A quantitative-theoretical analysis of syntactic microvariation: Word order in Dutch verb clusters


LANGUAGE

Jeroen van Craenenbroeck, Marjo van Koppen, Antal van den Bosch

Language, Volume 95, Number 2, June 2019, pp. 333-370 (Article)

Published by Linguistic Society of America
DOI: *https://doi.org/10.1353/lan.2019.0033*

A QUANTITATIVE-THEORETICAL ANALYSIS OF SYNTACTIC MICROVARIATION: WORD ORDER IN DUTCH VERB CLUSTERS

Jeroen van Craenenbroeck
KU Leuven, *CRISSP*, and *KNAW Meertens Instituut*

Marjo van Koppen
KNAW Meertens Instituut, *Utrecht University*, and *UiL-OTS*

Antal van den Bosch
KNAW Meertens Instituut

This article is a case study in how quantitative-statistical and formal-theoretical (generative) approaches to language variation can be combined. We provide a quantitative analysis of word-order variation in verb clusters in 185 dialects of Dutch and map the results of that analysis against linguistic parameters extracted from the theoretical literature on verb clusters. Based on this novel methodology, we argue that verb cluster ordering in Dutch dialects can be reduced to three grammatical parameters (largely similar to the ones described in Barbiers et al. 2018), and we identify the dialect groups that correspond to the various settings of those parameters.*

Keywords: verb clusters, Dutch dialects, linguistic variation, generative grammar, correspondence analysis, *k*-nearest neighbors classification

1. Introduction. Typological studies into linguistic phenomena typically reveal a bewildering amount of variation and flexibility on the one hand, combined with seemingly universal rigidity on the other. For instance, when we consider the possible orderings of demonstratives, numerals, adjectives, and nouns inside the noun phrase, the languages of the world use no fewer than fourteen different orders as their neutral word order, but at the same time this means that of the twenty-four (four factorial) theoretically possible orders, ten are universally unattested (Greenberg 1963, Cinque 2005, Abels & Neeleman 2012). The job of the comparative linguist, then, is to separate that which is fixed and necessary—the principles, in generative parlance—from that which is variable and contingent—the parameters. Within the theory of generative grammar, the notion of 'parameter' has seen various incarnations, ranging from binary choices in the workings of the core grammatical system (e.g. Baker 2001) over the lexical properties of functional items (e.g. Borer 1984, Chomsky 1995) to properties of the LF- and (in particular) PF-interface (e.g. Kayne 2005, Chomsky 2007).

Kayne (2000) has argued that a detailed comparison of large numbers of closely related languages or language varieties presents a powerful research tool for uncovering linguistic parameters, the idea being that such a comparison is the closest real-world alternative to a controlled laboratory experiment: one tries to keep all orthogonal variation constant so as to be able to examine the effect of minute changes to the phenomenon under investigation. In this article we follow Kayne's lead by examining in detail a

* We would like to thank Klaus Abels, Harald Baayen, Sjef Barbiers, Hans Bennis, Jonathan Bobaljik, Lotte Dros-Hendriks, Frans Hinskens, François Husson, Dany Jaspers, Göz Kaufmann, Richard Kayne, Ad Neeleman, John Nerbonne, Marc van Oostendorp, Cora Pots, Ian Roberts, Koen Roelandt, Martin Salzmann, Jolijn Sonnaert, Dirk Speelman, Gert De Sutter, Benedikt Szmrecsanyi, Tanja Temmerman, Guido Vanden Wyngaerd, Eline Zenner, the members of the Maps and Grammar project, and audiences at Mapping Methods (Tartu, May 2014), What Happened to Principles & Parameters? (Arezzo, July 2014), Methods in Dialectology XV (Groningen, August 2014), Maps and Grammar (Amsterdam, September 2014), Dialect Syntax: The State of the Art (Frankfurt am Main, December 2014), INALCO (Paris, February 2015), GLOW 38 (Paris, April 2015), CGSW 30 (Chicago, May 2015), Dealing with Bad Data in Linguistic Theory (Amsterdam, March 2016), CRISSP 10 (Brussels, December 2016), and the summer school Limits of Variability (Potsdam, June 2018) for discussion, comments, and references. Many thanks also to the referees and editors of *Language* for their extensive but always constructive comments and suggestions, and to Jan-Pieter Kunst for providing us access to the raw SAND data.

specific linguistic phenomenon, word order in clause-final verb clusters, in 185 dialects of Dutch. We argue that we can distill grammatical parameters from this large data set by looking for statistical patterns in the data and mapping those against the insights gleaned from the theoretical literature on verb clusters. The resulting picture is one in which quantitative-statistical and formal-theoretical (in particular, generative) approaches to linguistics go hand in hand, mutually benefiting from one another (see Merlo 2015 for an approach that is similar in spirit).

The article is organized as follows. The next section provides some general background for readers unfamiliar with the Dutch dialect landscape and/or the phenomenon of verb clusters. Section 3 introduces the data that form the basis for the analysis. As will become clear, even if we restrict ourselves to a relatively confined empirical domain such as verb cluster ordering, the amount of attested variation is substantial and a statistical approach quickly becomes appealing. This methodology, introduced and detailed in §4, combines quantitative-statistical methods (correspondence analysis in particular; cf. Greenacre 2007) with qualitative-theoretical analyses of the generative type. Section 5 presents the results of this analysis. We show that the variation in Dutch dialect verb cluster ordering can be reduced to two main dimensions, and we indicate for each of those dimensions how well it aligns with the theoretical literature on verb clusters. Section 6 then interprets these results from a formal-theoretical point of view. We propose a parametric account of verb cluster ordering (based heavily on the analysis in Barbiers et al. 2018) and show how the abstract parameter settings and dialect types predicted by this account can be detected in the actual data set, while at the same time taking into account additional sources of variation, such as priming effects or geographical proximity. Section 7 concludes.

2. Background: dutch dialects and verb clusters. This introductory section provides some general background on the Dutch dialect landscape as well as the phenomenon of verb clusters. A basic understanding of both of these topics will be useful for the reader to be able to interpret and situate the results obtained in the following sections.

2.1. The dutch dialect landscape. The Dutch language area in Europe comprises the Netherlands, the northern half of Belgium, and the northern tip of France. While the entire area—with the exception of the small French part—is roofed by the same standard language, Standard Dutch, there is considerable variation in the varieties of Dutch that are spoken. Not surprisingly, then, both Belgium and the Netherlands have a rich and long tradition of descriptive dialectological work. One of the most influential and authoritative ways of representing the Dutch dialect landscape is the map from Daan & Blok 1969, adapted in Figure 1.[1]

Daan and Blok based this map on perceived linguistic differences, that is, on intuitions from dialect speakers and dialectologists about which dialects were similar to one another and which ones were different (see Spruit 2008:15–19 for more background on this map and on Dutch dialectological maps in general). In this map we can discern roughly ten major dialect areas of Dutch. Starting in the southwest, the sand-colored region represents West Flemish. It is spoken in the north of France, the Belgian province of West Flanders and parts of East Flanders, and in parts of Zeelandic Flanders in the

[1] Note that this figure and several of the others are presented in full color in the electronic versions of this article, but in black and white in the print version; the color versions of the figures are also available open access at http://muse.jhu.edu/resolve/71.

FIGURE 1. The main Dutch dialect regions, adapted from Daan & Blok 1969.

Netherlands. It shows substantial similarities to the Zeelandic dialect spoken in the rest of the Dutch province of Zeeland (the yellow islands to the north of the West Flemish region). To the right of West Flemish, in a slightly darker hue, is East Flemish, which is spoken in parts of East Flanders and parts of Zeelandic Flanders. Further to the right is Brabantic (dark orange), which is spoken in parts of East Flanders, Flemish Brabant, Antwerp, and Belgian Limburg. The red area is the Limburgian dialects, which cover most of the Belgian and Dutch provinces of Limburg. The pinkish area to the north of Brabantic and Limburgian represents North Brabantic, which corresponds by and large to the province of North Brabant in the Netherlands. The light green areas in the center of the Netherlands represent the Hollandic dialects. They are flanked by North Hollandic in the north (the yellow/beige region) and the Saxon dialects in the east and northeast (various shades of dark green). Finally, there is Frisian (the blue area in the north). One can argue about whether Frisian is a language of its own rather than a dialect of Dutch, given that there is ‘a codified and legally recognized standard variety of Frisian, which is also in limited use in formal and public domains’ (Hinskens & Taeldeman 2013:3, citing Willemyns 2006:1762). Given that the distinction between a language and a dialect is a fairly arbitrary one to begin with (Chambers & Trudgill 1998: Ch. 1) and given that the Frisian dialects show interesting variation with respect to the topic under investigation here (word order in clause-final verb clusters), we include Frisian and its dialects in our analysis in the remainder of the article.

This concludes our overview of the Dutch dialect landscape. The regions described here should help the reader make more sense of the geographical interpretation of our analysis in §6.3. Given the introductory nature of the present section, our overview was necessarily brief. For more in-depth discussion of the regions outlined in Fig. 1, we refer the reader to Hinskens & Taeldeman 2013.

2.2. Verb clusters. Verbs tend to cluster at the right-hand side of the clause in Dutch and other (mainly) West Germanic languages. Moreover, such verb clusters typically display a fair amount of word-order variation. Consider in this respect the two-verb cluster in 1.

(1) a. dat hij heeft gelachen.
that he has laughed
b. dat hij gelachen heeft.
that he laughed has
'that he has laughed.'

The perfect auxiliary *heeft* 'has' can either precede (1a) or follow (1b) the past participle it selects, in this case *gelachen* 'laughed', thus leading to two possible cluster orders. Similarly, when the verb cluster consists of three verbs, even more orders—though not all; see §3 below—become possible. This is illustrated in 2.

(2) a. dat hij moet hebben gelachen.
that he must have laughed
b. dat hij moet gelachen hebben.
that he must laughed have
c. dat hij gelachen moet hebben.
that he laughed must have
'that he must have laughed.'

Over the years, the type of word-order variation illustrated in 1 and 2 has been the topic of extensive investigation in the dialectological, usage-based, and formal-theoretical literature (see Bennis & Coussé 2012, on which the following comments are based, for an overview and references). The dialectological interest in this phenomenon dates back to 1949 (van den Berg 1949) and was initially primarily focused on two-verb clusters, more specifically two-verb clusters of the type illustrated in 1, consisting of a perfect auxiliary and a participle. It was in this tradition that the terms 'red order' and 'green order' were coined to refer to the orders in 1a and 1b, respectively. The colors referred rather arbitrarily to the color scheme used in the maps drawn by Pauwels (1953), and in the prescriptive literature that later ensued, the red order confusingly referred to the preferred option, and the green one to the dispreferred option. In later dialectological work, the empirical focus was extended to include three-verb clusters; see in particular Stroop 1970. That interest in this phenomenon has remained high is witnessed by the fact that in a recent dialect atlas, which forms the empirical basis for the analysis in the following sections, an entire chapter is devoted to verb clusters (Barbiers et al. 2008:14–25).

The usage-based literature has—to the best of our knowledge—focused almost exclusively on two-verb clusters. The prime representative of this line of research is De Sutter (2009), who performs a logistic regression analysis on a corpus study of two-verb clusters of the type in 1, in which he examines the effect on cluster order of explanatory variables such as frequency of the participle, length of the middle field, information value of the last preverbal word, and so forth. The factors that come out in De Sutter's research as most significant in determining the cluster order are the choice of the auxiliary ('have' vs. 'be'), the morphology of the main verb (in particular, whether the verb has a separable particle), and syntactic persistence (essentially a priming effect: if the previous cluster had a red order, there is an increased chance that the next one will too). Other notable work in this tradition includes Coussé 2008 and Bloem et al. 2017.

The formal-theoretical interest in verb cluster order can be traced back to Evers 1975. Much of the discussion has centered around the question of how to derive the various orders in 1 and 2, and whether these facts are informative as to the underlying base order of the Dutch VP. Evers started out from a head-final base order, in which the green order in 1b is basic, and the red one in 1a is derived from it by moving the participle to the right (so-called VERB RAISING). Later accounts reversed that line of thinking and took 1a to be basic, with 1b derived from it through movement of the participle to the left (so-called VP-intraposition; see in particular Zwart 1993). The issue remains far from settled to this day; see Wurmbrand 2017 for extensive discussion and references.

This concludes our brief overview of the existing literature on verb clusters. The research we present in the following sections is the first to bring together various strands from the accounts just mentioned: it focuses both on two- and on three-verb clusters, it combines the quantitative rigor of the usage-based approaches with the analytical depth of the formal-theoretical accounts, and it shares with the dialectological literature an interest in focusing on geographically determined cooccurrence patterns. In the next section we introduce the data that form the basis of our analysis.

3. THE DATA: WORD-ORDER VARIATION IN DUTCH VERB CLUSTERS. We begin our discussion of the data by revisiting the examples in 1, repeated here as 3.

(3) a. dat hij heeft gelachen.
that he has laughed
b. dat hij gelachen heeft.
that he laughed has
'that he has laughed.'

These two examples mean exactly the same thing, and for most—if not all—speakers of Standard Dutch the choice between them is more or less optional. Let us pretend, however, for the sake of argument, that the data in 3 stem from two different dialects, with dialect A allowing only the auxiliary-participle order and dialect B only the opposite one. One could then postulate a parameter that captures this difference and argue that dialect A and dialect B have a different setting for this parameter. For instance, starting out from a head-initial base order (in which the auxiliary precedes the participle), one could argue that in dialect B the participle has moved to the left across the auxiliary, while in dialect A it has stayed put. This would yield the following parameter setting.

(4) a. dialect A: [−MoveParticipleToTheLeftOfAux]
b. dialect B: [+MoveParticipleToTheLeftOfAux]

Needless to say, actual linguistic data are never as clear-cut or black-and-white as this hypothetical example. In order for the reader to appreciate this, it suffices to include three-verb clusters in the discussion. An example is given in 5.

(5) Ik vind dat iedereen moet kunnen zwemmen.
I find that everyone must can swim
'I think everyone should be able to swim.'

The main verb *zwemmen* 'swim' is selected by the modal *kunnen* 'can', which is in turn selected by *moet* 'must'. All three verbs cluster at the end of the clause, with the linear order reflecting the selectional hierarchy: the most deeply embedded verb is also rightmost in the cluster. As is customary in the literature on verb clusters, we use number combinations to refer to the various cluster orders. The cluster in 5, for example, displays a 1-2-3 order, whereby '3' refers to the most deeply embedded verb of this three-verb cluster (i.e. *zwemmen* 'swim'), '2' refers to *kunnen* 'can', and '1' to *moet* 'must'. In three-verb clusters, there are six (three factorial) theoretically possible orders. However, a large-scale dialect investigation of 267 Dutch dialects in Belgium, France, and the Netherlands (the SAND project; see Barbiers et al. 2005 and Barbiers et al. 2008) has revealed that for the cluster type illustrated in 5—that is, modal-modal-infinitive—only four of those six orders are attested.

