



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

Empirical syntax: idiolectal variability in two- and three-verb clusters in regional standard Dutch and Dutch dialects

Cornips, L.

published in

Describing and modeling variation in grammar
2009

document version

Publisher's PDF, also known as Version of record

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Cornips, L. (2009). Empirical syntax: idiolectal variability in two- and three-verb clusters in regional standard Dutch and Dutch dialects. In A. Dufter, J. Fleischer, & G. Seiler (Eds.), *Describing and modeling variation in grammar* (Vol. Studies and Monographs 204, pp. 203-224). (Trends in Linguistics). De Gruyter Mouton.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
pure@knaw.nl

Empirical syntax: Idiolectal variability in two- and three-verb clusters in regional standard Dutch and Dutch dialects

Leonie Cornips

Abstract. The goal of this paper is to discuss idiolectal variability in a regional Dutch variety and Dutch dialects concerning word order in the two- and three-verb cluster (right periphery) on the basis of two solid empirical studies; that is, (i) a corpus 33.5 hours of recorded spontaneous speech of 67 speakers in the city of Heerlen in the southeastern part of the Netherlands and (ii) a corpus of written relative acceptability judgments collected for the *Syntactic Atlas of the Dutch Dialects*.

The subjects in both corpora reveal idiolectal variability. One of the variants included in the idiolectal variability always concerns standard Dutch word order, signalling a very advanced process of standardization. The degree of individual variation and word order preferences depend on the type of auxiliary and morphology of the deepest embedded verb. In this paper, it is argued that a basic head-initial structure with optional verb raising is capable of accounting for the empirical results.

1. Introduction*

Recent years have seen a growing interest in the interdisciplinary field of dialectology, sociolinguistics and formal syntax in the domain of microvariation; that is, syntactic variation between closely related dialects in geographical and/or social space (Barbiers, Cornips and van der Kleij (eds.) 2002; Kortmann (ed.) 2004; Cornips and Corrigan 2005; Cornips and Corrigan (eds.) 2005). This interdisciplinary field provides an opportunity to perform theoretical research on the basis of solid empirical studies. At the moment, not many large-scale studies exist that present minimal directions in which local dialects can vary and in which the role of fieldwork is discussed in addition to the procedures through which data are elicited, ordered and analyzed (cf. Cornips and Poletto 2005; Cornips 2006). In recent years, though, many researchers claim that a bridge between variation and theory is possible; that is, it is assumed that there exists a fairly direct connection between grammar and language usage (Meechan

and Foley 1995: 82; Pintzuk 1995; Kroch 1998; Cornips and Corrigan 2005, Cornips 2006). Wilson and Henry (1998: 8) phrase the merits of an overall approach as follows: "[...] we may be able better to understand language variation and change as they are driven by social factors but constrained (at one level) by the nature of possible grammars."

The goal of this paper is to discuss idiolectal variability, i.e., individual speaker variation concerning word order in the verbal cluster (right periphery) on the basis of two solid empirical studies. The first study is based on 33.5 hours of recorded spontaneous speech of 67 speakers in one community, namely Heerlen, located in the province of Limburg in the southeastern part of the Netherlands (cf. Cornips 1994, 1998, 2006). The second study involves written acceptability judgments collected for the project of the *Syntactic Atlas of the Dutch Dialects* (so-called SAND project; Barbiers, Cornips and Kunst 2007) that has resulted in a large-scale geographical investigation of Dutch dialects in the Netherlands and Belgium. This corpus contains written judgments of 370 native speakers of local dialects throughout the whole area. By searching these corpora, possible intra- and inter-speaker variation can be distinguished.

