken and the written modality. Corpus studies also show that 1-2 orders are more likely to occur after other 1-2 orders (De Sutter 2005; Bloem et al. 2017).

3.1.3 Semantics of the main verb

Collostructional analysis studies have found that certain main verbs are used statistically significantly more often in either the 2-1 or 1-2 order in written text (De Sutter 2005; Bloem 2016a) and that certain semantic classes of verbs also appear to have such usage preferences (Bloem 2016a). For example, verbs with negative semantics are more likely to be used in the 2-1 order (Bloem 2016a).

3.1.4 Auxiliary verb class

De Schutter (1964) noted that the different auxiliary verbs have different frequencies of occurrence in the 1-2 and 2-1 order. This finding also holds when grouping instances of auxiliary verbs by their semantic class, such as modal verbs, auxiliaries of time (De Sutter 2005).

3.1.5 Information weight

Some studies have discussed evenly distributing information weight throughout a sentence (De Sutter et al. 2007) and the verb cluster (Bloem 2016b) as a factor, though this could also be considered a processing effect.

3.1.6 Rhythmic factor

Evidence has also been found that Dutch speakers move verbs in the verb cluster around to avoid stress clashes and to adhere to the rhythm of the Dutch language (De Schutter 1996; 2012).

Along with these potential sources of intraspeaker variation, there is also interspeaker variation. Speakers from different regions have different word order preferences (Pauwels 1953) and younger speakers have a stronger preference for the 1-2 order than older speakers (Olthof et al. 2017). Some other contextual factors play a role as well — the mode of communication (spoken or written), the interactivity of the discourse, the immediacy of the discourse, and stylistic preferences of the speaker or language community (De Sutter 2005).

3.2 Processing verb cluster orders

It has been argued that, in more complex contexts, the ‘easier’ verb cluster order is used (De Sutter 2009; Bloem et al. 2017). Various arguments have been made regarding which order is ‘easier’ or ‘default’ in Dutch: De Sutter (2009) discusses the 2-1 order as easier to process, while Bloem et al. (2017) conclude that it is likely to be the 1-2 order on the basis of a large corpus study. Other arguments for an 1-2 default are that it seems to be acquired earlier by children (Meyer et al. 2015) and has a more uniform information density (Bloem 2016b).

The arguments used for these claims could apply to Frisian as well, although they have not been tested. The 1-2 order is claimed to be acquired earlier than the 2-1 order by Dutch children because the 1-2 order seemingly violates verb-finality and is therefore a new construction to be learned. The most verb-like element, the finite verb, is not in the final position in a 1-2 order cluster (Meyer et al. 2015). Verb-finality is learned quite early by Dutch children, and presumably by Frisian children as well, as Frisian is also a verb-final language. Therefore, this same argument should apply to Frisian. Furthermore, cognitive constraints on processing are unlikely to differ for speakers of different, closely related languages.

If both verb cluster orders are available in the mental grammars of the authors who wrote texts in Early-Modern Frisian, and this language change occurred ‘from below’, we would expect to find similar factors affecting Early-Modern Frisian verb cluster variation: the 1-2 order is used in more complex contexts. If the 1-2 order was a learned borrowing, stylistically marked, it would not be used in this way. Bloem et al. (2017) used the

following factors as indicators of processing complexity in their study of Dutch verb clusters:

3.2.1 Syntactic priming

When a word or construction is perceived or produced it is activated, and some traces of this activation remain for a while, making it easier to re-activate them. In Dutch, verb cluster orders can syntactically prime each other (Hartsuiker & Westenberg 2000) and priming or persistence effects can be found in written text too (De Sutter 2005).

3.2.2 Morphological structure of the main verb

Dutch has separable complex verbs, which contain a particle that may appear in a different place in the sentence in some contexts. Separable verbs have been argued to be more complex to process because, unlike regular verbs, syntactic processes can operate on them in addition to morphological processes (Bloem et al. 2017). Separable complex verbs exist in Frisian too. If the 1-2 order is an available option in the writer's grammar without being stylistically marked, we would expect it to be used in the more complex condition (i.e. when the verb is separable as opposed to inseparable).

3.2.3 Length of the middle field

Verb clusters preceded by longer middle fields can be considered more difficult to process, because a longer middle field is more likely to lead to long dependency lengths and more dependencies between various parts of the middle field and the verb cluster at the end. According to resource-limitation models of language processing, this would lead to greater integration costs at the verb cluster.

3.2.4 Structural depth

Verb clusters that are further down in the syntactic tree of a sentence may be considered more difficult to process, assuming that the apparent hierarchical syntactic structure of sentences is cognitively real.

3.2.5 Information value of the last preverbal word

This factor concerns the word class of the word that is used directly before the verbal cluster. Psycholinguistic evidence suggests that closed-class words are processed differently than open-class words, however, it is not clear which word class is actually easier to process.

3.2.6 Definiteness of the last preverbal word

When a definite NP occurs before the verb cluster, it has to be linked to a previously given discourse element, which may be more complex than processing an indefinite NP.

3.2.7 Frequency of the main verb

Psycholinguistic theories generally assume that more frequent words are more easily activated, making them easier to process.

3.2.8 Extraposition

In Dutch, it is possible to extrapose prepositional objects, positioning them after the verb cluster. This is called PP-over-V and is a common construction (Willems & De Sutter 2015) and it has been hypothesized to make long middle fields easier to process (van Haeringen 1956). PP-over-V is grammatical in Frisian too.

3.2.9 Multi-word units

A multi-word unit (MWU) is a lexical unit, consisting of multiple words that together carry a different meaning than that of its constituent parts. They are processed faster and remembered better than other sequences of words.

These factors could also be studied in Early-Modern Frisian. In this paper, we will limit ourselves to factors that can be extracted from a part-of-speech tagged historical corpus automatically. In section 4.4, we discuss which of these factors can be studied in the available Early-Modern Frisian corpus data without manually combing through the corpus.

4 Data and methodology

To study Early-Modern Frisian verb cluster orders and their usage contexts, we have extracted verb clusters from a corpus of Early-Modern Frisian texts, the Integrated Language Database⁹ (Taalatabank, TDB) of the Fryske Akademy. This corpus contains all of the attested Frisian texts from the years 1550-1800. The texts have been tokenized, lemmatized and part-of-speech tagged manually. This section describes the corpus annotation, the selected texts and the process of verb cluster extraction and data pre-processing.

4.1 Texts

From the corpus, we have selected several of the longer texts and text collections for analysis. We decided to focus on authors of whom the corpus contains more material, because this is more likely to give us large samples of verb clusters for a particular author, avoiding interspeaker variation. Larger samples are advantageous because it is easier to make inferences about the data when there is more of it.