(6) a. Ik vind dat iedereen moet kunnen zwemmen. (✓1-2-3)
b. Ik vind dat iedereen moet zwemmen kunnen. (✓1-3-2)
c. Ik vind dat iedereen zwemmen moet kunnen. (✓3-1-2)
d. Ik vind dat iedereen zwemmen kunnen moet. (✓3-2-1)
e. *Ik vind dat iedereen kunnen zwemmen moet. (*2-3-1)
f. *Ik vind dat iedereen kunnen moet zwemmen. (*2-1-3)

Moreover, it is not the case that in every one of those 267 dialects the orders in 6a–6d are well-formed. Quite the contrary: there is a substantial amount of variation when it comes to which dialect allows which subset of these four cluster orders. For example, while in the dialect of Midsland (illustrated in 7) only 1-3-2 and 3-2-1 are well-formed, Langelo Dutch (shown in 8) allows only for 1-2-3 and 3-1-2.

(7) Midsland Dutch
a. *dat elkeen mot kanne zwemme. (*1-2-3)
that everyone must can swim
'that everyone should be able to swim.'
b. dat elkeen mot zwemme kanne. (✓1-3-2)
c. *dat elkeen zwemme mot kanne. (*3-1-2)
d. dat elkeen zwemme kanne mot. (✓3-2-1)
e. *dat elkeen kanne zwemme mot. (*2-3-1)
f. *dat elkeen kanne mot zwemme. (*2-1-3)

(8) Langelo Dutch
a. dat iedereen mot kunnen zwemmen. (✓1-2-3)
that everyone must can swim
'that everyone should be able to swim.'
b. *dat iedereen mot zwemmen kunnen. (*1-3-2)
c. dat iedereen zwemmen mot kunnen. (✓3-1-2)
d. *dat iedereen zwemmen kunnen mot. (*3-2-1)
e. *dat iedereen kunnen zwemmen mot. (*2-3-1)
f. *dat iedereen kunnen mot zwemmen. (*2-1-3)

More generally, there are sixteen (two to the fourth power) possible subsets or combinations of word orders that a dialect can select from 6a–6d.[2] Of those sixteen options, twelve are attested in the SAND data. They are listed in Table 1, each accompanied by a sample dialect in which this particular combination occurs.

It should be clear that a data pattern such as this is not straightforwardly amenable to the type of (overly) simple parameter account outlined above. Looking at the combinations in Table 1, it is not obvious which parameters are responsible for this variation or even how to go about trying to identify those parameters. Things get even worse when we expand our empirical viewpoint further and consider ALL cluster orders that were

[2] There are only fifteen if we exclude the option whereby none of the orders is allowed in the dialect in question. That would be a dialect in which three-verb clusters of the type modal-modal-infinitive simply do not occur. As far as we know, no such dialect exists in Dutch.

SAMPLE DIALECT	1-2-3	1-3-2	3-2-1	3-1-2
Beetgum	✓	✓	✓	✓
Hippolytushoef	✓	✓	✓	*
Schermerhorn	✓	✓	*	✓
Warffum	✓	✓	*	*
Visvliet	✓	*	✓	✓
Kollum	✓	*	✓	*
Langelo	✓	*	*	✓
Oosterend	✓	*	*	*
Midsland	*	✓	✓	*
Waskemeer	*	✓	*	*
Bakkeveen	*	*	✓	✓
Lies	*	*	✓	*

TABLE 1. Word-order combinations in modal-modal-infinitive clusters in the SAND dialects.

part of the SAND questionnaire. There were a total of eight questions in the questionnaire that dealt exclusively with verb cluster order, which are briefly described in 9.

(9) a. three questions about two-verb clusters of the type auxiliary-participle
b. one question about two-verb clusters of the type modal-infinitive
c. one question about three-verb clusters of the type modal-modal-infinitive
d. one question about three-verb clusters of the type modal-auxiliary-participle
e. one question about three-verb clusters of the type auxiliary-motion verb-infinitive
f. one question about three-verb clusters of the type auxiliary-modal-infinitive

The cluster types in 9a and 9c were already illustrated above (see examples 3 and 5, respectively). For the four remaining types we provide representative examples in 10.

(10) a. dat jij het niet **mag zien**.
that you it not may see
'that you are not allowed to see it.' (modal-infinitive)
b. dat hij haar **moet hebben gezien**.
that he her must have seen
'that he must have seen her.' (modal-auxiliary-participle)
c. dat hij **is gaan zwemmen**.
that he is go swim
'that he went for a swim.' (auxiliary-motion verb-infinitive)
d. dat hij mij **had kunnen roepen**.
that he me had can call
'that he could have called me.' (auxiliary-modal-infinitive)

Together, the eight questions listed in 9 yielded a total of thirty-one cluster orders. If we now list, for each of the 267 SAND dialects, which dialect has which combination of those thirty-one cluster orders, we arrive at 137 different patterns of verb cluster orders. Put differently, when considering the data on verb cluster orders from the SAND questionnaire, we can discern 137 different dialect types.[3]

What does this mean for parameter theory? In its most extreme form, the theory of parameters would posit that any and every observable morphosyntactic difference be-

[3] The data table contains a number of gaps (see §4.2 below for detailed discussion). Those NAs were not taken into consideration when counting the number of word-order patterns, which means that 137 is a conservative estimate.

tween two languages should be reducible to a different setting for at least one parameter. Applied to the case at hand, this would mean that each of the 137 dialect types differs from all of the others in at least one parameter setting. Alternatively, it could be that some of the variation found in the SAND database is not due to linguistic parameters, in the sense that it reflects the effects of, for example, dialect mixing or contact, influence from the (normative) standard language, or even speech errors. What theoretical linguists try to do, then, is to determine which parts of the variation are due to the grammatical system, and which parts are not. Barbiers (2005) takes this approach in his analysis of verb clusters. With respect to the modal-modal-infinitive clusters introduced in 6, he proposes that the grammar rules out the 2-3-1 and the 2-1-3 orders, but that the four remaining orders are grammatical in all varieties of Dutch. Any interdialectal differences in the acceptability of these orders—such as the contrast between Midsland Dutch and Langelo Dutch in 7 versus 8—is then due to 'sociolinguistic factors' such as 'geographical and social norms as well as considerations of register and context' (Barbiers 2005:234–35).

In this article we follow the same general principle as Barbiers—that is, we assume that only part of the variation found in the SAND data should be derived from the grammatical system—but with a different methodology and a different conclusion. We show how a statistical analysis of the SAND verb cluster data can be mapped against the findings from the formal-theoretical literature on this phenomenon and conclude that a substantial portion of the variation can be accounted for through the interaction between three linguistic parameters. The next section describes this methodology in more detail.

4. Methodology.

4.1. Introduction. This section is organized as follows. In the next subsection we describe how the raw data from the SAND questionnaires were preprocessed so as to make them amenable to a statistical analysis. Section 4.3 outlines the difference between active (locational) and supplementary (linguistic) variables and makes clear how the latter were extracted from the theoretical literature and how they were operationalized. Section 4.4 describes the correspondence analysis we performed on the data set, and the results of the analysis are presented in §5.

4.2. Preparation of the data. All data discussed and analyzed in this article come from the SAND project. As pointed out above, this four-year dialect atlas project (2000–2004) investigated the variety of Dutch spoken in 267 dialect locations in Belgium, France, and the Netherlands. It has yielded two atlases: Barbiers et al. 2005 and Barbiers et al. 2008. The SAND data stem from three sources: a written questionnaire, a series of oral dialect interviews, and an additional set of telephone interviews (see Cornips & Jongenburger 2001 for a detailed description of the SAND methodology). The analysis carried out in this article uses the raw data from the oral dialect interviews, which also form the basis for the maps in Barbiers et al. 2005 and Barbiers et al. 2008. Before we introduce the data in more detail, it is worth commenting on how they were collected, that is, how the interviews were carried out. In controlling for various extralinguistic factors, the SAND methodology was designed to maximize the chances of detecting and identifying the linguistic variables of dialectal variation. To give but a couple of examples (see Barbiers 2009:1609 for a more detailed discussion): all informants belonged to the same age group (fifty-five to seventy years old) and socioeconomic class (no higher education, lower middle class), all of them were interviewed by a speaker of their dialect, and all of them used the dialect in at least one public domain and had lived in the same location all their lives, as had their parents. In building in such restrictions, the SAND researchers wanted to reduce the odds that the variation that was found was due to sociolinguistic variables, dialect mixing, accommodation toward the (normative and prestigious) standard language, and so forth. As such, these factors contribute to making the SAND data set ideally suited for the type of research envisioned in this article, which focuses specifically on the grammatical parameters of language variation.

The SAND data files contain a list of all the data points contained in the two atlases, including information about the type of phenomenon under investigation, the number of the map and atlas, the element listed in the key of the map, and of course the dialect location where the phenomenon was attested. We have converted these data into a 31 × 267 matrix, in which each verb cluster order occupies a row and each dialect location a column. Cells are filled with '1' when that cluster order is attested in that dialect location, '0' when it is absent, and 'NA' in case the data point is missing. Cluster orders are identified by their English glosses, so as to make the data tables and graphs more readable for a non-Dutch-speaking audience. A small sample of this data table—the upper left-hand corner—can be seen in Table 2.

	MIDSLAND	LIES	WEST-TERSCHELLING	OOSTEREND	…
IS_DIED	0	0	0	NA	…
DIED_IS	1	1	1	NA	…
HAS_TOLD	0	0	0	0	…
TOLD_HAS	1	1	1	1	…
HAVE_CALLED	0	0	0	0	…
CALLED_HAVE	1	1	1	1	…
MAY_SEE	0	0	1	0	…
SEE_MAY	1	1	1	1	…
CAN_SWIM_MUST	0	0	0	0	…
MUST_CAN_SWIM	0	0	0	1	…
MUST_SWIM_CAN	1	0	0	0	…
…	. . .	. . .	. . .	. . .	…

TABLE 2. Upper left-hand corner of the raw data table.

The first row of this table contains data pertaining to 'is died': a two-verb cluster consisting of a singular form of 'to be' used as a perfect auxiliary followed by a participle, that is, a 1-2 order. As the values in the subsequent cells indicate, this order is not attested in the varieties spoken in Midsland, Lies, and West-Terschelling, and there is no data for Oosterend. The second line provides the data for the 2-1 order of that same cluster, and the remaining rows provide similar information for other verb clusters and their word orders.

Three additional aspects of the data preparation warrant further comment. First, as can be seen in Figure 2, the cluster orders show substantial differences in terms of their frequencies: while some occur in nearly all of the 267 dialect locations, others are limited to only a handful of places. In order to ensure that the rare orders were not spurious, we manually checked the interview data for the orders that yielded fewer than five occurrences—six orders in total. This included going back to both the written transcriptions and the sound recordings of the interviews (both of which can be accessed via the online version of the SAND database; Barbiers et al. 2006). Based on this, we excluded three cluster orders from the data set, as they turned out to be false positives: the two instances of a 2-1-3 order in the case of auxiliary-motion verb-infinitive (cf. 10c), the single instance of a 2-3-1 order in the case of modal-modal-infinitive (cf. 5), and finally the two instances of a 3-1-2 order in the case of auxiliary-modal-infinitive (cf. 10d). The other three rare orders showed no abnormalities, either in their transcriptions or in their sound recordings, and accordingly, they were retained in the data set. This means that the final data set contains twenty-eight cluster orders.[4]

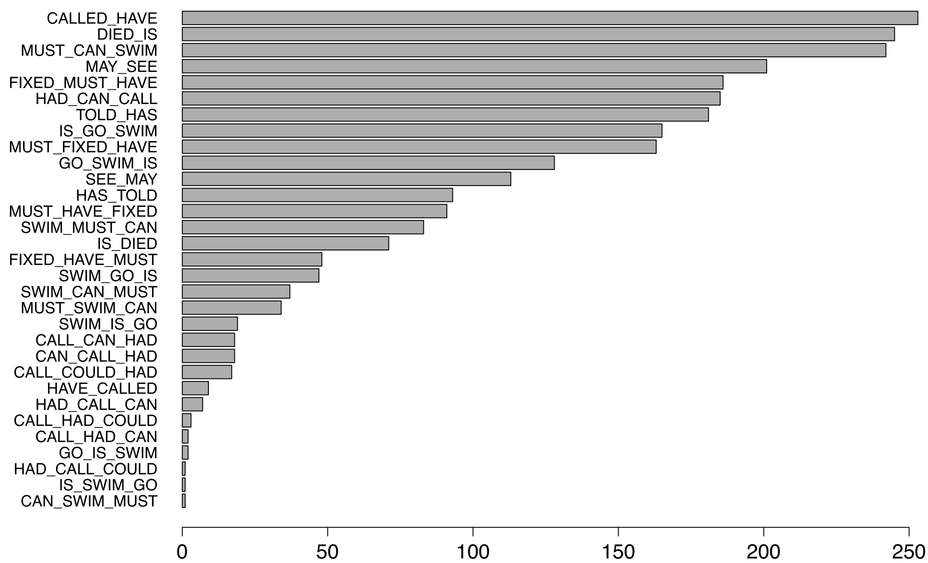

FIGURE 2. Frequency of the thirty-one verb cluster orders found in the SAND data.

The second aspect of the data-preparation stage that deserves some further discussion concerns the question methodology used in the oral dialect interviews of the SAND project. The data pertaining to verb clusters are based on two types of questions: translation tasks and elicitation questions. In the former, the informants were given a Standard Dutch sentence and were asked to translate it into their dialect, while the latter involved (prerecorded) oral versions of dialect sentences for which they had to provide a grammaticality judgment (see Cornips & Jongenburger 2001 for a more detailed description of the various question methodologies used in the SAND project).[5] For verb clusters this means that while in elicitation questions every possible cluster order was presented to the informants and explicitly judged by them,[6] in translation tasks they were presented with a single order in Standard Dutch, which they translated into their dialect with the same order, a different order, or—in the case of multiple responses—a combination of orders. Accordingly, some of the zeros in Table 2 are based on informants explicitly rejecting a particular order, while others are a reflection of the absence of this order in the informants' translation of the Standard Dutch sentence that was offered to them. For the remainder of this section as well as the next, we ignore this methodological complication, but we return to the difference between elicitation and translation in §6.3.

[4] We also ran the analysis without excising these false positives, and the results were highly comparable. In other words, the overall effect of this data-cleaning operation might turn out to be minimal.

[5] Of the twenty-eight cluster orders in the data set, fifteen are based on translation tasks, corresponding to three of the cluster types listed in 9, namely 9a, 9b, and 9f.

[6] This is not strictly true for all clusters: the extensive written questionnaire and dialect literature review that preceded the dialect interviews had turned up systematic gaps in verb cluster ordering. Orders that were unattested in any variety of Dutch were not included in the oral interviews, not even in elicitation questions. A well-known example is the 2-1-3 order. See Barbiers 2005 for a detailed overview.