The aim of this paper is to show that speakers reveal idiolectal variability with respect to different word orders per type of two- and three-verb cluster. Second, spontaneous production data from speakers in Heerlen provides such orderly heterogeneity that they can be considered as members of one sociolinguistic unit (Guy 2004). Third, the degree of individual variation and word order preferences depend on the type of auxiliary and morphology of the deepest embedded verb; that is, the verbal clusters in which the auxiliary (perfective/passive) selects a past participle allows for more variability than those in which it (modal) selects an infinitive. Finally, one of the variants included in the idiolectal variability always concerns the standard Dutch word order, signalling a very advanced process of standardization.

1.1 A comparison of both types of corpora

In general, both types of corpora – e.g., spontaneous speech and acceptability judgments – do not reflect the structure of the language directly since they deal with various intruding factors (cf. Cornips and Poletto 2005). One issue comes to the fore when comparing both types of corpora: What is the exact relationship between acceptability judgments and spontaneous speech data? In particular, speakers may judge certain types of constructions as acceptable despite the fact that they never surface empirically. Thus, on the one hand, subjects may share introspective judgments regarding constructions that are grammatically

possible intuitively but are not recoverable from observation alone (cf. Labov 1975; Cornips and Corrigan 2005); on the other, it is never certain whether a non-occurring structure in spontaneous speech is due to ungrammaticality or chance. To this end, a comparison between the two types of corpora deepens our insight in the relationship between naturally-occurring data and acceptability judgments – especially as to why these do not always converge. Written acceptability judgments also make it possible to derive a scale of acceptability concerning possible word orders in the verbal cluster, which would never arise through the observation of spontaneous speech alone. What is most important, though, is that the Heerlen Dutch corpus provides insight into the phenomenon of social microvariation; that is, syntactic variation between closely-related (clustering of) individual grammars in social space (according to independent social speaker variables, the geographical variable being kept constant), whereas the SAND corpus gives us understanding in variation between closely-related (clustering of) individual grammars in geographical space (the geographical grid is the independent variable, social variables are kept constant).

The present paper is organized as follows. In the second and third part, regional spontaneous spoken Dutch and acceptability judgments of 370 speakers of a local dialect throughout the Netherlands and Belgium will be examined regarding possible word order in two-verb and three-verb clusters, respectively. The final part concludes the paper.

2. Heerlen Dutch spontaneous speech corpus: idiolectal variability in two-verb clusters

The corpus of Heerlen Dutch consists of 33.5 recorded hours of spontaneous speech between two speakers who did not know each other but belonged to one cell (in-group conversation) with respect to their age, language background and level of education. The recordings took place at one of the speakers' homes between 1989 and 1990. A total of 67 male speakers participated in this survey (cf. Cornips 1994, 1998).¹ Heerlen is situated in the province of Limburg in the southeast of the Netherlands, near the Belgian and German border. Heerlen Dutch as a *regional standard* variety is the result of a process of abrupt language shift in the beginning of the twentieth century with the local dialect as the source and standard Dutch as the target language. The language shift was due to the explosive growth of the coal mining industry that attracted numerous workers from outside Limburg and the Netherlands to Heerlen and its surroundings. As a result of massive immigration, local inhabitants became a minority in their community within a time span of thirty years.

This corpus of 67 speakers in one community contains 1,230 tokens of two-verb clusters, which is considered a high frequency in sociolinguistic research (Cornips 2006). The corpus reveals the two possible types of two-verb clusters, namely (i) a verbal cluster containing a perfect or passive auxiliary combined with a past participle, as illustrated in (1) and (ii) a cluster involving a modal combined with an infinitive, as in (2):^{2,3}

- 1: Henk
 (1) a. *dat wij die ruit hadden ingetrapt*
 that we that window had kicked.PART
 15: Jan
 b. *want als daar iemand gearresteerd wordt*
 because when there someone arrested.PART is
- 3: Jansen
 (2) a. *dus die een beetje lezen kunnen*
 thus those a bit read.INF can
 3: Jansen
 b. *die dat ... redelijk kunnen opbrengen*
 those that reasonably can yield.INF

There are 558 tokens of a verbal cluster containing a perfective or passive auxiliary combined with a past participle, as in (1a) and (1b), respectively. What's more, this corpus also contains 554 tokens of a two-verb cluster involving a modal combined with an infinitive, as in (2).