We also make a distinction between rhyming texts and prose texts, selecting some of each, including a set of rhyming texts and prose texts by the same author. This factor of text type is relevant because in rhyming texts, metre and rhyme requirements restrict the constructions that an author can use, forcing authors to use other stylistic options that they might not use in prose, such as different word orders. A previously studied example of this can be found in Old English and Old High German poetry, where verb-first (V1) declarative clauses are more common in poetry than in prose written in these languages. This is a result of emphatic verb-fronting, a stylistic technique used by Old English and Old High German poets (Kaminska 2007). Interestingly, this word order is also common in prose translated from Latin to these languages (Kaminska 2007: p. 67), indicating that this use of the V1 word order may have been borrowed from Latin, becoming a stylistic option for Old English and Old High German, at least in poetry. It might be the case that such different uses of word orders occur in Frisian poetry as well. This section lists the texts and text collections that we have selected from the TDB, along with their year of writing or publication. We also include year as a factor, as word order preferences may have shifted over time. We have not used author as a factor in our main analysis, as it

⁹ <https://argyf.fryske-akademy.eu/en/undersyk/taalkunde/yntegrearre-taalatabank/>

correlates strongly with the factor year and partially with the factor of text type.

4.1.1 Gysbert Japicx's collected work (1639-1666)

We included all Frisian texts by Gysbert Japicx that are in the Integrated Language Database (excluding shorter editions when longer editions of the same text were available).¹⁰ The texts were split by text type: the collected prose texts of Gysbert Japicx and the collected rhyming texts of Gysbert Japicx. Gysbert Japicx (1603-1666) was a highly influential poet and writer from Bolsward, a city in the west of Frisia. He wrote Dutch and Frisian poetry, as well as Frisian prose, some of which was translated from French. He also translated Latin psalms to Frisian through Dutch, and may have written poetry in Latin. Frisian-language publications were exceptional at this time, and Gysbert Japicx was responsible for the majority of published Frisian writings in the 17th century (Breuker 1989). These text collections were chosen because of their importance and large size, as well as the fact that rhyme and prose texts by the same author are available.

4.1.2 Reyner Bogerman's Frisian Sayings (1542 and 1551)

These rhyming texts¹¹ were selected because they are relatively old — it's Early-Modern Frisian, but it's still fairly close to the Frisian of the 15th century.

4.1.3 Van Hichtum's wedding poetry (1609)

These rhyming texts¹² are said to have inspired Japicx to write in Frisian.

4.1.4 Burmania Sayings (1614)

This text¹³ was selected because it is the oldest longer Early-Modern Frisian prose text in the corpus.

¹⁰ Specifically, the following text numbers in the TDB were included: 2192 of text group 2, and 3195 of text group 3, all of text group 5 except 5001, 5149 and 5002, and texts 7154 to 7189 and 7191 of text group 7.

¹¹ Texts 1542a & 1551a in the TDB.

¹² Texts 1639b and 1639cl in the TDB.

¹³ Text 1614a in the TDB.

4.1.5 Waatze Gribberts Bruyloft (1701)

This text¹⁴ was selected because it is a longer prose text, written in a register that appears to be relatively close to spoken language.

4.1.6 Aagtje IJsbrants (1779) by Eelke Meinerts

This text¹⁵ was also selected because it is a longer prose text, written in a register that appears to be relatively close to spoken language. It is of a slightly later date than *Bruyloft*.

4.1.7 Eelke Meinerts' rhyming texts (1777-1783)

These texts were selected to also include some rhyming texts from the author of Aagtje IJsbrants and from the later end of the Early-Modern Frisian period.

4.2 Annotation

The texts in the TDB have been manually annotated by an expert on a word-by-word basis (Visser 1996), mainly for lexical properties. This section summarizes what has been annotated and what has not been annotated, and how this annotation can provide information about verb clusters. We aim to use this information to identify verb clusters in the text automatically.

The annotation does not include information on the syntactic structure of sentences, such as constituent structure or dependency structure. Since verb clusters are a syntactic phenomenon, and usually defined in terms of hierarchical structures, this means that the available annotation does not tell us which verbs form a verb cluster. For example, not every sequence of verbs in the linear order of the sentence is a verb cluster — some of the verbs may belong to a different verb phrase or clause. The annotation also does not include information on clause boundaries or sentence boundaries, further adding to the difficulty of identifying which verbs form a syntactic cluster. The annotations do not include information on the type of clause either. Verb clusters in subordinate clauses work differently from verb clusters in main clauses, because the syntactically highest verb is generally positioned in the second position in the main clause (V2-effect). Word order

¹⁴ Text 1701c in the TDB.

¹⁵ Text 1779m in the TDB.

preferences in the two clause types are also known to be different in other Germanic languages, except for English and Icelandic (Askedal 2006). For these reasons, the two clause types need to be distinguished.

There are various annotations at the word level however, including lemmatization. An uninflected, modern Frisian form of each word is annotated for each token. Numbers are used to distinguish lemmas that are spelled the same. This makes it possible to aggregate over verb forms or different spellings, i.e. to know that *mot* and *môtte* are both forms of *moatte* ‘must’. Information on the form of verbs is also included. This is useful because knowing whether it is an infinitive, participial or finite verb can help in automatically determining the verb’s syntactic position in the verb cluster. Furthermore, the annotation indicates whether a verb is a main verb, auxiliary verb or copula, though this property does not seem to take syntactic position into account — if ‘to have’ is the main verb of a clause, it is still annotated as being an auxiliary verb. The annotation of verb types is fairly extensive, also including information on transitivity (for main verbs) and tense (for auxiliary verbs). Five types of auxiliary verbs are distinguished in the annotation: modal, passive, aspectual, present and future. Copular verbs are included as a different verb type, besides auxiliaries. This classification is also relevant to the study of verb clusters, as the different auxiliary verb types exhibit different word order preferences in other Germanic languages, i.e. Dutch (De Sutter 2005). Furthermore, prefixes and suffixes are separated from lemmas by a dash. This includes verb particles, making it possible to distinguish separable verbs, which exhibit different word order preferences in Dutch.

4.3 Automatic extraction

From the chosen texts, we automatically extracted (potential) verb clusters and their properties. We wrote a Python script (van Rossum & Drake 2012) that detects verb clusters in a rule-based manner, using the information available in the annotation. A verb cluster is defined roughly as follows:

- A sequence of verbs within the same ‘clause’ (sequence of text not interrupted by punctuation)
- The sequence may be interrupted by a verb particle or ‘te’ (to), which mark a *gerundium* (to-infinitive)
- The sequence must have at least one auxiliary verb and one main verb

- Auxiliary verbs are verbs that are annotated as modal verbs, auxiliaries of time, and passive verbs. They can also be aspectual verbs and copula verbs, but only when that verb is not an infinitive, te-infinitive, past participle or present participle (because then it would be a main verb).
- Main verbs are verbs that are annotated as infinitives, *to*-infinitives, past participles and present participles.

The word order of the verb cluster is then determined on the basis of the relative positions of its constituent verbs in the linear order of the sentence. It should be noted that this procedure is not 100% reliable, especially in clusters with infinitival auxiliary verbs, where auxiliary verbs and main verbs may have the same form.

4.4 Factors potentially affecting order variation

In section 3.2, we saw that some factors correlate with verb cluster word order variation in modern Dutch, and can account for this variation in terms of processing effects. While we would like to test these factors for Early-Modern Frisian too, it is not possible to extract all of them from the TDB corpus. The TDB annotation is more limited than the annotation of the Dutch corpus that was used by Bloem et al. (2017): the Dutch Lassy Large corpus also contains syntactic annotation. In this section, we will discuss whether each of the relevant factors can be extracted from the Early-Modern Frisian data in the TDB, and if so, how we operationalize them.