The third and final methodological point concerns the NAs in Table 2, which indicate that some data points are missing. There are several possible reasons for this: sometimes a question was not included in a particular interview, sometimes informants gave an irrelevant translation, and so forth. In total, there are 480 cells that contain NA, which is 6.42% of a total of 7,476 (= 28 × 267) cells. Missing values occur in eighty-two of the 267 dialect locations, spread out evenly across the entire language area. Given that the correspondence analysis described in §4.4 cannot be applied to a data table containing missing values, we performed the analysis on only those dialects for which a full data set was available. This means that the final data table contains twenty-eight cluster orders and 185 dialect locations.[7]

4.3. ACTIVE AND SUPPLEMENTARY VARIABLES. As was pointed out in §3, the goal of this article is to determine which—if any—combination of grammatical parameters can best account for the observed word-order variation in Dutch verb clusters. In order to do so, we need to include in the analysis insights and results from the theoretical literature on verb clusters. In this subsection we describe how these theoretical analyses can be operationalized for a quantitative analysis.

The first step involves decomposing theoretical accounts into their constitutive parts. We illustrate how this works using the analysis by Barbiers, Bennis, and Dros-Hendriks (2018; henceforth BBD).[8] A central assumption in their account is that verb clusters are built via iterative application of the operation Merge. More specifically, no movement (neither head nor XP) is involved in the word-order variation attested in verb clusters. Moreover, given that Merge as a syntactic operation does not impose any linear order, variation in cluster order is due (at least in part) to postsyntactic linearization requirements. One such requirement, which is also the first parameter proposed by BBD, is the following.

(11) A dialect is uniformly {descending/ascending} in the linearization of verbs.

This principle presents dialects with a binary choice: either the linear order directly mimics the selection order, or it mirrors it. In the first case, all verb clusters are strictly ascending: that is, 1-2 and 1-2-3, while in the second case they are all descending, yielding 2-1 and 3-2-1 as the only orders.[9] Recall from the discussion in §3, however, that strictly ascending and strictly descending orders are not the only ones that are attested in Dutch dialects. BBD derive the additional orders by introducing parameters related to the categorial status of the elements making up the cluster.

First, they start out from the well-known observation that past participles have not only a verbal but also an adjectival use in Dutch. This is supported by their ability to occur in the attributive position of noun phrases.

(12) het gelezen boek
the read book
'the book that was read'

If this categorial ambiguity persists in contexts involving verb clusters, it makes an interesting prediction with respect to the linearization of such structures. Given that in all

[7] We also performed the analysis on a full 28 × 267 data table, but with imputed values for the missing data. This analysis is described in Appendix C.

[8] This analysis is not chosen at random: as will become clear in §§5 and 6, the Barbiers et al. 2018 approach is highly successful in accounting for word-order variation in Dutch verb clusters. In other words, this account reappears in later sections of the article.

[9] Here and throughout this article we are setting aside clusters containing four or more verbs, as such data were not part of the SAND questionnaires. See Abels 2016 for discussion.

varieties of Dutch nonverbal complements cannot follow the verb (see 13), adjectival versions of past participles are predicted to necessarily precede the (other) verbs in the cluster.

(13) a. dat Jan ziek is.
that Jan sick is
'that Jan is sick.'
b. *dat Jan is ziek.
that Jan is sick
intended: 'that Jan is sick.'

Let us see how this applies to an example like 14, a case of a 3-1-2 order in a modal-auxiliary-participle cluster. BBD's analysis of this cluster can be schematically represented as in 15: in spite of first appearances, this example contains not a three- but a two-verb cluster. Those two verbs, *moet* 'must' and *worden* 'become', are linearized in an ascending order that mimics their selection order. In addition, the example contains the adjectivally used participle *geholpen* 'helped', which is linearized to the left of the verb cluster, as is standard for adjectives; see 13. In a way, then, the 3-1-2 order is spurious: there is a two-verb cluster with a 1-2 order, but the adjectival participle preceding the cluster gives the whole sequence the overt appearance of a 3-1-2 order.

(14) dat Jan geholpen moet worden.
that Jan helped must become
'that Jan should be helped.'

(15) geholpen$_A$ moet$_{V_1}$ worden$_{V_2}$

This analysis thus accounts for all cases in which a past participle precedes the verbal cluster in spite of the parameter in 11 being set to 'ascending'. BBD formulate their second parameter as follows.[10]

(16) A dialect {does/does not} have verbal participles.

BBD employ a similar line of reasoning when it comes to infinitives unexpectedly occurring at the front of the cluster. They start from the basic observation that bare infinitives can be productively nominalized in Dutch. An example is given in 17.

(17) Ik vond het zwemmen erg leuk.
I found the swim.INF very nice
'I really enjoyed the swimming.'

If this categorial ambiguity carries over to verb clusters, then by the reasoning developed above, we expect infinitives to be able to precede otherwise strictly ascending (or descending; see n. 10) verb clusters. Consider in this respect the example in 18 and BBD's analysis of it in 19.

(18) dat iedereen zwemmen moet kunnen.
that everybody swim must can
'that everyone should be able to swim.'

(19) zwemmen$_N$ moet$_{V_1}$ kunnen$_{V_2}$

Accordingly, BBD formulate the parameter regulating the occurrence of these cluster orders as follows.

[10] Note, as is also pointed out by BBD, that the parameter in 16 is arguably also operative in dialects with a descending cluster order. The problem is that in such dialects a different setting for the parameter in 16 does not lead to a difference in surface order: given that a past participle is always the most deeply embedded verb, and given that this verb is always cluster-initial in descending dialects, there is no discernible difference between a cluster with a verbal participle and one with an adjectival predicate.

(20) A dialect {does/does not} have nominalized infinitives in verb cluster constructions.

The final ingredient of BBD's account concerns 1-3-2 orders. They follow from none of the parameters proposed so far: on the one hand, 1-3-2 is not a strictly ascending or descending order, while on the other, the main verb does not precede the entire cluster like an adjective or a noun would. Instead, BBD argue that these are cases of cluster interruption, whereby nonverbal material appears between the elements of a verbal cluster. This phenomenon is well known from the literature on verb-particle stranding; what BBD do is extend this account to include adjectival participles.[11] For the example in 21, this yields the representation in 22.

(21) dat Jan moet geholpen worden.
that Jan must helped become
'that Jan should be helped.'

(22) moet$_{V_1}$ geholpen$_A$ worden$_{V_2}$

Although BBD do not explicitly formulate a fourth parameter, it is clear that there is (at least potentially) an additional point of variation here, namely, the question of whether a dialect allows for cluster interruption of the type illustrated in 22.

This concludes our summary of the Barbiers et al. 2018 analysis of verb cluster ordering in Dutch: it starts out from a uniformly ascending or uniformly descending linearization order, and derives additional orders via the categorial ambiguity of participles (A or V) and infinitives (N or V), and via the option of cluster interruption. Note that we can reformulate this account (and the variation predicted by it) as a number of simple binary parameters. They are summed up in 23.

(23) a. **Asc:** Is the order compatible with an ascending linearization?
b. **Desc:** Is the order compatible with a descending linearization?
c. **VerbPart:** Does the order involve a verbal participle?
d. **NomInf:** Does the order involve a nominalized infinitive?
e. **ClustInterr:** Does the order involve cluster interruption?

Each of these parameters splits verb cluster orders into mutually exclusive subsets. More specifically, the twenty-eight cluster orders from the SAND data set can be encoded in terms of the linguistic parameters in 23. This means that Table 2 can now be extended with columns representing not geographical but theoretical information. This is illustrated for part of the data set in Table 3.

In total, we have added sixty-four linguistic variables to the data table, representing not just the analysis in Barbiers et al. 2018, but also those in Haegeman & van Riemsdijk 1986, Schmid & Vogel 2004,[12] Barbiers 2005, Barbiers & Bennis 2010, Bader 2012, and Abels 2016. Moreover, we have included a head-initial head movement analysis, a head-final head movement analysis, a head-initial XP movement analysis, and a head-final XP movement analysis, all as described in Wurmbrand 2017. Finally, we added nine additional variables that are not tied to a specific analysis.

[11] As BBD point out, the geographical distribution of the 1-3-2 order speaks in favor of this analysis: it occurs in roughly the same area where traditional cases of cluster interruption are also found. Note that in principle this account can also be extended to 1-3-2 orders that involve an infinitive as the most deeply embedded verb. BBD end up not taking this route, however, and propose that this order is 'a transitional phenomenon' (Barbiers et al. 2018:182). See the original paper for details.

[12] One aspect of the Schmid & Vogel 2004 analysis that we were not able to implement is the effect of focus/stress on verb cluster ordering, as this feature was neither tested nor transcribed in the SAND questionnaires.

	ASC	DESC	VERBPART	NOMINF	CLUSTINTERR
CALL_HAD_COULD	yes123	no321	yesVerbPart	yesNomInf	noClustInt
CALL_CAN_HAD	no123	yes321	noVerbPart	yesNomInf	noClustInt
CALL_COULD_HAD	no123	yes321	yesVerbPart	yesNomInf	noClustInt
CALLED_HAVE	yes123	yes321	noVerbPart	noNomInf	noClustInt
CAN_CALL_HAD	yes123	no321	noVerbPart	noNomInf	noClustInt
DIED_IS	yes123	yes321	noVerbPart	noNomInf	noClustInt
FIXED_HAVE_MUST	no123	yes321	noVerbPart	noNomInf	noClustInt
FIXED_MUST_HAVE	yes123	no321	noVerbPart	noNomInf	noClustInt
GO_SWIM_IS	yes123	no321	noVerbPart	noNomInf	noClustInt
HAD_CALL_CAN	yes123	no321	yesVerbPart	yesNomInf	noClustInt
HAD_CALL_COULD	yes123	no321	yesVerbPart	yesNomInf	noClustInt
HAD_CAN_CALL	yes123	no321	yesVerbPart	noNomInf	noClustInt
MUST_FIXED_HAVE	yes123	no321	noVerbPart	noNomInf	yesClustInt
…	…	…	…	…	…

TABLE 3. Encoding of the SAND data according to five linguistic parameters taken from Barbiers et al. 2018.

For example, we encoded for every cluster order whether or not it involves INFINITIVUS PRO PARTICIPIO (so-called IPP; see Schmid 2005 for discussion and references). A full list of all linguistic variables used in the analysis is given in Appendix A.[13]

In the statistical analysis described in the next subsection, these sixty-four linguistic variables are treated as supplementary variables. Unlike active variables (here: the geographical information extracted from the SAND database), supplementary variables do not contribute to the construction of the principal components. Instead, they serve to interpret or illustrate those components (see Greenacre 2007:Ch. 12, Husson et al. 2011: 20–24, Levshina 2015:354 for general discussion of supplementary variables). The next subsection provides more details about this.

4.4. THE ANALYSIS: CORRESPONDENCE ANALYSIS. CORRESPONDENCE ANALYSIS (CA) is a principal component method that can be applied to tables containing categorical data (for general discussion, see Greenacre 2007 and Levshina 2015:Ch. 19). The analysis proceeds in three steps: first, the raw data table is transformed into a distance matrix, then the number of dimensions of that distance matrix is reduced, and finally the result of that dimension reduction is matched against the supplementary (here: linguistic) variables. We now proceed to describe these steps in more detail.[14]

In a first step, the raw data table is converted into a distance matrix, a small portion of which is represented in Table 4. This is a 28 × 28 table that has the verb cluster orders from the SAND data both as rows and as columns. Each cluster order is compared pairwise with every other cluster order and a numeric value is assigned to that comparison, indicating how distinct these two cluster orders are from one another: the higher the value, the more different they are.[15] This distance is determined by looking at the active variables in the data table, that is, the geographical data.

[13] Note that the nitty-gritty details of each individual proposal do not matter at this point; we only discuss and elaborate upon these linguistic parameters to the extent that they become relevant in the analysis.

[14] All calculations were carried out in R (R Core Team 2014) using the FactoMineR package (Husson et al. 2014). See Appendix C for technical details as well as a link to all of the data and the R code used to perform the analysis.

[15] Given that the distance between a cluster order A and a cluster order B is identical to that between B and A, the distance matrix is symmetrical across the diagonal. Accordingly, only the lower half of the table is (partially) represented here. Moreover, given that each cluster order is identical to itself, the diagonal of the distance matrix contains only zeros.

	CALL_CAN_HAD	CALL_COULD_HAD	CALL_HAD_COULD	…
CALL_CAN_HAD	0.00			…
CALL_COULD_HAD	4.00	0.00		…
CALL_HAD_COULD	3.74	3.46	0.00	…
CALLED_HAVE	13.03	13.03	13.41	…
CAN_CALL_HAD	5.29	5.29	4.24	…
DIED_IS	13.00	13.00	13.37	…
FIXED_HAVE_MUST	4.89	4.89	5.65	…
FIXED_MUST_HAVE	11.83	11.83	11.66	…
GO_SWIM_IS	10.14	10.14	9.64	…
HAD_CALL_CAN	3.87	4.12	3.00	…
…	…	…	…	…

TABLE 4. Upper left-hand corner of a distance matrix based on the SAND verb cluster data.

Concretely, the more two cluster orders occur in the same dialect locations, the smaller the distance between them will be. The result is a measure of the degree of similarity (or conversely, difference) between the various cluster orders based on their geographical distribution. Note that the notion of 'geographical distribution' is not dependent on those dialect locations forming a contiguous dialect region. Rather, the geographical data are merely used as binary variables to determine which cluster orders typically go together and which ones do not (though see §6.3 below for a geographical analysis of the data).

The second step of the analysis involves dimension reduction. As in principal component analysis, the goal of CA is to reduce a (typically large) set of possibly correlated variables to a smaller group of linearly uncorrelated ones. Put differently, in the distance matrix shown in Table 4, each cluster order is situated in a twenty-eight-dimensional space, and the dimensionality of this space needs to be reduced for those data to be visualized and interpreted. For example, a two-dimensional representation of the verb cluster data under investigation is given in Figure 3.

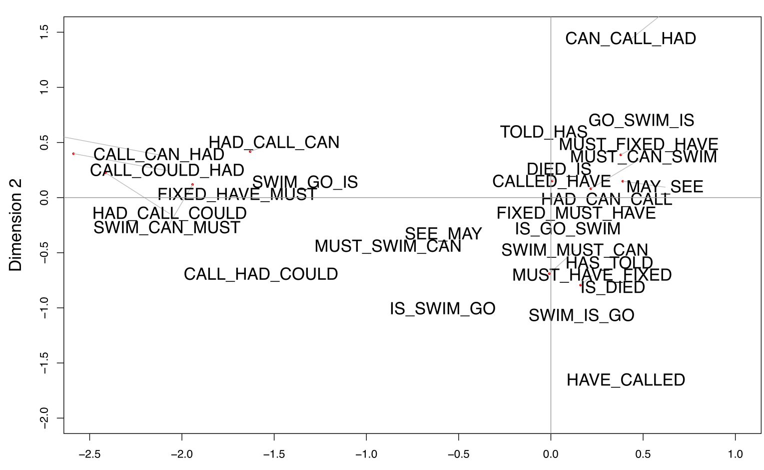

FIGURE 3. Two-dimensional representation of the SAND verb cluster data.