2.1 Word order variation in the perfective/passive two-verb cluster

In the late 1960s, Labov and others (Cedergren and Sankoff 1974) introduced the concept of variable rule as an extension of the optional rule used in formal syntax to distinguish variable elements of grammar. The variable rule differs considerably from the optional rule in that probability was built into the former and its application was conditioned by both social and stylistic, i.e., external dimensions of language use along with linguistic, i.e., internal dimensions. The variable rule consists of syntactic *variants* that are "alternate ways of saying 'the same' thing" (Labov 1972: 118). We can illustrate the make up of a specific syntactic variable by the utterances in (3). When we examine a two-verb cluster containing a perfective auxiliary combined with a past participle, we find that one and the same speaker (Peter) may produce both orders even while maintaining

the same level of speech style. The numbers 1 and 2 in (3) refer to the status of the verbs in terms of embedding, i.e., the verb 2 is structurally more embedded than verb 1:

- 15: Peter
 (3) a. *waar [...] X heeft gewoond* 1-2 order
 where X has lived.PART
 b. *waar X gewoond heeft* 2-1 order
 where X lived.PART has

In traditional analyses, Dutch is hypothesized to be a SOV language with a head-final verb phrase. Consequently, in earlier derivational models of formal syntax, the order 1-2 in the verbal cluster in (3a) does not reveal a basic order but is derived by (verb) movement, whereas the order 2-1 in (3b) represents the deep or underlying structure. In sociolinguistics, the fact that (i) (3a) and (3b) were considered different surface realizations of the same underlying structure and (ii) these variants were semantically equivalent (Labov 1972: 188) was the prerequisite to consider these variants as constituting the linguistic variable (cf. Cornips and Corrigan 2005, Cornips 2005).

The Heerlen Dutch corpus provides us an excellent opportunity to examine the interplay between idiolectal (I-language) and community variability (E-language). Let us first examine the variable use of the 1-2 (cf. (3a)) and 2-1 order (cf. (3b)) within the verbal cluster containing a perfective/passive auxiliary and a past participle. Table 1 reveals that at the level of the group of speakers, i.e., community grammar, there is considerable variability in usage:

Table 1. The use of the orders 1-2 or 2-1 in the verbal cluster containing a perfective/passive auxiliary and past participle at the level of group of speakers in Heerlen

order of the two verbs	tokens	%
2-1	376	67.4
V _{Part} +AUX		
1-2	182	32.6
AUX+V _{Part}		
total	558	100

A closer examination shows that it is not the case that this variation is due to two homogeneous subgroups in the community using 2-1 order *or* 1-2 order. Instead, Table 2 reveals that fifty out of 67 speakers (75%) use both orders and,

hence, reveal a 'choice' or idiolectal variability. Importantly, only one speaker reveals a categorical 1-2 order, whereas sixteen speakers reveal a categorical 2-1 order:

Table 2. Categorical and variable use of 1-2 and/or 2-1 order within the verbal cluster containing a perfective/passive auxiliary and past participle in Heerlen at the level of the individual speaker

	<i>n</i> speakers (67)
categorical use 1-2	1
categorical use 2-1	16
variable use 1-2/2-1	50

Although there is considerable intra- and inter-speaker variation in Heerlen, word order differences do not have the value of a sociolinguistic marker, that is to say, no significant social correlations arise when the speaker data are divided according to the independent variables language background, i.e., whether the subject is a local dialect or Heerlen Dutch speaker from birth onwards, age and education/occupation (see Cornips and Ribbert 2006).

Let us now consider a second type of verbal cluster containing two verbs in Heerlen Dutch, e.g., a verbal cluster containing a modal or aspectual auxiliary combined with an infinitive.