4.4.1 Syntactic priming

Priming may have taken place when the order of two subsequent verb clusters is the same. This factor does not involve any syntactic dependency between the two verb clusters, so we can operationalize it in the same way as Bloem et al. (2017). We record whether the previous verb cluster was in the 1-2 or 2-1 order, if there was one.

4.4.2 Morphological structure of the main verb

Just like Dutch, Frisian has separable complex verbs. In the TDB, such morphological properties are not annotated, but these verbs often appear to be written with a space or dash between the separable particle and the rest of the verb. Therefore, we take any verb containing a space or dash to

be a separable complex verb. When a separable verb is actually separated, this is annotated both on the particle and on the main verb. These cases are also counted as separable.

4.4.3 Length of the middle field

As the TDB lacks syntactic annotation, it is difficult to determine the length of the middle field of a clause before a cluster, or even of the clause. We measure the number of consecutive words without punctuation, or a line end (in the case of poetry) as an approximation of clause length, but this is likely to be inaccurate. For example, if the subordinate clause has an embedded clause, the length will only be counted from the end of the embedded clause. In section 5.1, we evaluate the accuracy of these estimates.

4.4.4 Structural depth

This measure relates to the position of a verb cluster in the syntactic tree of a sentence. This information cannot be obtained from the TDB, and we have not included this factor in the present study.

4.4.5 Information value of the last preverbal word

This factor concerns the word class of the word that is used directly before the verbal cluster. As this depends only on linear order and word class annotation, it can be extracted from the TDB. We operationalize this factor in roughly the same way as [De Sutter \(2007\)](#), in terms of three classes: **highly** informational (nouns, verbs, numerals, proper names), **intermediate** informational (adjectives and adverbs) and **low** informational (pronouns, conjunctions, prepositions, articles and interjections).

4.4.6 Definiteness of the last preverbal word

This factor cannot be operationalized accurately for the TDB data. While some morphological properties of nouns are included in the annotation, such as gender, definiteness is not annotated. And because there is no syntactic annotation such as dependency links, we cannot easily retrieve whether a preverbal noun has a determiner that indicates definiteness. Therefore, we have not included this factor in our study.

4.4.7 Frequency of the main verb

It is difficult to get an accurate estimation of general word frequencies in Early-Modern Frisian, especially since they may have varied over time. We estimated this factor by counting the frequencies of lemmas (not word forms) in all texts selected from the TDB. This estimate is likely biased due to the limited number of texts and the limited topics that these texts cover, but it is the best we can do.

4.4.8 Extraposition

Extraposition of the prepositional object, also called PP-over-V, is another source of variation that can occur around verbal clusters. It cannot reliably be detected using the annotation available in the TDB, as it depends on syntactic links between the prepositional object and the clause before the verb cluster. Therefore, we have not included this factor in the present study.

4.4.9 Multi-word units

No annotation on multi-word units, idiomatic expressions or similar information is available in the TDB, therefore we have not included this factor in the present study.

Table 1 summarizes the list of factors we could operationalize for the TDB, with frequency information for each possible value. Besides these processing-related factors, there are a few other relevant factors that are known to affect verb cluster order variation in Dutch that may be relevant here. As mentioned before, one is the type of auxiliary verb: the 1-2 order is more common with modal auxiliaries than with *hebben* ‘to have’ in Dutch, for example. Another is the use of *to*-infinitives, or *te*-infinitives in Dutch, where order preferences appear to be different. We have included these two factors in the study.

Another factor is clause type: word order preferences in main clause clusters are different than in subordinate clause clusters. In the present study, we only focus on subordinate clause clusters, as in Frisian main clauses the finite verb is placed in verb-second position, and due to the lack of syntactic annotation in the TDB we cannot determine which verbs go together. However, the annotation also does not distinguish between main clauses or subordinate clauses, so our data set is likely to include some

Factor	Levels	Frequency
TEXT TYPE	Rhyming text	377
	Prose text	1,169
TEXT YEAR	Approximate year in which text was published	—
TYPE OF AUXILIARY	Modal	692
	Copula	12
	Aspectual	43
	Passive	233
	Past tense	466
	Future tense	94
TO-INFINITIVE	To-infinitive	79
	Not a to-infinitive	1,467
PRIMING	1-2 order is primed	367
	2-1 order is primed	1,171
SEPARABLE VERB	Separable main verb	135
	Inseparable main verb	1,411
INFORMATION VALUE	High (nouns and verbs)	597
	Medium (adverbs and adjectives)	536
	Low (function words)	397
FREQUENCY OF MAIN VERB	Log frequency of the main verb in the overall corpus	—

Table 1: Summary of factors and the frequency of occurrence of their levels.

main clause clusters that look like subordinate clause clusters (for example, if there is nothing between the finite verb in V2-position and the final verb). In section 5.1, we have evaluated the extent of this problem.

5 Results

In this section we present our results regarding verb cluster word order in Early-Modern Frisian. First, we present an evaluation of the reliability of our automatically extracted data by manually checking a part of it. Then, we present the results of this automatically extracted data from the Early-Modern Frisian corpus.

5.1 Evaluation

In the previous section, we have discussed how the data was automatically extracted from the corpus, and that the annotation of the TDB corpus was not always sufficient to extract verb clusters, their word order, and their associated information unambiguously. Specifically, the script can make the following errors:

- Categorize a part of a 3-verb cluster as a 2-verb cluster
- Categorize (a part of) a main clause verb cluster or group as a 2-verb subordinate clause cluster
- Categorize a sequence of verbs that spans clause boundaries as a verb cluster
- Incorrectly determine the order of a cluster, if verbs that can be both a main verb and an auxiliary verb are involved (i.e. ‘to have’)
- Incorrectly estimate the length of the clause due to punctuation inside the clause, or lack of punctuation

Furthermore, there may be errors in the annotation. The first four error types affect the verb cluster order that the script reports, the last error types affects the clause length that the script reports. We have evaluated to what extent these errors influence the factors that we use in our study by having a Frisian linguist manually check the annotation of a random sample of 50 1-2 order clusters from Gysbert Japicx’s collected prose work and 50 2-1 order clusters from the same subcorpus. We chose to use prose text for this evaluation because sentence structure is more complex in these texts and the script is more likely to make mistakes. A pilot evaluation showed fewer errors in poetry. Furthermore, prose texts are expected to contain fewer 1-2 orders, so it is more important that the 1-2 orders that are identified, are correct. The sample is balanced for word order — it is not completely random, otherwise there would be more 2-1 orders in the sample, as the texts contain more 2-1 orders. We evaluate only for precision, not for recall (whether the script finds all the relevant clusters). That would involve reading a sample of the corpus and checking it for any verb clusters the script might have missed.