In this graph, each of the twenty-eight cluster orders is situated on a two-dimensional plane. When two cluster orders are close together (e.g. CALL_COULD_HAD and CALL_CAN_HAD on the left-hand side of the graph), this means that they have a highly similar geographical distribution, while when two orders are far apart (like

HAVE_CALLED at the bottom and CAN_CALL_HAD at the top of the graph), they typically do not cooccur in the same dialect locations. In other words, Fig. 3 offers a visual representation of the degree of similarity between the twenty-eight cluster orders.

A dimension reduction of this sort always involves a trade-off between, on the one hand, explaining as much of the variance from the original data set as possible and, on the other hand, keeping the number of dimensions as small as possible for easy visualization and interpretation. In order to determine the appropriate cutoff point, we can make use of a so-called scree plot. This two-dimensional graph represents the dimensions on the x-axis and indicates on the y-axis the percentage of variance explained by that dimension. The scree plot for the SAND verb cluster data is represented in Figure 4. This graph shows that the first two dimensions combined explain roughly 35% of the variance found in the data.[16] After the second dimension there is a sharp drop in explanatory power, and the scree plot begins to flatten out: adding additional dimensions to the analysis represents only a modest increase in explanatory power. In other words, the scree plot suggests that it is the first two dimensions of the CA that should be subject to further exploration and interpretation.

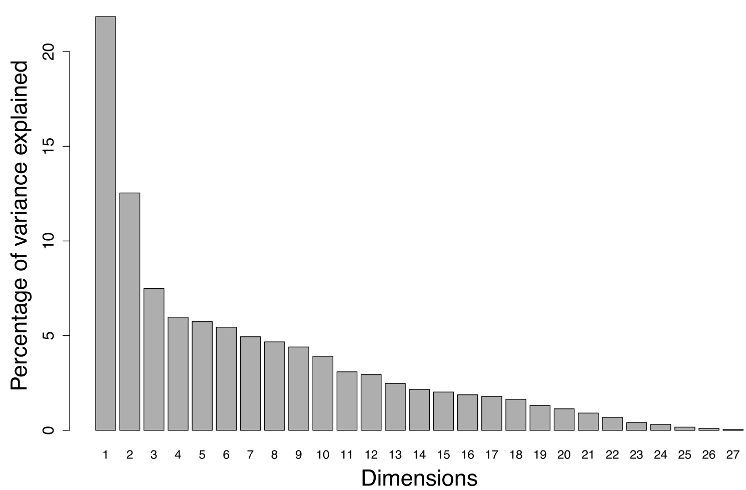

FIGURE 4. Scree plot for the CA of the SAND verb cluster data.

Consider again the graph in Fig. 3. It shows which verb cluster orders typically cluster together, and which ones do not. If the microvariation in Dutch dialect verb cluster ordering is to be reduced (at least partially) to grammatical parameters (see the discussion in §3), then we expect the pattern in Fig. 3 to be determined (at least partially) by such parameters. Put differently, cluster orders that are close together in the graph should be the result of the same or a highly similar parameter setting, while orders that are further apart should have fewer parameter values in common. Grammatical parameters thus create NATURAL CLASSES of verb cluster orders. This is where the supplementary (linguistic) variables come in: we use them to interpret the two dimensions that were retained in step two of the analysis. This is achieved by comparing the first two di-

[16] The fact that the percentages in Fig. 4 are generally fairly low is a consequence of the sparseness of the data table: with twenty-eight individuals and 185 variables it contains 5,180 cells, but only 1,988 of them are filled by a '1'; the majority of the cells are empty. See Husson et al. 2011:92–101 for a comparable data table with similar CA results.

mensions of the CA to those variables. The central question is if—or to what extent—the data patterns in Fig. 3 align with a theoretical property of the clusters in question. There are two basic ways of testing this (both of which are illustrated in detail in the next section). The first is to color-code the plot in Fig. 3 according to (the values of) a linguistic variable and to see if cluster orders that have the same color (i.e. that share some grammatical property) are also close together in the graph (i.e. have a similar geographical distribution). The second is to calculate, for each combination of linguistic variable and CA dimension, the squared correlation ratio (η^2), which provides a measure for the proportion of variance on that particular dimension that is explained by that linguistic variable. The value of η^2 is between 0 and 1, and the higher the number, the stronger the correlation between the dimension and the linguistic variable.

This concludes the methodological section of this article. We have made explicit how both the raw data from the SAND project and the theoretical linguistic literature on verb clusters were operationalized for a quantitative, statistical analysis that takes the form of a CA. This method allows us to reduce the multidimensional variational space to a two-dimensional one. In the next section we examine for these two dimensions the extent to which they align with the supplementary linguistic variables, and in §6 we interpret those results from a formal linguistic point of view.

5. Results.

5.1. Introduction. In this section we present the results of the analysis outlined in the previous section. For each of the two dimensions retained in the analysis we indicate which of the linguistic variables correlate most strongly with that dimension. The section is highly descriptive in nature; for a linguistic interpretation of the results presented here, we refer the reader to §6.

Before proceeding with the discussion of the results, we need to make one preliminary remark. As pointed out by Richardson (2011), the value of η^2 for the combination of a dimension and a particular categorical variable is sensitive to the number of values this variable can have: the higher the number of possible values, the higher the value of η^2.[17] This means that when evaluating the results, one should be wary of variables that have a high η^2-value merely (or mostly) because they have many different values. Accordingly, in what follows we mainly concentrate on two- or three-valued variables when discussing the results of the CA.[18]

5.2. Dimension 1. Table 5 lists which of the supplementary (linguistic) variables in the CA described above have the highest squared correlation ratio for the first dimension. In light of the preliminary remark made above, we also list for each variable how many values it has.

The variable Add.ClusterOrder is one of the nine variables that was added on the basis of the linguistic literature, but not tied to a specific analysis. It encodes all clusters in terms of their basic order (i.e. 1-2, 1-2-3, 3-2-1, etc.), regardless of the verbs making up the cluster. It has seven possible values (two orders for two-verb clusters plus five

[17] One way of clearly demonstrating this is by introducing a fake variable into the data set, which assigns a different value to each of the twenty-eight cluster orders. Such a variable has a ‘perfect’ η^2-value of 1.

[18] One type of variable that we systematically leave undiscussed in this section is the summary variables we constructed for each analysis (see Appendix A for the full list of variables). Given that the values for these summary variables are a concatenation of the values of all the variables making up that analysis, they inevitably have too many values to be properly evaluated in terms of their η^2-value. It is our hope, though, that in future work variables such as these will provide a measure for directly comparing entire analyses in terms of their success in accounting for a particular data set.

VARIABLE	η^2	# OF VALUES
Add.ClusterOrder	0.805	7
SchmiVo.MAPlrV	0.769	4
BarBenDros.asc	0.696	2
HaegRiems.inversion.modal	0.599	3
Add.LightHeavyOrdering	0.508	13
BarBen.NomInf	0.460	3
BarBenDros.NomInf	0.422	2

TABLE 5. The highest η^2-values for dimension 1.

orders for three-verb clusters—the 2-1-3 order is missing from the data set), which arguably (at least partially) explains its high position in this ranking. This variable will also make a reappearance in the η^2-ranking of the second dimension (see §5.3 below), possibly for the same reason.

A similar fate befalls the variable Add.LightHeavyOrdering. It is also one of the nine additional variables, and it is inspired by Abels (2011) and Bobaljik (2004), who suggest that cluster ordering might be sensitive to the 'morphological size' of the verb forms involved in the cluster, the idea being that participles are 'smaller' than infinitives (see Abels 2011:24). In order to test the effect of the morphological shape on cluster ordering, we encoded the twenty-eight cluster orders in terms of their morphological make-up. For example, the cluster IS_DIED was encoded as FinPart (a finite verb followed by a participle), SWIM_CAN_MUST as InfInfFin, and so forth. Given that this method of encoding yielded thirteen different values,[19] however, its high η^2-value is arguably an artefact, which seems corroborated by the fact that this same variable also has a high η^2-value for the second dimension (see §5.3 below). Accordingly, we set this variable aside for the remainder of the discussion.

Things get more interesting when we consider the other variables in Table 5, in particular the ones that have only two values. Let us start with BarBenDros.asc. It concerns one of the variables based on Barbiers et al. 2018 that was introduced in §4.3 above. It is set to 'yes' when the cluster order is compatible with a uniformly ascending order in BBD's analysis, and to 'no' when it is not. As indicated in Table 5, this variable has a high squared correlation ratio of 0.696.[20] The color-coded plot in Figure 5 provides a visual representation of this.

The information to focus on in this graph is the extent to which the distribution of points along the x-axis (i.e. dimension 1) correlates with the color coding, in particular the contrast between black (a negative setting for the variable) and red (a positive setting).[21] As is clear from the graph, the red and black points cluster sharply along this horizontal dimension. This further confirms that the variable BarBenDros.asc provides a good match for the first dimension of the CA.

The variable SchmiVo.MAPlrV is taken from Schmid & Vogel 2004. Their account is based in optimality theory (OT), and this variable corresponds to a constraint in their analysis. The constraint in question is given in 24.

[19] FinInf, FinInfInf, FinInfPart, FinPart, FinPartInf, InfFin, InfFinInf, InfFinPart, InfInfFin, InfPartFin, PartFin, PartFinInf, and PartInfFin.

[20] It is hard to find absolute measures for η^2 to determine the size of the effect. Some authors cite Cohen 1962, in which case an η^2-value of 0.0099, 0.0588, and 0.1379 would correspond to a small, medium, and large effect, respectively, but see Richardson 2011 for critical discussion.

[21] Note that in the black and white versions of this and similar figures, red corresponds to dark gray, green to light gray, and black to black.

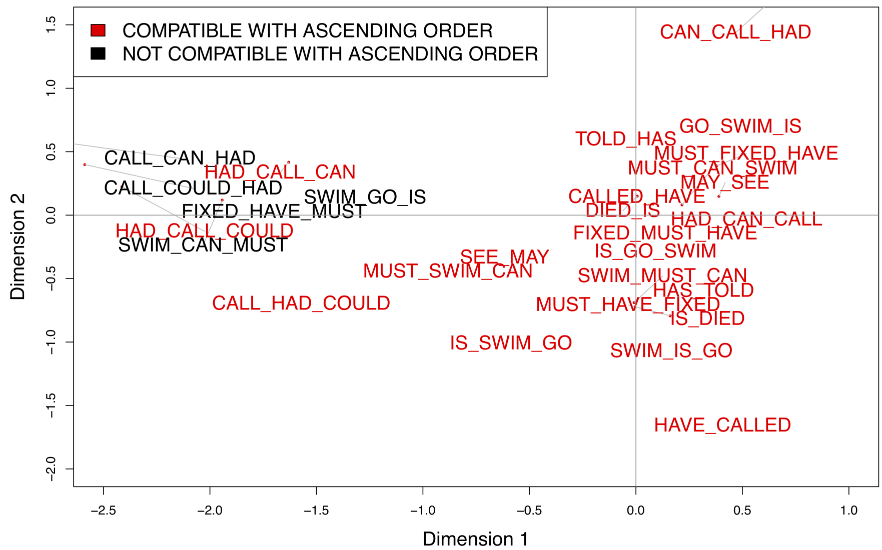

FIGURE 5. Two-dimensional representation of the SAND verb cluster data color-coded according to the first parameter from Barbiers et al. 2018.

(24) **MAPlrV:** The heads of an extended projection of V are linearized in a left-to-right fashion; that is, if head A asymmetrically c-commands head B at LF, then the PF correspondent of A precedes the one of B at PF.

As is clear from the definition, this constraint sets head-initial orders apart from non-head-initial ones. As such, it should come as no surprise that it patterns similarly to the head-initial (ascending) setting for the first parameter from Barbiers et al. 2018. At the same time, though, Schmid and Vogel's (2004) variable has four possible values (depending on the number of times the constraint in 24 is violated in the OT tableau), which means that its η^2-value in Table 5 is partially artificially inflated. The variable HaegRiems.inversion.modal is taken from Haegeman & van Riemsdijk 1986. It concerns the word order of a modal verb vis-à-vis its complement. The variable is set to 'yes' when a modal precedes its complement, to 'no' when the modal is preceded by its complement, and to 'dna' when the cluster does not contain a modal. The color-coded graph in Figure 6 provides a visual representation of this variable. Once again, the contrast between the positive setting of the parameter (the green points) and the negative setting (the red points) lines up with the horizontal dimension.

The final two variables from Table 5 are virtually identical. We focus on BarBenDros.NomInf, as this one stems from Barbiers et al. 2018, the article discussed above.[22] It concerns BBD's third parameter, that is, the one in 20. This variable is set to 'yes' when the order contains—or can contain; cf. n. 10—a nominalized infinitive according to BBD's analysis, and to 'no' when it does not. As the graph in Figure 7 shows, this variable also provides a clear match for the first dimension of the CA.

[22] The analysis in Barbiers & Bennis 2010 (from which the variable BarBen.NomInf was taken) can be seen as the precursor of the one in Barbiers et al. 2018. The only difference between the two variables is the fact that BarBen.NomInf has a 'dna' value for clusters that do not contain an infinitive, whereas in the variable BarBenDros.NomInf such clusters are marked as 'no'.

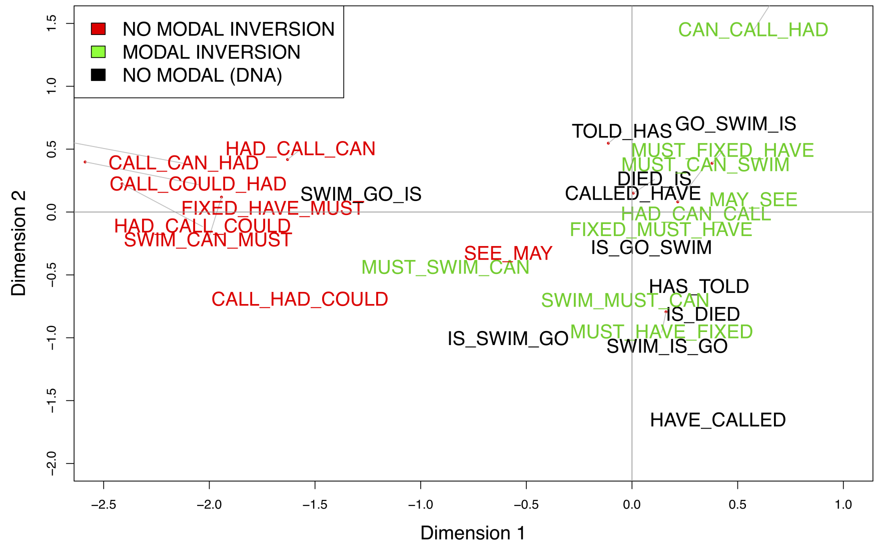

FIGURE 6. Two-dimensional representation of the SAND verb cluster data color-coded according to the modal inversion parameter from Haegeman & van Riemsdijk 1986.