2.2 Word order variation in the modal/aspectual two-verb cluster

A two-verb cluster may also contain a modal such as *kunnen* 'can', *moeten* 'must', *willen* 'want', *mogen* 'may', *zullen* 'shall' or an aspectual auxiliary such as *doen* 'do', *laten* 'let', *gaan* 'go' and *blijven* 'remain'. The modal selects an infinitive instead of past participle, as illustrated in (4):

- (4) *dat hij de auto gaat_{ASP} / wil_{MOD} stelen_{INF}* 1-2 order
 that he the car goes / wants steal
 'that he wants to steal the car'

In contrast to the preferred order 2-1 in the perfective/passive verbal cluster (cf. Table 1), the order 1-2 is now the most preferred one and even nearly categorical in the Heerlen Dutch modal verbal cluster:

Table 3. The use of the orders 1-2 or 2-1 in the modal two-verb cluster combined with an infinitive at the level of group of speakers in Heerlen

order of the two verbs	tokens	%
2-1	14	.03
V _{Inf} -MOD		
1-2	540	99.97
MOD-V _{Inf} (cf. (4))		
total	554	100

The same holds for the aspectual two-verb cluster, as shown in Table 4:

Table 4. The use of the orders 1-2 or 2-1 in the aspectual two verb cluster combined with an infinitive at the level of group of speakers in Heerlen

order of the two verbs	tokens	%
2-1	1	0
V _{Inf} -ASP		
1-2	117	100
ASP-V _{Inf}		
total	118	100

The individual level reflects the considerable decrease in variable use of the two possible orders at the community level. Hence, only twelve speakers show variable use between the 1-2 and 2-1 orders in contrast to the fifty speakers who are displayed in Table 2. The twelve speakers together produce only 15 tokens (see Tables 3 and 4) of the 2-1 order:

Table 5. Categorical and variable use of 1-2 and/or 2-1 order within verbal cluster containing a modal/aspectual and infinitive in Heerlen at the level of the individual speaker

	<i>n</i> speakers (67)
categorical use 1-2	55
categorical use 2-1	0
variable use 1-2/2-1	12

Finally, unlike the perfective/passive cluster, the modal/aspectual cluster includes a categorical 1-2 order but not a categorical 2-1 order. This is an interesting empirical result since in the generative literature it is argued (on the basis of intuitions) that both orders are freely available in standard Dutch regardless the type of the two-verb cluster (see Den Besten and Broekhuis 1989: 81 and

references cited; Zwart 1996: 232; Wurmbbrand 2006: 237). Therefore, let us consider the variability between the two orders in each type of verbal cluster in more detail.

2.3 Dependencies between the two types of verbal clusters

The Heerlen Dutch speakers reveal an implicational relationship between the two possible orders at the individual and community level. This implicational relation is different regarding the type of auxiliary (perfective versus modal/aspectual) and the morphology of the least embedded verb (past participle versus infinitive). With respect to the perfective cluster, it appears that (with the exception of one speaker) the 1-2 order only arises in combination with the 2-1 order at the level of the individual speaker. Thus, the 1-2 order never emerges categorically but is dependent on the occurrence of the 2-1 order. Similarly, the 2-1 order in the modal/aspectual cluster never emerges categorically, that is to say, it only combines with the 1-2 order at the level of the individual and community:

perfective/passive auxiliary + past participle	modal/aspectual + infinitive
implicational relation	implicational relation
1-2 iff 2-1/1-2	2-1 iff 1-2/2-1

Figure 1. Implicational order relation in both types of verbal cluster in embedded clauses

Consequently, the differences between individual speakers are not just random, but systematic in their categorical and variable uses along different morphosyntactic features displaying *orderly heterogeneity*. Thus, syntactic variation is not constant 'throughout', but individual speakers reveal consistent behavior according to the community pattern (Guy 2004).