5.1.1 Word order errors

Of the 50 automatically extracted candidate 1-2 clusters, 34 were found to be actual two-verb clusters from subordinate clauses: a precision of 68%. Of the 50 2-1 clusters, all 50 met this requirement (100% precision). Most of the erroneous candidate clusters were cases of a finite auxiliary verb in V2

position in a main clause, immediately followed by the main verb in final position, with no intervening objects. This looks exactly like a 1-2 order cluster consisting of a finite auxiliary verb and a main verb at the end of a subordinate clause. Main clause clusters cannot look like 2-1 order clusters, which explains the 100% precision for the 2-1 order. This evaluation shows that a statistical model of this data is likely to overestimate the probability of 1-2 orders somewhat, at least for prose text.

5.1.2 Clause length errors

The Waatze Gribberts subcorpus contained a few verb clusters for which the length of the clause was estimated to be 484 words, due to a lack of punctuation in a section of the text. For these outliers, the length value was manually removed. All other estimates of clause lengths are 40 words or less.

For the 100 manually checked clusters described above, the correct clause lengths were manually determined. We used linear modeling to test whether there is a significant difference between the automatically determined clause lengths and the manually determined clause lengths.

The model's estimate of the mean clause length (weighted equally over manual and automatic measurements) is 8.9 words (95% confidence interval 8.3 .. 9.5 words). The estimated mean of the automatically determined lengths does not differ significantly from the estimated mean of the manually corrected lengths ($t[168] = 1.1$, $p = 0.27$). We conclude that there is no evidence for an overall difference between automatic and manually corrected lengths (estimated difference = 0.7 words; 95% confidence interval -0.6 .. 1.9 words). The estimated mean of 1-2 order cluster clause lengths also does not differ significantly from the estimated mean of 2-1 order cluster clause lengths ($t[168] = -0.1$, $p = 0.91$). We conclude that there is no evidence for an overall difference between 1-2 and 2-1 order clause lengths (estimated difference = 0.06 words; 95% confidence interval -1.3 .. 1.2 words). However, both conclusions are qualified by a significant interaction between correction and word order ($p = 0.003$): the estimated correction-related length difference (corrected minus automatic) is 3.8 words greater for 1-2 orders than for 2-1 orders. In other words, there is evidence for a cross-over interaction where the effect of manual correction is opposite for 1-2 orders compared to 2-1 orders. The script significantly overestimates the length of clauses with 2-1 clusters and underestimates the length of clauses with 1-2 clusters. We must conclude that the automatically determined clause lengths are not reliable predictors of cluster order, even

though they do not differ significantly from the manually corrected clause lengths overall. Therefore, we have excluded this factor from the model.

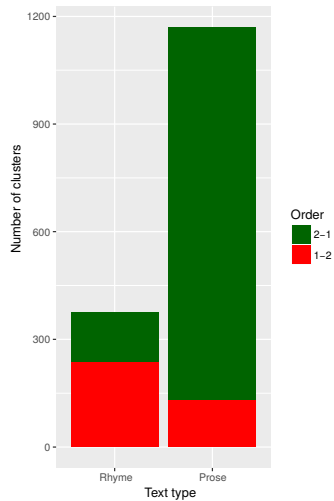
5.1.3 Annotation errors

The expert annotator found a few errors that are likely caused by annotation errors in the TDB. There were three cases where the to-infinitive status of a main verb was not detected. Furthermore, the annotator disagreed with the auxiliary type labels of two verbs. *Litte* ‘let’ is annotated as an aspectual verb, but the annotator considered it to be more like a modal verb, and *behearre* ‘should’ is annotated as a modal verb, but the annotator considered it to be more like an aspectual verb. Two of the 100 sample clusters contained *litte* and four contained *behearre*. We decided to follow the annotation on this point and not make modifications to the auxiliary verb classes in the TDB.

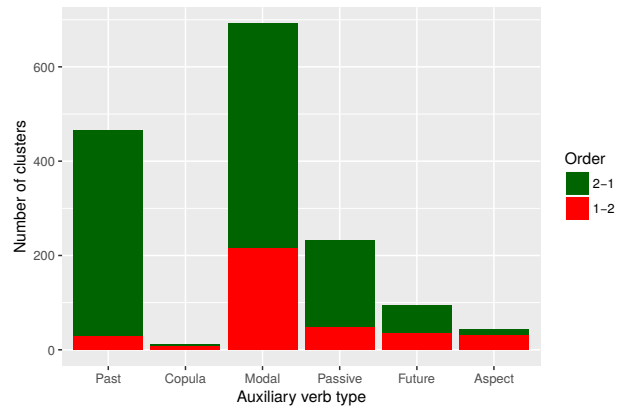
5.2 Multifactorial model of order variation

To be able to discuss the effect that different factors have on verbal cluster order variation in early-modern Frisian, we have created a multifactorial model using the factors described in section 4.4. We model verbal cluster order as a binary variable, in which the order can be 1-2 (ascending) or 2-1 (descending). This is the dependent variable. This variable is modeled in terms of the factors of the multivariate model, the independent variables, using logistic regression.

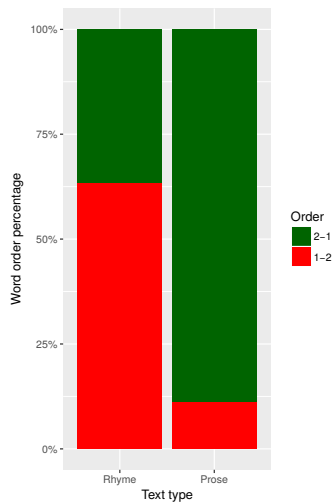
Verbal cluster order variation is one of many language variation phenomena where it has been established that multiple factors contribute significantly to the observed variation (De Sutter 2005). Such phenomena are best studied using multifactorial models, rather than testing each factor one by one. Starting with Gries (2001), such multifactorial models have been successfully used in the study of language variation. Multifactorial models allow the researcher to examine how much a particular factor contributes to the choice of a construction or word order, while controlling for the other factors. Studying several factors in isolation may cause the same variation to be attributed to multiple factors. Testing each factor while applying the necessary corrections for running multiple tests also increases the chance of type II errors, a failure to reject a false null hypothesis, and the statistical power of a multifactorial model is greater.



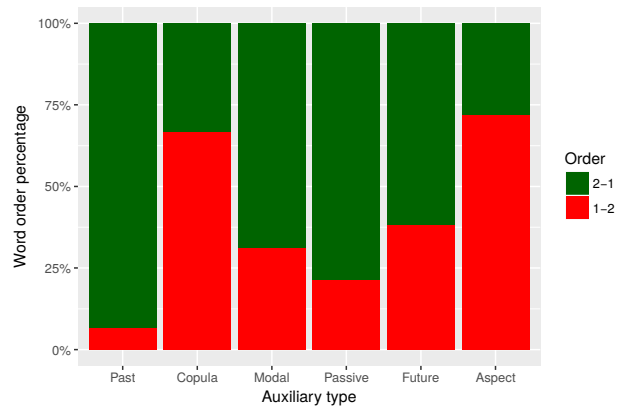
(a) Verb cluster word order by text type



(b) Verb cluster word order by type of auxiliary verb

Figure 1: Absolute counts of verb clusters per order.