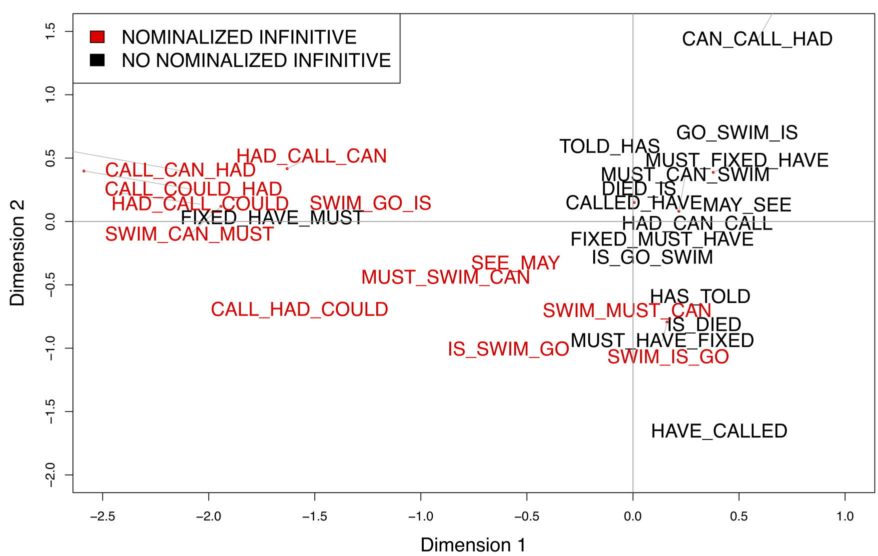

FIGURE 7. Two-dimensional representation of the SAND verb cluster data color-coded according to the third parameter from Barbiers et al. 2018.

This concludes our discussion of dimension 1. We have outlined and illustrated which of the linguistic variables have the highest η^2-values with respect to this dimension. In §6 we translate these findings into a theoretical analysis, but before doing so, we turn to the second dimension.

5.3. DIMENSION 2. The linguistic variables with the highest squared correlation ratio for the second dimension are listed in Table 6.

The variables Add.LightHeavyOrdering and Add.ClusterOrder were discussed in the previous subsection and are set aside for the reasons outlined there. As in the previous

VARIABLE	η^2	# OF VALUES
Add.LightHeavyOrdering	0.794	13
Add.ClusterOrder	0.457	7
SchmiVo.MAPlrVfunc	0.451	4
Add.PartPrecedes(IPP).Aux	0.429	3
BarBen.VerbPart	0.408	3
Add.Slope	0.390	4
BarBenDros.VerbPart	0.292	2

TABLE 6. The highest η^2-values for dimension 2.

section, we first turn our attention to variables that have only two values. For the second dimension, this is once again a variable extracted from Barbiers et al. 2018, namely the second parameter (see 16). This variable is set to 'yes' when the cluster (necessarily) involves a verbal participle according to BBD, and to 'no' when it does not.[23] The graph in Figure 8 visualizes the relationship between this variable and the second dimension of the CA.

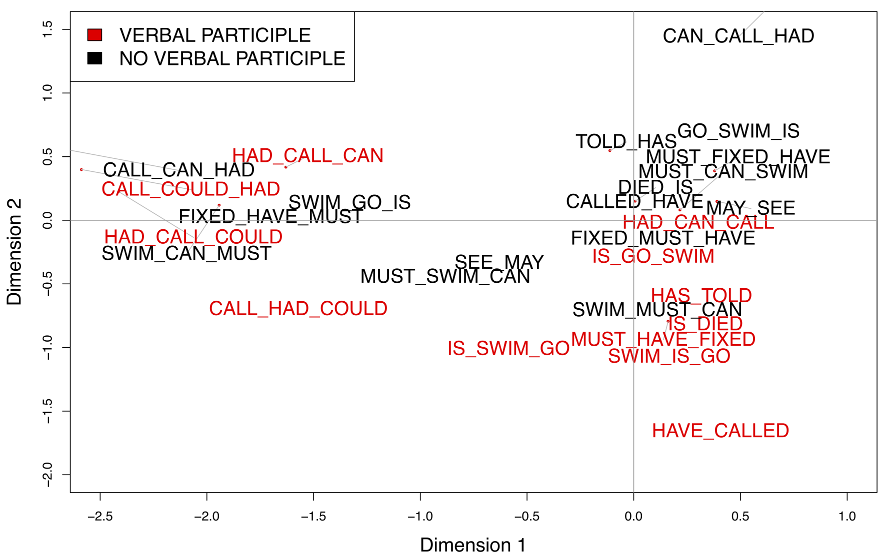

FIGURE 8. Two-dimensional representation of the SAND verb cluster data color-coded according to the second parameter from Barbiers et al. 2018.

This time we are focusing on the distribution of the points along the y-axis, and the extent to which that distribution matches the different colors. Although the picture is less clear than the ones discussed in the previous subsection—note also that overall, the η^2-values in Table 6 are smaller than those in Table 5—there does seem to be a tendency for the red points (the clusters containing a verbal predicate) to be below the x-axis, and for the black points to be above it. The distinction becomes sharper when we consider the variable Add.PartPrecedes(IPP).Aux. This is an additional variable that is set to 'yes' when a participle precedes its auxiliary and to 'no' when it follows it. It differs from BarBenDros.VerbPart in that it has a 'dna' value for clusters that do not contain a

[23] The parameter BarBen.VerbPart (from Barbiers & Bennis 2010) is once again identical to this variable, save for the addition of a 'dna' value. Note that assigning values for a parameter requires a certain amount of analysis. In the case under discussion here, for example, we treat IPP infinitives as participles.

participle. As the graph in Figure 9 shows, this offers a somewhat clearer picture, with the green points (positive setting) above the x-axis, the red points (negative setting) below, and the (essentially irrelevant) black points somewhere in the middle.

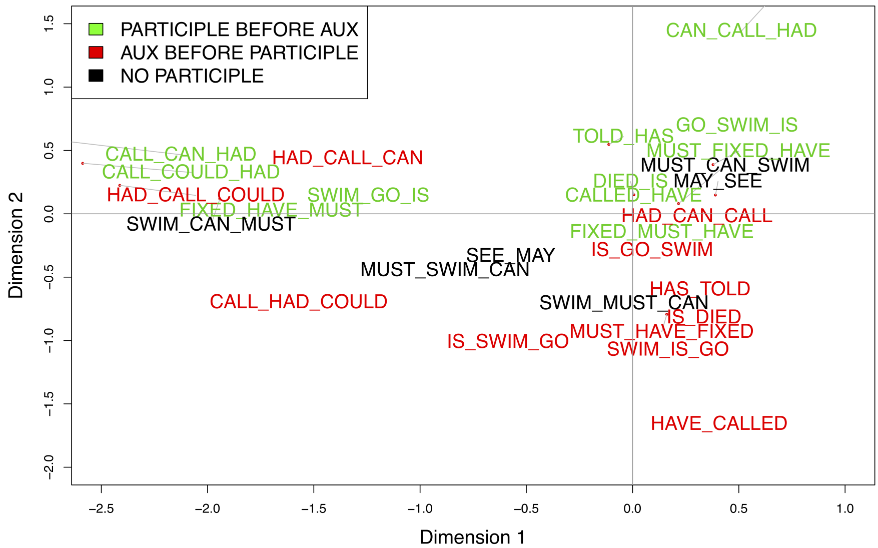

FIGURE 9. Two-dimensional representation of the SAND verb cluster data color-coded according to the additional parameter Add.PartPrecedes(IPP).Aux.

Finally, there are two variables in Table 6 that have four possible values. The first once again stems from Schmid and Vogel (2004) and corresponds to a constraint in their OT account. It is defined in 25.

(25) **MAPlrVfunc:** If A is a functional verb (or a verb containing functional features) that asymmetrically c-commands at LF another verb B that belongs to the same extended projection, then the correspondent of A precedes that of B at PF.

Given that auxiliaries that select a participle fall under Schmid and Vogel's (2004) definition of 'functional verb', it should come as no surprise that a variable based on the constraint in 25 patterns with the other variables discussed in this subsection.

The final variable in Table 6 is Add.Slope. It annotates whether the cluster is ascending or descending. Given that three-verb clusters are not necessarily uniformly ascending or descending, the variable has four possible values: ascending, descending, ascending-descending (e.g. 1-3-2), and descending-ascending (e.g. 3-1-2). This variable has a squared correlation ratio of 0.390, but just like the preceding one it has four values and therefore a somewhat artificially inflated η^2-value.

This concludes our presentation of the results of the CA. For each of the two dimensions we have listed which of the linguistic variables correlates most strongly with that dimension. In the next section we turn to the linguistic analysis of these findings.

6. INTERPRETATION.

6.1. INTRODUCTION. In this section we translate the results from the CA into a linguistic microparametric account of verb cluster ordering in Dutch. We proceed in two steps. First, in §6.2, we propose an analysis that is highly similar to the one put forward in Barbiers et al. 2018 and outlined in §4.3 above. Based on this account we define which data patterns or dialect types we expect to find in the data. Section 6.3 sets out to find these types in the actual data set, and in so doing, examines the role of two additional extragrammatical sources of variation, namely priming caused by translation questions from Standard Dutch and geographical proximity. We propose an implementation of our parameters that circumvents the effects of priming and introduce an additional experiment (based on *k*-nearest neighbor classification) to compare the predictive power of our analysis vis-à-vis one based on geographical location.

6.2. THE ANALYSIS. Recall from the discussion in §5 that three variables based on the Barbiers et al. 2018 account of Dutch verb clusters had high η^2-values for the first two dimensions of the CA, in spite of the fact that all three of them are bivalued. We take this to be a strong indication that a theoretical account along the lines proposed by BBD is correct. In particular, like BBD, we take verb clusters to be the result of Merge, a narrow syntactic operation that is inherently free of linear ordering. Word-order variation in verb clusters comes about as a result of the interaction between the three parameters in 26.

(26) a. **AscDesc:** Verb clusters are linearized in a strictly {ascending/descending} fashion.
b. **PartV:** Participles {can/cannot} be verbal in a verb cluster.
c. **InfN:** Infinitives {can/cannot} be nominal in a verb cluster.

The parallelism between the parameters in 26 and the first three parameters of BBD (see 11, 16, and 20) should be obvious. In fact, there is only one aspect in which our account differs from that of BBD and it concerns their fourth parameter (see 21): the question of whether a dialect allows for cluster interruption. The reason our account makes no reference to BBD's cluster interruption parameter is because that parameter received very little support in the analysis. Table 7 lists the η^2-values for this variable for the first two dimensions of the CA. The low values in this table—compare and contrast with those in Tables 5 and 6—make clear that this variable plays virtually no role in accounting for the variance in the data set.[24] As a result, we leave a discussion of the relation between cluster order and cluster interruption as a topic for further research.

VARIABLE	DIM 1	DIM 2
BarBenDros.ClustInterr	0.027	0.050

TABLE 7. The η^2-values of the cluster interruption parameter from Barbiers et al. 2018 for the first two dimensions of the CA.

We can now move to the more concrete implementation details of the account sketched in 26. Suppose the parameter AscDesc in a particular dialect is set to 'descending'. This implies that strictly descending orders such as 2-1 and 3-2-1 should be well-formed in that dialect. In addition, just as was the case in BBD's account (see n. 10), the setting of the other two parameters in 26 now becomes irrelevant—or rather, their effect becomes undetectable. Regardless of whether the main verb is verbal, adjectival (in the case of a participle), or nominal (in the case of an infinitive), it will al-

[24] It is interesting to speculate why this might be the case. If BBD are right in assuming that the similarity in geographical distribution between the 1-3-2 order and bona fide cases of cluster interruption provides a strong argument in favor of their fourth parameter (cf. n. 11), adding such cluster interruption data to the data set might help bring out that parameter. Another thing to note is that this parameter receives a positive value in only one word order of one cluster type, namely MUST_FIXED_HAVE, as this is the only one that contains a participle as V_3 in the 1-3-2 order. This scarcity of positive evidence might have also contributed to the low η^2-scores in Table 7.

ways precede the other verbs in the cluster and no new cluster orders arise. Summing up, a dialect in which AscDesc is set to 'descending' is expected to display only the orders 2-1 and 3-2-1 in its verb clusters.

When AscDesc is set to 'ascending', however, things become more complicated. The first thing we need to draw attention to concerns the specific way in which our PartV-parameter—as well as BBD's parameter in 16, on which ours is modeled—is formulated. Note that it makes reference not to whether a dialect allows for ADJECTIVAL participles, but to whether it allows for VERBAL ones.[25] This implies that all dialects should allow a participle to precede the cluster, but only in a subset—those for which PartV is set to 'yes'—can a participle also occur at the end of the cluster. Dialects with an 'ascending' setting for AscDesc and negative settings for both PartV and InfN should then allow for the following cluster orders: 1-2 and 1-2-3 in case the main verb is an infinitive, 2-1 and 3-1-2 in case the main verb is a participle, and 2-3-1 in IPP contexts.[26] An ascending dialect with a positive setting for PartV but a negative one for InfN would in addition allow for the orders 1-2 and 1-2-3 in cases where the main verb is a participle. Ascending dialects with the inverse parameter setting (negative for PartV, positive for InfN) would allow for the following orders: 1-2 and 1-2-3 in case the main verb is an infinitive, 2-1 and 3-1-2 in all cluster types, and 2-3-1 in IPP contexts. Finally, an ascending dialect with a positive setting for both PartV and InfN would allow for all of the above-mentioned orders. Table 8 sums up the variation predicted by the current account.

DIALECT TYPE	ASCDESC	PARTV	INFN	PREDICTED ORDERS
1	Desc			2-1, 3-2-1
2	Asc	−	−	1-2_{Inf}, 1-2-3_{Inf}, 2_{Part}-1, 3_{Part}-1-2, 2-3-1
3	Asc	+	−	1-2, 1-2-3, 2-3-1, 2_{Part}-1, 3_{Part}-1-2
4	Asc	−	+	1-2_{Inf}, 1-2-3_{Inf}, 2-1, 3-1-2, 2-3-1
5	Asc	+	+	1-2, 1-2-3, 2-1, 3-1-2, 2-3-1

TABLE 8. The five dialect types predicted by the current account.

This sums up our microparametric analysis of verb clusters in Dutch. It should be clear, of course, that the classification in Table 8 represents an idealization of the data. In particular, recall that in §3 we showed there to be 137 different dialect types in the SAND data under investigation here. In the next subsection we explore the extent to which the five-way distinction in Table 8 can be reconciled with those 137 types, and we address a number of additional questions raised by our account so far.

6.3. THEORY VS. PRACTICE: ADDITIONAL SOURCES OF VARIATION.

INTRODUCTION. In the previous subsection we defined the five dialect types that are predicted to occur by our microparametric analysis of verb cluster ordering (which very closely matches the analysis in Barbiers et al. 2018). When we now list for each of those five types how many SAND dialects fit the given description—that is, how many dialects show all and only the predicted orders—it turns out that there is not a single SAND dialect that provides a perfect match for the microparametric options that fall

[25] See Barbiers et al. 2018:160–63 for argumentation as to why this way of formulating the parameter is preferred.