(a) Verb cluster word order percentages by text type



(b) Verb cluster word order percentages by type of auxiliary verb

Figure 2: Relative percentages of verb clusters per order.

We will start with some general observations regarding the data. The 2-1 order is more frequent overall: the data set contains 1546 clusters in total, of which 1175 clusters are in the 2-1 order and 371 clusters in the 1-2 order. Since we expect more word order variation in the rhyming text type, it is interesting to look at word orders per text type. Figures 1(a) and 2(a) shows the distribution of verb cluster orders by text type, revealing that the 1-2 order is relatively and absolutely more frequent in the rhyming texts, and that there is more word order variation there.

It is also interesting to look at word orders per type of auxiliary verb, as different auxiliary verbs exhibit clear word order preferences in Dutch. Figure 1(b) shows the absolute counts and figure 2(b) shows the relative percentages.

Table 2 lists all of the explanatory variables used in the model, along with their effect size. The model uses these variables to predict the binary dependent variable of verbal cluster word order. This table allows us to explore the effect of each factor on verbal cluster order. Most of the explanatory variables are categorical, and for these variables we coded orthogonal contrasts to be able to compare the estimated effect of categories to the estimated mean effect of multiple other categories. For example, for the variable TYPE OF AUXILIARY, the effect of the category of modal auxiliary verbs is compared to the mean effect of all other categories (shown in Table 2 as +Modal compared to -FutureAspectualCopulaPastPassive). The effect size listed for the +Modal auxiliary type then indicates the odds of a 1-2 order when there is a modal auxiliary verb, rather than another, all other conditions being the same.

One might wonder why we do not compare every type of auxiliary verb to the average of all other types. This is not possible in a single statistical test, as including more contrasts than the number of levels, minus one, would introduce multicollinearity — certain auxiliary types would be in the model more than once, and these ‘copies’ would then have a correlation of 1 with each other. An assumption of regression modeling is that the factors do not correlate with each other, so this cannot be done. Orthogonal contrasts are the next best option — another option would be to compare every type of auxiliary to one particular ‘baseline’ type, but there is no one type of auxiliary that is ‘default’ or ‘basic’ in verb cluster constructions.

We ordered the orthogonal contrasts such that the auxiliary verb types which deviate the most from the average in modern Dutch are tested first. Therefore, we start by comparing modal clusters to the average, as modal auxiliaries have a very strong 1-2 order preference in Dutch (the 2-1 order is sometimes even said to be ungrammatical). The same goes for future

auxiliaries, and the less frequent aspectual auxiliaries. Next come the copular verbs, which deviate from the average in the other direction in Dutch, having a 2-1 order preference. Passive and past auxiliaries come last, as they have more or less average word order preferences in Dutch (Bloem et al. 2017). In short, the order of comparison is: Modal, Future, Aspectual, Copula, Past, Passive.

The effect size of each variable is given as an odds ratio. An odds ratio further from 1 in either direction indicates a stronger effect. An odds ratio of 1 would mean that there is no difference in the odds of a 1-2 order when a modal is used, compared to another auxiliary verb. In line with previous work, we are reporting values for the 1-2 order. As a result, odds ratios > 1 indicate an effect of association with the 1-2 word order, odds ratios < 1 indicate the 2-1 order. In our results the odds ratio for the modal contrast is 1.19, showing that in our data, clusters with modal auxiliaries are estimated to be 1.19 times more likely to be in the 1-2 order than clusters with other auxiliary verbs. This difference is not significant however, so our model provides no evidence for an effect of modal auxiliaries ($t[1545] = 0.17$; 95% confidence interval = 0.76 .. 1.84 times; p from one = 0.44). In Table 2, asterisks indicate statistical significance: One asterisk (*) indicates statistical significance at the $p < 0.05$ level, two asterisks indicate significance at $p < 0.01$, and three asterisks indicate significance at $p < 0.001$.

The variables YEAR and FREQUENCY OF MAIN VERB are not categorical, but continuous. In these cases, the odds ratios show the increase or decrease in the odds of a 1-2 order per unit (i.e. per additional year).

It should be noted that the model is likely to over-estimate the number of 1-2 orders because, as we saw in section 5.1, the automatic extraction gives us false positives for 1-2 orders but not for 2-1 orders. Since the script achieved a precision of 68% on 1-2 orders, the true odds are likely to be around 0.68 times lower than the odds estimated by the model we present here, although this does not control for correlations between specific factors and the error rate.

The results in Table 2 show that the model estimates 1-2 orders to be 18.69 times more likely in rhyming text than in prose text (95% confidence interval 12.8 .. 27.8 times). 1-2 orders are significantly more likely in rhyming text than in prose text ($t[1545] = 2.93$; $p = 5.51 \times 10^{-49}$). We conclude that Early-Modern Frisian rhyming texts are more likely to contain verb clusters in the 1-2 order than Early-Modern Frisian prose texts.

Regarding the type of auxiliary verb, we can observe several things. As mentioned, clusters with modal auxiliaries are estimated to be 1.19 times more likely to be in the 1-2 order than clusters with other auxiliary verbs,

Variable	Contrast	Odds ratio
TEXT TYPE	+ Rhyme / -Prose	*** 18.69
TEXT YEAR	Per additional year	0.998
TYPE OF AUXILIARY	+ Modal / -FutAspCopPastPassive	1.19
	+ Future / -AspCopPastPassive	0.98
	+ Aspectual / -CopPastPassive	** 7.15
	+ Copula / -PastPassive	** 7.88
	+ Past / -Passive	*** 2.50
TO-INFINITIVE	+ Yes / -No	*** 8.33
PRIMING	+ 1-2 / -2-1	0.95
SEPARABLE VERB	+ Separable / -Inseparable	0.64
INFORMATION	+ High / -LowMedium	0.91
VALUE	+ Medium / -Low	*** 0.24
FREQUENCY OF MAIN VERB	Per additional order of magnitude	0.959

Table 2: Effect of different variables on the likelihood of 1-2 verbal cluster orders.

a difference that is not significant ($t[1545] = 0.17$; 95% confidence interval = 0.76 .. 1.84 times; p from one = 0.44). Therefore, we have no evidence of an effect of the use of modal auxiliaries on verb cluster order in Early-Modern Frisian. In modern Dutch, verb clusters with modal auxiliaries have a strong 1-2 order preference (Bloem et al. 2017), and diachronically, Dutch modal verb clusters, and other clusters taking infinitival main verbs, shifted towards a 1-2 order preference earlier than verb clusters with other auxiliaries (Coussé 2008: p. 63).

We do observe a significant difference for the contrast of aspectual auxiliary verbs versus copular, passive and past ones: the model estimates aspectuals to be 7.2 times more likely to occur in the 1-2 order ($t[1545] = 1.97$; 95% confidence interval = 2.7 .. 18.8 times; $p = 5.58 \times 10^{-05}$), even though this category only occurs 43 times in the data. We conclude that aspectual verbs are more likely to occur in the 1-2 order than copular, passive and past auxiliary verbs. In Dutch, aspectual verbs behave like modal verbs in that they are almost always used in the 1-2 order, though 2-1 orders can be judged acceptable. They are also more widely used in Dutch than in Frisian.

Another significant difference here is between copular auxiliaries, and past and passive ones. The model estimates copular auxiliaries to be 7.9

times more likely to occur in the 1-2 order ($t[1545] = 2.06$; 95% confidence interval = 2 .. 36.5 times; $p = 0.0049$), although this category is even more infrequent, occurring only 12 times in the data. Nevertheless, we can conclude that copular verbs are more likely to occur in the 1-2 order than passive and past auxiliary verbs. Conversely, in Dutch, copular auxiliaries have the strongest association with the 2-1 order out of all the auxiliary verb types.

The final significant difference among the auxiliary verb types is between passive auxiliaries and auxiliaries of past tense. Passive auxiliaries are estimated to be 2.5 times more likely to occur in the 1-2 order ($t[1545] = 1.62$; 95% confidence interval = 2.5 .. 10.2 times; $p = 6.55 \times 10^{-06}$). We conclude that passive auxiliaries are more likely to occur in the 1-2 order than past auxiliary verbs. In Dutch, past tense auxiliaries (with participial main verbs) have a somewhat greater 1-2 order preference than passive auxiliaries (De Sutter 2009).

Even though the data set does not include many verb clusters with a to-infinitive (79 out of 1546 clusters), we see that the model finds a significant effect. Clusters with *to*-infinitives are estimated to be 11.5 times more likely to be in the 1-2 order than other clusters ($t[1545] = 2.12$; 95% confidence interval = 4.3 .. 16.0 times; $p = 2.06 \times 10^{-10}$). We conclude that *to*-infinitival clusters were more likely to be used in the 1-2 order in Early-Modern Frisian texts.

Lastly, we can observe a statistically significant effect of the information value of the preverbal word: when this word is medium-informative (an adverb or adjective) rather than low-informative (a function word), the cluster is estimated to be 0.24 times more likely to be in the 1-2 order, or in other words, 4.16 times more likely to be in the 2-1 order ($t[1545] = -1.44$; 95% confidence interval = 0.16 .. 0.35 times; $p = 4.04 \times 10^{-12}$). We conclude that 2-1 order clusters were more likely following medium-informative words and 1-2 clusters were more likely following function words in Early-Modern Frisian texts. In modern Dutch, both function words and medium-informative words are more associated with the 1-2 order, while high-informative words before clusters are more associated with the 2-1 order (Bloem et al. 2017).

We find no evidence for an effect of year. The direction of the model's estimate is a slight decrease in 1-2 orders over time, but this effect is not statistically significant. For the remaining factors listed in Table 2, we also found no evidence of an effect on verb cluster order in Early-Modern Frisian. In studies of modern written Dutch corpora, effects of all of these factors were found (De Sutter 2005; Bloem et al. 2017), though the fact that the

Dutch corpora that have been studied are much bigger plays a major role in this: with a larger corpus, effects of smaller size can be proven, as there will be more evidence for statistical significance.

When using such a large multifactorial model, there is the risk that the different factors are not independent, and partially explain the same portion of the variation. For example, it might be the case that modal auxiliary verbs are generally also more likely to be used in rhyming text. Statistically, we can test for this by computing the variance inflation factor (VIF) of the variables in the model. This is a measure of multicollinearity — correlation between predictor variables. In our model, all of the VIFs are very low (< 1.3 , less than 10 is considered acceptable). The factors therefore do not appear to correlate with each other, are independent, and do not statistically account for the same parts of the variation.

6 Discussion

Using the results above, we can now expand upon the three possible sources of verb cluster order variation in Early-Modern Frisian that we proposed in this article. These sources were the following:

- i. Variation in Early-Modern Frisian texts is not due to contact, but a continuation of the mediaeval situation.
- ii. Variation in Early-Modern Frisian texts is due to contact through bilingualism, with early acquisition of the optionality, similar to the modern Frisian situation and the arguments made by [Bremmer \(1997\)](#).
- iii. Variation in Early-Modern Frisian texts is due to learned borrowing, with late acquisition of the optionality, along the lines of [Blom's \(2008\)](#) account for the mixed Frisian texts.

To establish whether there was a language contact effect, and what type, we are interested in the similarities and differences to the factors that affect modern Dutch verb cluster order variation.

Firstly, our model of verb cluster order in the Early-Modern Frisian corpus showed us that there is a clear effect of text type on the use of 1-2 orders: the Dutch-like 1-2 order is used far more in rhyming text. While it can also be used in prose text, the stylistically demanding context of rhyme apparently prompted Early-Modern Frisian writers to make more use of the 1-2 option. The high usage percentages in the rhyming texts shows that the Early-Modern Frisian writers had no problem using this word order as a stylistic option, at least.

Furthermore, we observed an association between *to*-infinitives and the 1-2 order. This link was also noted by [Hoekstra \(2012\)](#), who found *to*-infinitives to be relatively frequent in the 1-2 order in a Gysbert Japicx text. Hoekstra argues that this indicates borrowing from Dutch, where *to*-infinitives are said to be head-initial (i.e. auxiliary-first). This construction may have been convenient to borrow because there was no similar construction in Frisian, and it seems like it came as a package with the 1-2 cluster order from Dutch. A similar argument could be made for our finding that aspectual verbs are more associated with the 1-2 order — they are also more commonly used in Dutch than in Frisian ([Hoekstra 2012](#)). These findings make the ‘no contact’ option a rather unlikely option.

[Hoekstra \(2012\)](#) provides the additional observation that larger verb clusters (with more than three verbs) are more likely to be in an ascending order (i.e. 1-2-3), but the sample size is small in that case, and we have not examined larger clusters in detail in the present study.

Generally, we did not observe much evidence for the processing effect that was observed for modern Dutch verb cluster order variation ([Bloem et al. 2017](#)). While we were only able to make use of a limited range of factors, we found no evidence for typical processing effects such as an increased 1-2 order preference for separable verbs in our Early-Modern Frisian data. This indicates that the Early-Modern Frisian 1-2 order did not have the function of reducing processing loads in certain contexts that it appears to have in Dutch. There was also no evidence of a priming effect, although this is difficult to measure in text — in modern written Dutch, writers are more likely to use a 1-2 order after another 1-2 order, showing that these orders are elements that are integrated well enough into the language system to undergo structural priming. We cannot say this for the Early-Modern Frisian data.

It seems more plausible that the 1-2 order is mainly a stylistic device used by these authors in the written modality, rather than a construction with the function of decreasing language processing load as in Dutch. This points at the conclusion that the contact effects we do observe are a consequence of learned borrowing, not of widespread bilingualism as in modern Frisian.