[26] We follow Barbiers et al. 2018 in assuming that the 2-3-1 order found in IPP contexts is dependent on their second parameter/our PartV. BBD's proposal is that in a 2-3-1 order, V_2 and V_3 form a syntactically complex cluster that as a whole functions like a participle (see also Zwart 2015 for similar ideas, implemented in terms of layered derivations). Given that the 2-3 complex precedes V_1, 2-3-1 orders represent cases of an adjectivally used participle.

out from our theoretical account. On the one hand, this should not come as too much of a surprise, given the data description we gave earlier. Recall from §3 that our 185 dialects represent no fewer than 137 different types when it comes to verb cluster ordering. Add to this the fact that there are 71,751,825 theoretically possible combinations of verb cluster orders in our data set,[27] and it becomes clear that the vast majority of those combinations are not attested, and that combinations represented by more than one dialect are expected to be very rare. Finally, recall from §4.4 that the first two dimensions of our CA—on which we based our theoretical analysis—account for only 35% of the variance in the data set. This means that there are important additional sources of variation at play.

On the other hand, the CA also clearly showed that the parametric account of Barbiers et al. 2018 (as partly replicated in the previous subsection) provides a good match for those first two dimensions and hence does carry quite a bit of weight in accounting for the variance in the data.

In this subsection we reconcile these two forces with one another. We first identify two additional sources of variation, namely priming (or repetition effects) caused by the translation questions, and geographical proximity. Given that individual cluster orders cannot be directly coded for these two factors, we cannot rely on CA to bring them to the surface and hence will have to resort to other, more indirect, means of detecting their presence in the data. Then we propose a way of operationalizing our analysis that, on the one hand, eliminates the role of priming, while, on the other hand, allowing us to precisely weigh the effect of our analysis vis-à-vis that of geographical proximity. The conclusion of this additional experiment will be that the analysis outlined in the previous subsection explains part of the variance in the verb cluster data set above and beyond the variation caused by geographical proximity.

PRIMING AND GEOGRAPHICAL PROXIMITY. Recall from §4.2 that the verb cluster data on which this article is based come from two different question types: translation tasks and elicitation questions. In the first type, informants were asked to translate a question from Standard Dutch into their dialect, while in the second they were asked to judge a (prerecorded) sentence uttered in their dialect. As pointed out by Cornips and Poletto (2005, 2007), translation tasks carry the risk of creating priming effects, whereby an informant is influenced by the order offered to them—especially if that order is grammatical in the standard language—and as a result produces a dialect sentence that is a word-for-word translation of the task sentence, but that is not grammatical in the informant's dialect. There are two reasons to think this kind of priming or repetition effect has played a role in our data set as well. First of all, previous research has shown that priming indeed influences word-order choice in verb clusters in Dutch (see the discussion of De Sutter 2009 in §2.2). A second reason to think priming is an additional source of variation in our data is suggested by the graph in Figure 10.

This graph is basically identical to the one in Fig. 3, but color-coded according to question type. As is clear from the key, we have split the translation questions into two subcategories: the matching translation, whereby the cluster order offered by the informant is identical to that in the task sentence, and the nonmatching one, where answer and task sentence show a different order. Note that the matching (i.e. green) orders are mostly located close to the origin of the graph. This shows that these orders did not con-

[27] This number takes into consideration the hypothesis that each dialect allows at least one order for each type of cluster. For example, for a cluster consisting of the auxiliary 'to be' and a participle (IS_DIED), each dialect will allow a 1-2 order, a 2-1 order, or both, but not neither.

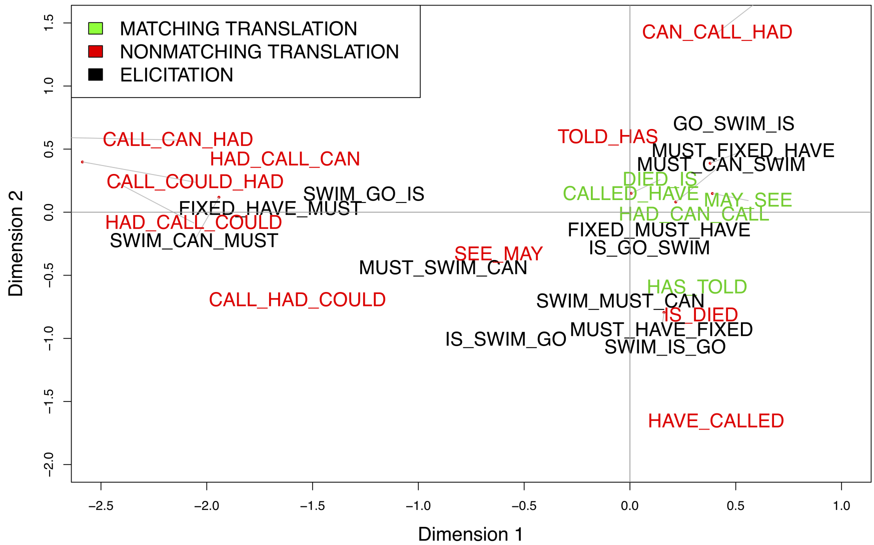

FIGURE 10. Two-dimensional representation of the SAND verb cluster data, color-coded according to question methodology.

tribute substantially to the construction of the first two dimensions; that is, they were less instructive in finding patterns in the data set. The nonmatching (red) orders, by contrast, can be found at the extremes of the graph, both horizontally and vertically, showing that they carried substantial weight in determining the dimensions. This is further confirmed by the frequency data in Fig. 2 above. Note how three of the four most frequent cluster orders correspond to matching translation questions and, moreover, how these orders are accepted by virtually all informants. This too suggests that there is a priming or repetition effect at play, which we take to be an additional source of variation in our data set.

Another additional factor is geographical proximity. In the data table partially represented in Table 2, dialect locations are treated as simple, independent binary variables. As is well known from the dialectological literature, this is a simplification, in that dialect phenomena tend to show coherent geographical patterns (see Spruit 2008 for a particularly clear illustration of this in the context of the SAND data). Needless to say, these geographical tendencies might form another source of variation: informants might be influenced by nearby dialects and thus develop repertoires of verb cluster orders that are partially similar to those of speakers of nearby dialects. Just as was the case with priming, there are reasons to think that geographical proximity indeed has an effect on the variation in our data set as well. We can show this by measuring the correlation between linguistic distance and geographical distance. Consider again Table 4 above. This is a distance matrix indicating how different verb cluster orders are from one another based on the dialect locations where those orders occur. We can also construct the mirror image of this distance matrix, in particular one that indicates how different DIALECT LOCATIONS are from one another, based on the number of verb cluster orders they share. A small portion of that matrix is given in Table 9.

What this table tells us is that the linguistic difference between, say, Midsland and Ferwerd is smaller than that between Midsland and Visvliet. In other words, Midsland and Ferwerd have more verb cluster orders in common than do Midsland and Visvliet.

	Midsland	Lies	West-Terschelling	Schiermonnikoog	…
Midsland	0.00				…
Lies	0.27	0.00			…
West Terschelling	0.20	0.27	0.00		…
Schiermonnikoog	0.20	0.30	0.40	0.00	…
Ferwerd	0.11	0.20	0.11	0.33	…
Visvliet	0.61	0.47	0.52	0.56	…
Bakkeveen	0.20	0.27	0.20	0.40	…
…	…	…	…	…	…

TABLE 9. Upper left-hand corner of a distance matrix of dialect locations based on the SAND verb cluster data.

Now, given that the rows (and columns) of Table 9 represent geographical locations, we can also construct an actual distance matrix of these locations, where the numbers in the cells represent the number of kilometers that separate these locations. A small portion of that distance matrix is given in Table 10.

	Midsland	Lies	West-Terschelling	Schiermonnikoog	…
Midsland	0.00				…
Lies	4.49	0.00			…
West Terschelling	8.15	12.65	0.00		…
Schiermonnikoog	97.27	92.85	105.33	0.00	…
Ferwerd	59.98	55.80	67.68	40.07	…
Visvliet	108.00	103.87	115.58	25.74	…
Bakkeveen	113.19	109.46	120.11	45.64	…
…	…	…	…	…	…

TABLE 10. Actual distance matrix (in kilometers, as the crow flies) of the SAND locations.

If the variation in our data set is partially caused by geographical proximity, then we predict there to be a positive correlation between Table 9 (linguistic distance) and Table 10 (geographical distance). Following Levshina 2015:348–49 we used the Mantel test to test this prediction. The Mantel statistic, which is identical to the Pearson correlation coefficient r, is 0.403 ($p = 0.001$), which indicates that there is indeed a (moderate) positive correlation between the number of cluster orders two dialects share and their geographical proximity.

Summing up, we have identified two extragrammatical factors that contribute to the variation in our data set: priming and geographical proximity. In the next subsection we combine these findings with the theoretical analysis from the previous subsection.

OPERATIONALIZING AND TESTING THE ANALYSIS. The reason why we found no dialects that fit the profiles predicted by the theoretical analysis was because each profile was defined as a unique combination of cluster orders—1 out of the 71,751,825 theoretically possible combinations. As such, the prediction did not take into account possible additional sources of variation. In other words, the criteria for falling into one of the predicted dialect types were too strict. In this section we propose a new, more relaxed way of defining our parameters, while at the same time taking into account one of the findings from the preceding subsection, namely the effect of priming created by the translation questions. Specifically, we avoid using matching answers to translation questions as a diagnostic for particular settings of our parameters. The criteria we propose to use are the following.

(27) a. **AscDesc:** In order to obtain the setting [Desc], the dialect needs to display at least one 3-2-1 order.

b. **PartV:** In order to obtain the setting [+PartV], the dialect needs to translate at least one 2-1 order into a 1-2 order.

c. **InfN:** In order to obtain the setting [+InfN], the dialect needs to display at least one 2-1, 3-1-2, or 1-3-2 order with an infinitive as main verb.

Before seeing how many dialects fit these descriptions, let us examine the criteria in a bit more detail. For the parameter AscDesc, we chose to make use of the one order that is unique to the [Desc]-setting of this parameter and that cannot be derived under any of the settings of the other parameters: a completely descending order in a three-verb cluster. This order was never offered as a task sentence in a translation question, so there is no risk of priming. For the parameter PartV, we started from Cornips and Poletto's (2005, 2007) observation that while a matching answer to a translation question might be a sign of priming, a nonmatching answer provides a strong signal that the alternative option is strongly preferred to the one that was offered. Thus, speakers who change a 2-1 order in a two-verb cluster with a participle and an auxiliary into a 1-2 order show a clear preference for the verbal status of that participle, and hence for a positive setting for PartV. Finally, for the parameter InfN, we listed orders that require or strongly prefer the nominal status of the infinitive, again while avoiding any order that was offered in a translation question. When we use these criteria to define dialect groups in our data set, we obtain the results outlined in Table 11.

AscDesc	InfN	PartV	# of dialects
Desc	+	+	10
Desc	+	−	39
Desc	−	+	1
Desc	−	−	4
Asc	+	+	29
Asc	+	−	29
Asc	−	+	17
Asc	−	−	56

Table 11. Dialect groups based on the parameter criteria in 27.

When evaluating the results presented in this table, it is useful to also take into account the geographical location of the dialect(group)s listed. In Figure 11 we have plotted the eight dialect groups onto a geographical map. Let us start by considering the dialects with a [Desc]-setting for the first parameter. As Table 11 shows, by far the most common parameter setting for these dialects is [Desc/+InfN/−PartV]. This is exactly as expected, given that (i) one of the criteria for a positive setting of the InfN-parameter was the occurrence of a 2-1 order (see 27c above) and this order is also compatible with a strictly descending dialect, and (ii) a negative setting for the PartV-parameter entails that the translation questions involving two-verb clusters with a participle yielded only (matching) descending orders. Moreover, the thirty-nine dialects that have this parameter setting form a contiguous and geographically coherent region in the north of the language area. Dialects that combine a [Desc]-setting of AscDesc with a negative setting for InfN are predicted not to occur, and as Table 11 shows, such patterns are extremely rare indeed. This leaves the final type of [Desc]-dialects, those with the setting [Desc/+InfN/+PartV]. These are not predicted to occur either, given that a positive setting of PartV entails the occurrence of ascending 1-2 orders in two-verb clusters with a participle, which would in turn be incompatible with the general descending nature of the dialect. The fact that we nonetheless find ten dialects that fit this description might indicate that this is an area where additional factors of variation come into play. In that respect it is quite striking to see that these ten dialects (the brownish dots on the map in Figure 11) are typically situated in border areas, either between the (descending) north and the (ascending) south, or between Belgium/the Netherlands and Germany. This suggests that the patterns of cluster orders displayed by these dialects might be a contact phenomenon.

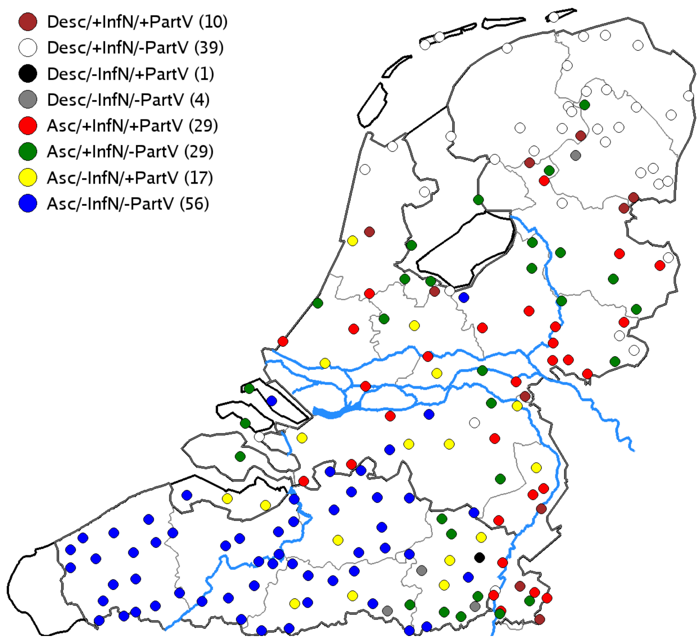

FIGURE 11. Geographical representation of the dialect groups in Table 11.

The dialects with an [Asc]-setting for the AscDesc-parameter are more numerous. The largest group is the one with the parameter setting [Asc/−InfN/−PartV] (the blue dots on the map). This setting represents the variety spoken in the Belgian provinces of West and East Flanders, Brabant, and Antwerp. These dialects generally have a strictly ascending order, except in the case of their participles, which they prefer to treat as adjectival, and which hence show up to the left of the cluster. The other three patterns are all situated in the center and south of the Netherlands and in the two provinces of Limburg. The setting [Asc/−InfN/+PartV] (the yellow dots) unsurprisingly are situated rather close to the Belgian pattern, since it differs from it only in allowing its participles to also occur to the right of their selecting auxiliaries. The other two patterns are more distinct, in that they have a positive setting for InfN, which allows infinitives to occur to the left of the verb cluster, a phenomenon that is rarely found in Belgium. Of these two patterns, the [Asc/+InfN/−PartV]-setting (green dots) is the least prototypically ascending, in that it allows descending orders both in clusters involving infinitives (due to the positive setting of InfN) and in clusters involving participles (due to [−PartV]). Unsurprisingly, then, this pattern seems to form a transition area to the descending dialects in the north (the white dots) and to German (also descending) in the southeast.