Other examples of learned borrowings and similar phenomena can be found in historical Frisian texts. [Brandsma \(1936: p. 23\)](#) notes Gysbert Japicx’s overuse of *-je*-verbs. Frisian has two classes of weak verbs. Verbs in class I end in *-e*, which corresponds to standard Dutch verb endings, and verbs in class II end in *-je*. There is no Dutch or Low Saxon analogy for this second class. [Brandsma’s \(1936\)](#) study shows that Gysbert Japicx has mixed up the two classes in his writings. There were not much more than

ten verb types of class I where he always used class I endings, while there were ‘many’ verbs that he used with different class endings — mainly class II endings on verbs that were originally class I, but some examples of the reverse were also found. It appears that Japicx overused the ending that is typical of Frisian. This might be a case of hypercorrection caused by knowledge of Dutch or Dutch writing, in order to avoid Dutch interference. It might also be a case of late acquisition (either adult L1 acquisition or L2 acquisition) in which not all the subtleties of the system were acquired by Japicx, resulting in this overgeneralization of class II endings. Similar late acquisition appears to have taken place with Japicx’s 1-2 order verb clusters.

7 Conclusion

Our data suggests that learned borrowing is the primary source of 1-2 verb cluster orders in Early-Modern Frisian texts. The lack of a function for the 1-2 order beyond stylistic usage indicates late acquisition, perhaps when the writers of these texts learned to write in another language than Frisian, as Frisian did not have much of a written tradition. In this case, the subtleties of the usage patterns of the 1-2 order in Dutch were either not fully acquired or not fully transferred into their Frisian. At the same time, there are clear signs of language contact, as some verb cluster constructions (such as the ones with *to*-infinitives) appear to have been borrowed from Dutch.

In the older, language-mixed Basle Wedding Speeches however, we did not find signs of borrowing of the 1-2 order, as there was no correlation between 1-2 verb cluster word order and the use of Dutch word forms. This indicated to us that Middle Frisian retained the 1-2 order option from Proto-West-Germanic. We can therefore also not exclude that this language-internal source still played a role in Early-Modern Frisian, given the fact that the authors from the early-Modern period apparently did not consider the use of 1-2 orders to be ungrammatical in their writing. However, in early 20th century Frisian the 1-2 order appears to be completely ungrammatical, so the present day usage of the 1-2 order cannot be a remnant of Proto-West-Germanic, but is rather a consequence of widespread bilingualism among the Frisian-speaking population, which is evidenced by the fact that it has been observed in children as young as 4 to 6 years old (Meyer et al. 2015).

We conclude that these 1-2 verb cluster word orders used in Early-Modern Frisian texts likely came from Dutch, but do not have the same

underlying cause as the modern Dutch word orders. The contact-induced change that led to the present modern Frisian situation appears to be recent.

Competing interests

The author(s) has/have no competing interests to declare.

References

- Askedal, John Ole. 2006. Some typological differences between the modern Germanic languages in a historical and geographical perspective. *Tidsskrift for Sprogforskning* 4(1). 187–207. <https://doi.org/10.7146/tfs.v4i1.317>. <https://tidsskrift.dk/tfs/article/view/317>.
- Barbiers, Sjef, H Bennis, G De Vogelaer, M Devos & M van der Ham. 2006. Dynamische syntactische atlas van de Nederlandse dialecten (DynaSAND). Amsterdam, Meertens Instituut. URL: <http://www.meertens.knaw.nl/sand>.
- Bischoff, Karl. 1985. Reflexe gesprochener Sprache im Mittelniederdeutschen. In Werner Besch, Oskar Reichmann & Stefan Sonderegger (eds.), *Sprachgeschichte. ein handbuch zur geschichte der deutschen sprache und ihrer erforschung.*, vol. 2, 1263–1268. De Gruyter Mouton.
- Bloem, Jelke. 2016a. Lexical preferences in Dutch verbal cluster ordering. In Kate Bellamy, Elena Karvovskaya, Martin Kohlberger & George Saad (eds.), *ConSOLE XXIII: Proceedings of the 23rd Conference of the Student Organization of Linguistics in Europe*. 70–93. Leiden: Leiden University Centre for Linguistics. <http://media.leidenuniv.nl/legacy/5-bloem.pdf>.
- Bloem, Jelke. 2016b. Testing the processing hypothesis of word order variation using a probabilistic language model. In Dominique Brunato, Felice Dell’Orletta, Giulia Venturi, Thomas François & Philippe Blache (eds.), *Proceedings of the workshop on computational linguistics for linguistic complexity (cl4lc)*. 174–185.
- Bloem, Jelke, Arjen Versloot & Fred Weerman. 2014. Applying automatically parsed corpora to the study of language variation. In Jan Hajic & Junichi Tsujii (eds.), *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. 1974–1984. Dublin: Dublin City University and Association for Computational Linguistics. <http://www.aclweb.org/anthology/C14-1186>.
- Bloem, Jelke, Arjen Versloot & Fred Weerman. 2015. An agent-based model of a historical word order change. In *Proceedings of the Sixth Workshop*

- on *Cognitive Aspects of Computational Language Learning*. 22–27. Lisbon, Portugal: Association for Computational Linguistics. <http://aclweb.org/anthology/W15-2404>.
- Bloem, Jelke, Arjen Versloot & Fred Weerman. 2017. Verbal cluster order and processing complexity. In Enoch Aboh (ed.), *Complexity in human languages: A multifaceted approach*, vol. 60, 94–119. Elsevier. <http://dx.doi.org/10.1016/j.langsci.2016.10.009>.
- Blom, Alderik H. 2008. Language admixture in the Old West Frisian Basle Wedding Speeches? *Amsterdamer Beiträge zur älteren Germanistik* 64(1). 1–27.
- Boonen, Ute K. 2007. Syntactische variatie in 13de-en 14de-eeuwse oorkonden: Rood en groen in de adresformule. *Handelingen van de koninklijke commissie voor Toponymie & Dialectologie* 79. 165–194.
- Bor, Arie. 1971. *Word-groups in the language of the skeltana riucht: a syntactic analysis with occasional lexicological observations; followed by an inquiry into its punctuation and the possibility of the influence on the text of spoken language*. H. Veenman & Zonen.
- Brandsma, Wytze Lammert. 1936. *Het werkwoord bij gysbert japix*: University of Groningen dissertation.
- Bremmer, Rolf H. 1997. Bad Frisian and bad Low German: Interference in the writings of a medieval West Frisian. *Multilingua* 16. 375–388.
- Breuker, Philippus H. 1989. *It wurk fan gysbert japix: pts. 1-2. oerlevering en ûntstean*: Vrije Universiteit Amsterdam dissertation.
- Buma, Wybren Jan. 1957. *Aldfryske houlikstaspraken*. van Gorcum.
- Campbell, Alistair. 1950. Review of Buma, *Houlikstaspraken*. *IB* 21. 207–208.
- Coupé, Griet. 2015. *Syntactic extension. the historical development of Dutch verb clusters*: Radboud University Nijmegen dissertation. <http://repository.ubn.ru.nl/handle/2066/141109>.
- Coussé, Evie. 2008. *Motivaties voor volgordevariatie. Een diachrone studie van werkwoordvolgorde in het Nederlands.*: Ghent University dissertation. <http://dx.doi.org/1854/12253>.
- de Haan, Ger. 1996. Recent changes in the verbal complex of frisian. In A Petersen & H F Nielsen (eds.), *A frisian and germanic miscellany published in honour of nils Århammar on his sixty-fifth birthday, 7 august 1996*, 171 – 184. Odense University Press.
- de Haan, Germen J. 1997. Contact-induced changes in modern West Frisian. *Us Wurk* 46. 61–89.
- De Schutter, G. 1964. De dubbele werkwoordgroep in het Nederlands. *Studia germanica gandensia* 45–84.