Summing up, the criteria given in 27 for defining and detecting the parameters of our syntactic analysis in the data set have proven to be successful in identifying and defining geographical regions with a particular parameter setting. Moreover, by not including matching answers to translation questions in our definition of the parameters, we have excluded the part of the data most at risk for influence from priming. At the same time, however, the role of geographical proximity has not yet been taken into consideration. What is more, the map in Fig. 11 seems to show a gradual transition from a strictly ascending system in the southwest—albeit one with adjectival participles—to a strictly descending one in the north, and with 'unexpected' types limited to transition areas. In other words, the role of geographical location seems to be strong. In order to ascertain the predictive power both of the three linguistic variables, AscDesc, InfN, and PartV, and of the geographical location of each of the 185 dialect types, all with respect to predicting the twenty-eight cluster orders, we performed an experiment with *k*-nearest neighbor classification (Daelemans & van den Bosch 2005). The *k*-nearest neighbor classifier operationalizes the assumption that a new instance of a classification problem can be classified by copying the solution of already encountered examples that are similar to the new instance. Take, for instance, a new dialect type characterized only by its longitude and latitude; whether this dialect uses a particular cluster order, such as CALL_CAN_HAD, would be inferred by the *k*-nearest neighbor classifier by comparing it to all memorized examples of dialect types and copying the class of the *k* dialect types closest to the new dialect type in terms of some geometrical distance (e.g. Euclidean distance); these *k*-closest memorized examples are its 'nearest neighbors'. The number of neighbors taken to predict the outcome is set by the hyperparameter *k*; throughout the current test we set $k = 1$.

For each cluster order we performed three experimental runs with a *k*-nearest neighbor classifier: we used (i) only longitude and latitude as predictive features, or (ii) the binary values of the variables AscDesc, InfN, and PartV, or (iii) all five features combined. Each run was set up as a leave-one-out experiment, where each of the 185 dialect types acted as the unknown dialect type, while all of the remaining 184 dialect types acted as the memorized dialect types. In all runs we left hyperparameters to their default values; aside from $k = 1$ we used the default similarity metric provided by the TiMBL software[28] (Daelemans et al. 2010). This metric computes a distance between two instances X and Y, $\Delta(X, Y)$, by computing the sum of distances per feature, for all features $i = [1 \ldots n]$ (equation 28).

(28) $$\Delta(X,Y) \sum_{i=1}^{n} w_i \delta(x_i, y_i)$$

The distance at feature i, $\delta(x_i, y_i)$, is computed differently for numeric features (longitude and latitude) and symbolic features (the three linguistic variables), as specified in equation 29.

Numeric distance is normalized by the maximal distance in the numerical dimensions found across memorized instances. Symbolic distance is simply 0 when the two instances share the same value at feature i; otherwise the distance is 1.

(29) $$\delta(x_i, y_i) = \begin{cases} \frac{x_i - y_i}{max_i - min_i} & \text{if numeric, otherwise} \\ 0 & \text{if } x_i = y_i \\ 1 & \text{if } x_i \neq y_i \end{cases}$$

In computing the overall distance between X and Y, the distance at feature i is then weighted by feature weight w_i (equation 28). As weighting function we use gain ratio, a data-driven information-theoretic metric that estimates the discriminative power of a

[28] https://languagemachines.github.io/timbl/, version 6.4.13

feature. To do that, first the INFORMATION GAIN of the feature is computed as the difference in uncertainty (i.e. entropy) between the situations without and with knowledge of the value of that feature, as given in equation 30, where C is the set of class labels, V_i is the set of values for feature i, and $H(C) = -\sum_{c \in C} P(c) \log_2 P(c)$ is the entropy of the class labels.

(30) $$w_i = H(C) - \sum_{v \in V_i} P(v) \times H(C|v)$$

A well-known problem with information gain, analogous to the problem mentioned in §5 in CA with the value of η^2 for the combination of a dimension and a particular categorical variable—that it is sensitive to the number of values this variable can have—is that it tends to overestimate the relevance of features with large numbers of values. To normalize information gain over features with high numbers of values, Quinlan (1993) introduced a variant called GAIN RATIO (GR), given in equation 31, which is information gain divided by SPLIT INFO, $si(i)$, the entropy of the feature values (equation 32).

(31) $$w_i = \frac{H(C) - \sum_{v \in V_i} P(v) \times H(C|v)}{si(i)}$$

(32) $$si(i) = H(V) = -\sum_{v \in V_i} P(v) \log_2 P(v)$$

The raw result of one leave-one-out experiment is 185 classifications into + or −, some of which could be incorrect in two ways: either as false positives (cases of − that are labeled by the classifier as +) or as false negatives (cases of + that are labeled by the classifier as −). Measuring accuracy as the percentage of correctly predicted cluster-order values would take the ratio of all correct classifications (i.e. all correctly predicted cases of + and −) over the total number of classifications, including false positives and false negatives. Doing this would leave us vulnerable to effects of class skew when interpreting and comparing our results. For instance, if for a particular construction the positive class + occurs in only ten out of 185 cases, always predicting the negative class would lead to an accuracy of 175 out of 185 cases, or 94.6%. This high accuracy hides the fact that none of the + cases were correctly identified. A proper comparable score would compensate for class skew and would focus on +.

As a common countermeasure for this problem, we computed the AREA UNDER THE ROC CURVE (AUC) of the positive class + (Fawcett 2004). The AUC value is based on the true positive rate of the classifier (also known as recall, or the ratio between the number of true positives of + over the total number of cases of + in the data), and the false positive rate (the ratio between the false positives of + over the total number of cases of − in the data). Plotting false positive rate on the x-axis and true positive rate on the y-axis, a single experiment represents a point in this space; AUC is then the surface between this point and the three coordinates (0,0), (1,0), and (1,1). An AUC score of 0.5 indicates chance behavior, while 1.0 indicates flawless predictions. For more details, see Fawcett 2004.

Averaging over the twenty-eight constructions, computing the mean AUC and standard deviation, we observe the following:

- Using only longitude and latitude as features, the mean AUC is 0.6289 with a standard deviation of 0.1026;
- using only the linguistic analysis features, the mean AUC is 0.5589 with a standard deviation of 0.1241;
- using both feature sets, the mean AUC is 0.6444 with a standard deviation of 0.1187.

A two-tailed paired t-test reveals that the combination of features is significantly different from using only the geographical features ($p = 0.0082$, $t = 2.85$), while the combination of features does not perform significantly differently from using only the linguistic features ($p = 0.081$, $t = -1.81$). The results obtained with the two feature sets separately do not differ significantly either ($p = 0.163$, $t = 1.44$). In other words, although the two feature sets appear to carry information that leads to comparably accurate predictions, they contribute to some extent complementary informational clues that in combination provide the best predictive accuracy.[29] This means that the linguistic analysis proposed in the previous section explains part of the variance in our data set above and beyond the variation that can be attributed to geographical proximity.

6.4. Conclusion. In this section we have interpreted the results of the correspondence analysis carried out and described in the previous two sections. Based on the η^2-scores of the complementary variables, we proposed a microparametric analysis of verb cluster ordering that closely resembles the analysis proposed in Barbiers et al. 2018. A first attempt to identify the five dialect types predicted by that account in our data set failed, which led us to look for (and find) additional, extragrammatical sources of variation: priming caused by translation questions from Standard Dutch, and geographical proximity. The second attempt at defining our parameters was able to circumvent the issues raised by priming and successfully identified the dialect groups corresponding to the various parameter settings in the data set. In order to weigh these results against the role played by geographical location, we performed an experiment with k-nearest neighbor classification, which showed that the information provided by the syntactic analysis is to some extent complementary to that of the geographical information, and that an account that relies on both outperforms either measure individually, significantly so compared to using only the geographical information.

7. Conclusion. This article is situated at the intersection of quantitative and qualitative linguistics. It uses quantitative-statistical methods to further our theoretical understanding of variation in verb cluster ordering in Dutch dialects. In so doing, it harnesses and combines the strengths of both approaches: quantitative linguistics has sophisticated means of dealing with large and highly varied data sets, while hypotheses and analyses from qualitative linguistics can be used to guide and narrow down the interpretation of the statistical results. For the case at hand—verb clusters—we have shown how the 137 dialect types that were manifested in the raw data can be whittled down to the interaction between three grammatical parameters. The method thus allows one to make a detailed proposal about the amount of variation that is due to the grammatical system itself and the portion that should be relegated to extragrammatical factors.

While this article was concerned exclusively with word order in Dutch verb clusters, it should be clear that the method outlined here can be extended to any empirical domain for which both a large data set and a number of explicit competing theoretical proposals are available. One domain that comes to mind in this respect is word-order variation in the noun phrase (Greenberg 1963, Cinque 2005, Abels & Neeleman 2012;

[29] See Appendix B for a table containing all of the AUC scores of our experiment. The mean scores are above-chance performance ($AUC = 0.5$), but they are not high. For several of the twenty-eight cluster orders, the AUC scores for some or all of the leave-one-out experiments are in fact close to chance performance, while other cluster orders are predicted fairly accurately. A general observation from the results in the table is that the combination of the two feature sets indeed tends to improve on the scores of the two sets individually, although this is not always the case.

see also Merlo 2015). More generally, we hope this article illustrates the viability and mutual benefits of an increased collaboration between formal-theoretical (in particular generative) and quantitative-statistical linguists working on language variation.

APPENDIX A: LIST OF LINGUISTIC VARIABLES USED IN THE CORRESPONDENCE ANALYSIS

VARIABLE NAME	INTERPRETATION
BarBenDros.asc	Is the order compatible with an ascending linearization order?
BarBenDros.desc	Is the order compatible with a descending linearization order?
BarBenDros.VerbPart	Does the order involve a verbal participle?
BarBenDros.NomInf	Does the order involve a nominalized infinitive?
BarBenDros.ClustInterr	Does the order involve cluster interruption?
BARBENDROS	The analysis of Barbiers et al. 2018, that is, a summary variable of the preceding five variables.
Haeg.Riems.branch.nonbranch	Is the node undergoing inversion branching or nonbranching?
HaegRiems.3left	Does the order involve leftward displacement of V3 (as discussed in Wurmbrand 2017)?
HaegRiems.base.order	Does the order represent a base-generated one?
HaegRiems.inversion.aux	Does the order involve inversion of the complement of an auxiliary?
HaegRiems.inversion.modal	Does the order involve inversion of the complement of a modal?
HAEG.RIEMS	The analysis of Haegeman & van Riemsdijk 1986, that is, a summary variable of the preceding five variables.
SchmiVo.MAPch	Do complements precede their heads?
SchmiVo.MAPhc	Do heads precede their complements?
SchmiVo.MAPlrV	Is the extended projection of V linearized in an ascending fashion?
SchmiVo.MAPlrVfunc	Do auxiliary verbs precede their verbal complement?
SCHMID.VOGEL	The analysis of Schmid & Vogel 2004, that is, a summary variable of the preceding four variables.
Barbiers.base.generation	Does the order represent a base-generated one?
Barbiers.feature.checking.failure	Does the order represent a feature-checking violation?
Barbiers.spec.pied.piping	Does the derivation of the order involve pied-piping via the specifier?
BARBIERS	The analysis of Barbiers 2005, that is, a summary variable of the preceding three variables.
BarBen.VerbPart	Does the order involve a verbal participle?
BarBen.base.order	Does the order represent a base-generated one?
BarBen.NomInf	Does the order involve a nominalized infinitive?
BarBen.VPR	Does the order involve verb projection raising/cluster interruption?
BARBIERS.BENNIS	The analysis of Barbiers & Bennis 2010, that is, a summary variable of the preceding four variables.
Bader.AuxMod	Do auxiliaries precede modals they c-command?
Bader.base.order	Does the order represent a base-generated one?
Bader.VMod	Do modals precede their verbal complement?
BADER	The analysis of Bader 2012, that is, a summary variable of the preceding three variables.
Abels.base.order	Does the order represent a base-generated one?
Abels.high.order	Does the verbal head precede its complement at the highest node in the cluster?
Abels.low.order	Does the verbal head precede its complement at the lowest node in the cluster?
Abels.prosody	Does the order represent a misalignment between syntactic structure and prosodic structure?
ABELS	The analysis of Abels 2016, that is, a summary variable of the preceding four variables.
HIniHmvt.2LeftAdjoinTo1	Does V_2 left-adjoin to V_1?
HIniHmvt.3LeftAdjoinTo1	Does V_3 left-adjoin to V_1?

(*Appendix A table continues*)

VARIABLE NAME	INTERPRETATION
HIniHmvt.3LeftAdjoinTo2	Does V_3 left-adjoin to V_2?
HIniHmvt.base.order	Does the order represent a base-generated one?
HEADINITIAL.HEADMVT	A head-initial head movement analysis, that is, a summary variable of the preceding four variables.
HFinHmvt.2RightAdjoinTo1	Does V_2 right-adjoin to V_1?
HFinHmvt.3RightAdjoinTo2	Does V_3 right-adjoin to V_2?
HFinHmvt.base.order	Does the order represent a base-generated one?
HEADFINAL.HEADMVT	A head-final head movement analysis, that is, a summary variable of the preceding four variables.
HIniXPmvt.base.order	Does the order represent a base-generated one?
HIniXPmvt.VP2toVP1	Does the order involve movement of VP_2 to VP_1?
HIniXPmvt.VP3toVP1	Does the order involve movement of VP_3 to VP_1?
HIniXPmvt.VP3toVP2	Does the order involve movement of VP_3 to VP_2?
HEADINITIAL.XPMVT	A head-initial XP movement analysis, that is, a summary variable of the preceding four variables.
HFinXPmvt.base.order	Does the order represent a base-generated one?
HFinXPmvt.VP2toVP1	Does the order involve movement of VP_2 to VP_1?
HFinXPmvt.VP3evacuation	Does the order involve movement of VP_3 to the left of the cluster?
HFinXPmvt.VP3toVP2	Does the order involve movement of VP_3 to VP_2?
HEADFINAL.XPMVT	A head-final XP movement analysis, that is, a summary variable of the preceding four variables.
Add.1vs2	Does V_1 precede V_2?
Add.1vs3	Does V_1 precede V_3?
Add.2vs3	Does V_2 precede V_3?
Add.ClusterOrder	What is the cluster order (e.g. 1-2, 1-2-3, 3-1-2, …)?
Add.harmonic	Is the cluster order harmonic (uniformly ascending or uniformly descending) or not?
Add.IPP	Does the cluster order involve IPP?
Add.LightHeavyOrdering	What morphological forms is the cluster made up of (infinitives, finite verbs, participles)?
Add.PartPrecedes(IPP).Aux	Do participles (including IPP-infinitives) precede their selecting auxiliary?
Add.Slope	Is the cluster order ascending, descending, first ascending and then descending (e.g. 1-3-2), or first descending and then ascending (e.g. 3-1-2)?