- De Schutter, G. 1996. De volgorde in tweeledige werkwoordelijke eindgroepen met voltooid deelwoord in spreek-en schrijftaal. *Nederlandse taalkunde* 1. 207–220.
- De Schutter, Georges. 2012. De werkwoordelijke eindgroep en nog steeds geen einde? *Verslagen & Mededelingen van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde* 122(1). 1–38.
- De Sutter, Gert. 2005. *Rood, groen, corpus! Een taalgebruiksgebaseerde analyse van woordvolgordevariatie in tweeledige werkwoordelijke eindgroepen*: University of Leuven dissertation. <http://hdl.handle.net/1979/98>.
- De Sutter, Gert. 2007. Naar een corpusgebaseerde, cognitief-functionele verklaring van de woordvolgordevariatie in tweeledige werkwoordelijke eindgroepen. *Nederlandse Taalkunde* 12(4). 302–330.
- De Sutter, Gert. 2009. Towards a multivariate model of grammar: The case of word order variation in Dutch clause final verb clusters. In A Dufter, J Fleischer & G Seiler (eds.), *Describing and modeling variation in grammar*, 225–255. Walter De Gruyter. <http://hdl.handle.net/1854/LU-537795>.
- De Sutter, Gert, Dirk Speelman & Dirk Geeraerts. 2007. Luisteren schrijvers naar hun innerlijke stem? De invloed van ritmische factoren op de woordvolgorde in geschreven werkwoordelijke eindgroepen. *Neerlandistiek.nl* 7. 1–24.
- Gries, Stefan T. 2001. A multifactorial analysis of syntactic variation: Particle movement revisited. *Journal of quantitative linguistics* 8(1). 33–50. <http://dx.doi.org/10.1076/jqul.8.1.33.4092>.
- Hartsuiker, Robert J & Casper Westenberg. 2000. Word order priming in written and spoken sentence production. *Cognition* 75(2). B27–B39. [http://dx.doi.org/10.1016/S0010-0277\(99\)00080-3](http://dx.doi.org/10.1016/S0010-0277(99)00080-3).
- Hoekstra, E. 2012. Reade wurdfolchoarders en dêrmei gearhingjende aspekten yn 17e-ieusk Frysk. In ûndersyk nei de tiidwurdkloft yn Gysbert Japicx syn "Yen suwnerlinge forhânlinge Fen it Libben In fenne Deade". *It Beaken* 72 (2010)(3-4). 223–239.
- Hoekstra, Eric & Arjen Versloot. 2016. Three-verb clusters in Interference Frisian: A stochastic model over sequential syntactic input. *Language and Speech* 59(1). 43–58. <https://doi.org/10.1177/0023830915577009>. <http://las.sagepub.com/content/59/1/43.abstract>.
- Kaminska, Anna. 2007. The word order of Old English and Old High German non-conjoined declarative clauses in different text types. *Studia Anglica Posnaniensia: International Review of English Studies* 43. 49–74.
- Koeneman, Olaf & Eilien Postma. 2006. Veranderingen in Friese werkwoordsclusters. *Nederlandse Taalkunde* 11(2). 124–145. https://pure.knaw.nl/portal/files/464064/OlafKoenemanVeranderingen1088_001.pdf.

- Labov, W. 1994. *Principles of linguistic change, Volume 1: Internal Factors*. Oxford: Blackwell.
- Labov, William. 1965. On the mechanism of linguistic change. *Georgetown University Monographs on Language and Linguistics* 18. 91–114. https://repository.library.georgetown.edu/bitstream/handle/10822/555458/GURT_1965.pdf;sequence=1#page=105.
- Larooij, Pieter. 1991. *Undersyk nei de posysje fan it helptiidwurd in de bysin yn Oudfryske Oorkonden, Bewerkt door O. Vries, Vierde Deel.*: University of Groningen MA thesis.
- Meyer, Caitlin, Doatske de Haan, Martina Faber & Fred Weerman. 2015. Embracing ascending orders: Verb clusters in bilingual acquisition. Presented at Generative Approaches to Language Acquisition, Université de Nantes.
- Olthof, Marieke, Maud Westendorp, Jelke Bloem & Fred Weerman. 2017. Synchronic variation and diachronic change in Dutch two-verb clusters. *Tijdschrift voor Nederlandse Taal- en Letterkunde* 133(1). 34–60.
- Pauwels, Anita. 1953. *De plaats van hulpwerkwoord, verleden deelwoord en infinitief in de Nederlandse bijzin*. M. & L. Symons, Leuven.
- Popkema, Jan. 2006. *Grammatica Fries*. Prisma Woordenboeken en Taaluitgaven.
- van der Meulen, P. 1937. Iets over woordschikking in het Stadfriesch. *Onze Taaltuin* 5. 337–341.
- van Haeringen, Coenraad Bernardus. 1956. *Nederlands tussen Duits en Engels*. Servire Den Haag.
- van Rossum, Guido & Fred L Drake. 2012. *The Python language reference, release 2.7.2*. Python Software Foundation Amsterdam, Netherlands. www.python.com.
- Versloot, Arjen, Hanneke Hengst & Marjan Huisman. 2012. Werkwoordvolgordes in het Amsterdams van de 17e eeuw. *Taal en Tongval* 64(1). 41–71.
- Versloot, Arjen P. 2004. Why old frisian is still quite old. *Folia Linguistica Historica* 38(Historica vol. 25, 1-2). 253–304.
- Visser, Willem. 1996. *Ta it bruken fan it Midfryske korpus fan de Taaldatabank fan de Fryske Akademy*. Fryske Akademy.
- Weerman, Fred. 2011. Diachronic change: Early versus late acquisition. *Bilingualism: Language and Cognition* 14(2). 149–151. <https://doi.org/10.1017/S1366728910000581>.
- Weerman, Fred, Mike Olson & Robert Cloutier. 2013. Synchronic variation and loss of case: Formal and informal language in a dutch corpus of 17th-century amsterdam texts. *Diachronica* 30(3). 353–381. <https://doi.org/10.1017/S0012267312000058>.

- [//doi.org/doi:10.1075/dia.30.3.03wee](http://doi.org/doi:10.1075/dia.30.3.03wee). <http://www.ingentaconnect.com/content/jbp/dia/2013/00000030/00000003/art00003>.
- Willems, Annelore & Gert De Sutter. 2015. Reassessing the effect of the complexity principle on PP Placement in Dutch. *Nederlandse Taalkunde* 20(3). 339–367. <http://dx.doi.org/10.5117/NEDTAA2015.3.WILL>.
- Wolf, Henk. 1996. *Wat en hoe yn it Ynterferinsjefrysk: In singroan en dia-groan ûndersyk nei feroarings yn de Fryske tiidwurdlike einrige*: University of Utrecht MA thesis.