Appendix B: AUC scores for all twenty-eight cluster orders

ORDER	ONLY GEOGRAPHICAL	ONLY LINGUISTIC	GEO + LINGUISTIC
IS_DIED	0.644	0.996	0.987
CALL_CAN_HAD	0.793	0.500	0.870
SEE_MAY	0.794	0.902	0.853
HAD_CAN_CALL	0.728	0.439	0.841
SWIM_GO_IS	0.700	0.940	0.831
SWIM_MUST_CAN	0.653	0.815	0.780
MUST_FIXED_HAVE	0.794	0.715	0.772
MUST_CAN_SWIM	0.764	0.500	0.758
SWIM_CAN_MUST	0.620	0.832	0.754
GO_SWIM_IS	0.719	0.725	0.746
MAY_SEE	0.745	0.817	0.739
HAD_CALL_CAN	0.478	0.500	0.728
CALL_COULD_HAD	0.716	0.500	0.713
IS_GO_SWIM	0.726	0.713	0.710
MUST_HAVE_FIXED	0.632	0.680	0.693
MUST_SWIM_CAN	0.587	0.500	0.674
FIXED_HAVE_MUST	0.681	0.872	0.669

(*Appendix B table continues*)

ORDER	ONLY GEOGRAPHICAL	ONLY LINGUISTIC	GEO + LINGUISTIC
FIXED_MUST_HAVE	0.646	0.493	0.628
SWIM_IS_GO	0.496	0.500	0.624
TOLD_HAS	0.551	0.690	0.623
HAS_TOLD	0.516	0.689	0.619
CAN_CALL_HAD	0.613	0.500	0.598
IS_SWIM_GO	0.497	0.500	0.500
HAD_CALL_COULD	0.497	0.500	0.497
CALLED_HAVE	0.495	0.500	0.497
CALL_HAD_COULD	0.497	0.500	0.497
DIED_IS	0.661	0.500	0.497
HAVE_CALLED	0.492	0.500	0.489

Appendix C: Supplementary materials and technical details

All calculations were carried out in R (R Core Team 2014) using the FactoMineR package (Husson et al. 2014), in particular the CA function in that package. For details about the algorithms and math underlying that function we refer the reader to Husson et al. 2011. The Mantel test discussed in §6.3 was carried out using the mantel function from the vegan package (Oksanen et al. 2018), and the *k*-nearest neighbor experiments were performed with the TiMBL software package (Daelemans et al. 2010), specifically TiMBL version 6.4.13; see https://languagemachines.github.io/timbl/. The data set that formed the input for the CA and the R code used to perform the analysis are available as a project on the Open Science Framework: https://osf.io/rpfvq/?view_only=413d47b0708748c78c47b519dc064028.

All scatterplots were drawn using the textplot function from the R package wordcloud (Fellows 2014) to ensure that the labels in the plot do not cover one another. This is why some cluster orders are connected by a short gray line to their actual coordinates (indicated by a red dot). The geographical maps were drawn using the cartographic software included in dynaSAND, the online dynamic version of the SAND atlas (Barbiers et al. 2006).

In addition to the complete-cases analysis reported on in the main text, where the CA was performed only on the 185 dialect locations that did not contain any NAs, we also performed a CA on the full data set, but with imputed values for the missing data. In particular, we imputed the missing data using the imputeCA function of the R package missMDA (Husson & Josse 2013). This function applies an iterative algorithm to impute missing values in a categorical data table (see Husson & Josse 2013 for more details and Josse et al. 2012 for general discussion of imputing missing data), which allowed us to perform the analysis on the complete 28 × 267 data table. The results of this alternative analysis turned out to not differ significantly from the ones reported in the main text. In order to test this we applied a simple linear regression to the coordinates on the first and second dimensions of both analyses. In particular, a simple linear regression was calculated to predict the coordinates of the data imputation analysis based on the coordinates of the complete-cases analysis. For the first coordinates a significant regression was found ($F(1,26) = 424.4$, $p < 0.000$), with an adjusted R^2 of .940, and for the second coordinates too ($F(1,26) = 232$, $p < 0.000$), with an adjusted R^2 of .895.

REFERENCES

Abels, Klaus. 2011. Hierarchy-order relations in the Germanic verb cluster and in the noun phrase. *Groninger Arbeiten zur Germanistischen Linguistik* 53.1–28.

Abels, Klaus. 2016. The fundamental left–right asymmetry in the Germanic verb cluster. *The Journal of Comparative Germanic Linguistics* 19.179–220. DOI: 10.1007/s10828-016-9082-9.

Abels, Klaus, and Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15.25–74. DOI: 10.1111/j.1467-9612.2011.00163.x.

Bader, Markus. 2012. Verb-cluster variations: A harmonic grammar analysis. Handout of a talk presented at New Ways of Analyzing Syntactic Variation, Nijmegen, November 2012.

Baker, Mark. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Barbiers, Sjef. 2005. Word order variation in three-verb clusters and the division of labour between generative linguistics and sociolinguistics. *Syntax and variation: Reconciling the biological and the social* (Current issues in linguistic theory 265), ed. by Leonie Cornips and Karen P. Corrigan, 233–64. Amsterdam: John Benjamins.

Barbiers, Sjef. 2009. Locus and limits of syntactic microvariation. *Lingua* 119.1607–23. DOI: 10.1016/j.lingua.2008.09.013.

Barbiers, Sjef, et al. 2006. *Dynamische Syntactische Atlas van de Nederlandse Dialecten (DynaSAND)*. Amsterdam: Meertens Institute. Online: http://www.meertens.knaw.nl/sand/.

Barbiers, Sjef, and Hans Bennis. 2010. De plaats van het werkwoord in zuid en noord. *Voor Magda: Artikelen voor Magda Devos bij haar afscheid van de Universiteit Gent*, ed. by Johan de Caluwe and Jacques van Keymeulen, 25–42. Ghent: Academia.

Barbiers, Sjef; Hans Bennis; Gunther De Vogelaer; Magda Devos; and Margreet van der Ham. 2005. *Syntactische Atlas van de Nederlandse Dialecten, Deel I*. Amsterdam: Amsterdam University Press.

Barbiers, Sjef; Hans Bennis; and Lotte Dros-Hendriks. 2018. Merging verb cluster variation. *Linguistic Variation* 18.144–96. DOI: 10.1075/lv.00008.bar.

Barbiers, Sjef; Johan van der Auwera; Hans Bennis; Eefje Boef; Gunther De Vogelaer; and Margreet van der Ham. 2008. *Syntactische Atlas van de Nederlandse Dialecten, Deel II*. Amsterdam: Amsterdam University Press.

Bennis, Hans, and Evie Coussé. 2012. Werkwoordsvolgorde in de rechterperiferie van de Nederlandse zin: Inleiding. *Taal en Tongval* 64.1–10. DOI: 10.5117/TET2012.1.BENN.

Bloem, Jelke; Arjen Versloot; and Fred Weerman. 2017. Verbal cluster order and processing complexity. *Language Sciences* 60.94–119. DOI: 10.1016/j.langsci.2016.10.009.

Bobaljik, Jonathan David. 2004. Clustering theories. *Verb clusters: A study of Hungarian, German and Dutch*, ed. by Katalin É. Kiss and Henk van Riemsdijk, 121–45. Amsterdam: John Benjamins.

Borer, Hagit. 1984. *Parametric syntax*. Dordrecht: Foris.

Chambers, J. K., and Peter Trudgill. 1998. *Dialectology*. 2nd edn. Cambridge: Cambridge University Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2007. Approaching UG from below. *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1–30. Berlin: Mouton de Gruyter.

Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36.315–32. DOI: 10.1162/0024389054396917.

Cohen, Jacob. 1962. The statistical power of abnormal–social psychological research: A review. *Journal of Abnormal and Social Psychology* 65.145–53. DOI: 10.1037/h0045186.

Cornips, Leonie, and Willy Jongenburger. 2001. Het design en de methodologie van het SAND-project. *Nederlandse Taalkunde* 6.215–32.

Cornips, Leonie, and Cecilia Poletto. 2005. On standardising syntactic elicitation techniques (part 1). *Lingua* 115.939–57. DOI: 10.1016/j.lingua.2003.11.004.

Cornips, Leonie, and Cecilia Poletto. 2007. Field linguistics meets formal research: How a microcomparative view can deepen our theoretical investigation Part 2 (sentential negation). Amsterdam: Meertens Institute, and Padua: ISTC-CNR, ms.

Coussé, Evie. 2008. Motivaties voor woordvolgorde: Een diachrone studie van werkwoordsvolgorde in het Nederlands. Ghent: University of Ghent master's thesis.

Daan, Jo, and Dirk Blok. 1969. *Van Randstad tot Landrand; toelichting bij de kaart: Dialecten en Naamkunde*. (Bijdragen en mededelingen der Dialectencommissie van de Koninklijke Nederlandse Akademie van Wetenschappen te Amsterdam 37.) Amsterdam: North-Holland.

Daelemans, Walter, and Antal van den Bosch. 2005. *Memory-based language processing*. Cambridge: Cambridge University Press.

Daelemans, Walter; Jakub Zavrel; Ko van der Sloot; and Antal van den Bosch. 2010. *TiMBL: Tilburg memory based learner, version 6.3: Reference guide*. (Technical report ILK 10-01.) Tilburg: Induction of Linguistic Knowledge Research Group, Tilburg Centre for Cognition and Communication.

De Sutter, Gert. 2009. Towards a multivariate model of grammar: The case of word order variation in Dutch clause final verb clusters. *Describing and modeling variation in* *grammar*, ed. by Andreas Dufter, Jürg Fleischer, and Guido Seiler, 225–54. Berlin: Mouton de Gruyter.

Evers, Arnold. 1975. *The transformational cycle in Dutch and German*. Bloomington: Indiana University Linguistics Club.

Fawcett, T. 2004. ROC graphs: Notes and practical considerations for researchers. Technical report HPL-2003-4. Palo Alto: Hewlett Packard Labs.

Fellows, Ian. 2014. wordcloud: Word clouds. R package. Online: https://cran.r-project.org/package=wordcloud.

Greenacre, Michael. 2007. *Correspondence analysis in practice*. 2nd edn. London: Chapman & Hall.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, ed. by Joseph Greenberg, 73–113. Cambridge, MA: MIT Press.

Haegeman, Liliane, and Henk van Riemsdijk. 1986. Verb projection raising, scope, and the typology of rules affecting verbs. *Linguistic Inquiry* 17.417–66. Online: https://www.jstor.org/stable/4178499.

Hinskens, Frans, and Johan Taeldeman (eds.) 2013. *Language and space: An international handbook of linguistic variation. Vol. 3: Dutch*. Berlin: De Gruyter Mouton.

Husson, François, and Julie Josse. 2013. missMDA: Handling missing values with/in multivariate data analysis (principal component methods). R package version 1.11. Online: http://CRAN.R-project.org/package=missMDA.

Husson, François; Julie Josse; Sébastien Lê; and Jeremy Mazet. 2014. FactoMineR: Multivariate exploratory data analysis and data mining with R. R package version 1.26. Online: http://CRAN.R-project.org/package=FactoMineR.

Husson, François; Sébastien Lê; and Jérôme Pagès. 2011. *Exploratory multivariate analysis by example using R*. Boca Raton: CRC Press.

Josse, Julie; Marie Chavent; Beno[i]t Liquet; and François Husson. 2012. Handling missing values with regularized iterative multiple correspondence analysis. *Journal of Classification* 29.91–116. DOI: 10.1007/s00357-012-9097-0.

Kayne, Richard S. 2000. *Parameters and universals*. Oxford: Oxford University Press.

Kayne, Richard S. 2005. Silent years, silent hours. *Movement and silence*, 241–60. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195179163.003.0010.

Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Merlo, Paola. 2015. Predicting word order universals. *Journal of Language Modelling* 3.317–44. DOI: 10.15398/jlm.v3i2.112.

Oksanen, Jari; F. Guillaume Blanchet; Michael Friendly; Roeland Kindt; Pierre Legendre; Dan McGlinn; Peter R. Minchin; R. B. O'Hara; Gavin L. Simpson; Peter Solymos; M. Henry H. Stevens; Eduard Szoecs; and Helene Wagner. 2018. vegan: Community ecology package. R package. Online: https://cran.r-project.org/package=vegan.

Pauwels, Anita. 1953. De plaats van hulpwerkwoord, verleden deelwoord en infinitief in de Nederlandse bijzin. Deel I: Tekst. Deel II: Kaarten. Leuven: KU Leuven master's thesis.

Quinlan, J. Ross. 1993. *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufmann.

R Core Team. 2014. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: http://www.R-project.org/.

Richardson, John T. E. 2011. Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review* 6.135–47. DOI: 10.1016/j.edurev.2010.12.001.

Schmid, Tanja. 2005. *Infinitival syntax: Infinitivus Pro Participio as a repair strategy*. Amsterdam: John Benjamins.

Schmid, Tanja, and Ralf Vogel. 2004. Dialectal variation in German 3-verb clusters: A surface-oriented optimality theoretic account. *The Journal of Comparative Germanic Linguistics* 7.235–74. DOI: 10.1023/B:JCOM.0000016639.53619.94.

Spruit, Marco René. 2008. *Quantitative perspectives on syntactic variation in Dutch dialects*. Amsterdam: Universiteit van Amsterdam dissertation.

Stroop, Jan. 1970. Systeem in gesproken werkwoordgroepen. *Taal en Tongval* 22.128–47.
van den Berg, Berend. 1949. De plaats van het hulpwerkwoord in de voltooide tijden in de Nederlandse bijzin. *Taal en Tongval* 1.155–65.
Willemyns, Roland. 2006. The Low Countries. *Sociolinguistics: An international handbook of the science of language and society*, vol. 3, ed. by Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, and Peter Trudgill, 1758–65. Berlin: De Gruyter.
Wurmbrand, Susi. 2017. Verb clusters, verb raising, and restructuring. *The Wiley Blackwell companion to syntax*, 2nd edn., ed. by Martin Everaert and Henk van Riemsdijk. Hoboken, NJ: Wiley-Blackwell. DOI: 10.1002/9781118358733.wbsyncom103.
Zwart, C. Jan-Wouter. 1993. *Dutch syntax: A minimalist approach*. Groningen: University of Groningen dissertation.
Zwart, C. Jan-Wouter. 2015. Top-down derivation, recursion, and the model of grammar. *Syntactic complexity across interfaces*, ed. by Andreas Trotzke and Josef Bayer, 25–42. Berlin: De Gruyter Mouton.

van Craenenbroeck
KU Leuven/CRISSP/KNAW Meertens Instituut
Blijde-Inkomststraat 21
3000 Leuven, Belgium
[jeroen.vancraenenbroeck@kuleuven.be]

van Koppen
KNAW Meertens Instituut/Utrecht University/UiL-OTS
PO Box 10855
1001 EW Amsterdam, The Netherlands
[marjo.van.koppen@meertens.knaw.nl]

van den Bosch
KNAW Meertens Instituut
PO Box 10855
1001 EW Amsterdam, The Netherlands
[antal.van.den.bosch@meertens.knaw.nl]

[Received 28 August 2017;
revision invited 21 February 2018;
revision received 21 September 2018;
revision invited 15 January 2019;
revision received 28 January 2019;
accepted 2 February 2019]