Let us now focus on the range of individual grammars in Heerlen Dutch on the basis of 1,230 occurrences of two-verb clusters in spontaneous speech, 67 speakers and two types of verbal clusters, i.e., perfective/passive + past participle and modal/aspectual + infinitive. Figure 2 reveals that 67 speakers produce six different grammars that take the form *yes* or *no* for each order in each type of verbal cluster. Importantly, Figure 2 also displays that the large majority of the speakers, e.g., 40 out of 67 (60%), produces both orders in the perfective/passive cluster and a categorical 1-2 order in the modal/aspectual cluster. Further, the patterns in Figure 2 confirm that (i) the perfective/passive cluster allows more individual variety than the modal/aspectual cluster ($n = 51$ and $n = 13$, respectively), and (ii) the perfective/passive cluster prefers the 2-1 order

as a single order ($n = 15$, 1-2 order; $n = 1$), whereas the modal/aspectual cluster prefers the 1-2 order as the single one ($n = 53$):

	perf./pass. cluster		modal/aspectual cluster	
	2-1	1-2	2-1	1-2
<i>individual grammars</i>				
<i>(n = 67)</i>				
n = 40 60%	yes	yes	no	yes
n = 12 17%	yes	no	no	yes
n = 10 14%	yes	yes	yes	yes
n = 3 5%	yes	no	yes	yes
n = 1 2%	yes	yes	yes	no
n = 1 2%	no	yes	no	yes
<i>non-occurrence</i>				
*	yes	no	yes	no
*	no	yes	yes	no
*	no	yes	yes	yes

Figure 2. The individual patterns regarding the 2-1- and 1-2 orders in the perfective/passive and aspectual/modal cluster

2.4 Analysis

Importantly, the degree of idiolectal variability appears to be conditioned by morphosyntactic features, e.g. the morphology of the deepest embedded verb and the type of auxiliary. From the above, it is clear that quantitative results based on an empirical study have the potential not only to assess but also to refine theoretical conceptions of linguistic structure. Hence, three structural analyses are possible depending on the model of syntax one employs. First, the classical analysis (as already discussed in 2.1) considers the 1-2 order as derived from the basic 2-1 structure by verb movement. Second, in a more recent framework all structures are head-first and movement is leftwards (Kayne 1994). If so, then 1-2 is the order from which 2-1 is derived by movement. Finally, from a Minimalist perspective, both structures can be available without a movement operation; in this case, it is assumed that a syntactic projection can be head-final (2-1 order) or head-initial (1-2 order): a head can either precede or follow the projection with which it is merged. The patterns of usage, in particular, the fact that only one order is categorical and that there exists a dependency relationship between the two orders provide evidence that it cannot be the case that both structures

are freely available to the individual grammar. Hence, in that case, it is expected that both orders would emerge as categorical in the community. Instead, one of the orders has to be derived by movement. I argue that a basic head-initial (1-2 order) structure is capable of accounting for the empirical results.⁴ A basic 1-2 order corresponds to the majority of the Heerlen Dutch individual grammars displaying a categorical 1-2 order, that is, 53 individual grammars (79%) in the modal/aspectual cluster (see Figure 2). Hence, the number of movements is kept to a minimum.

The larger variability in the perfective/passive cluster can be accounted for by verb raising of the past participle; the lower verb raises and attaches to the next verb higher up. Verb raising of the participle remains optional, resulting in the 2-1 order. The optional raising can be accounted for by assuming that the participle may be a verbal or adjectival element. If adjectival, it has to undergo verb raising; if verbal, the participle remains. In Dutch, participles can be used as adjectival predicates. The passive auxiliary *zijn* 'to have-been' is homophonous to the copular verb *zijn* 'to be'. The example in (5) is ambiguous between an adjectival (state) reading and a verbal (activity) reading (cf. Broekhuis 1999: 404 and 406):

- (5) *dat de muur versierd is*
 that the wall decorated is
copular construction: 'The wall has been decorated (for years).'
passive construction: 'The wall was decorated (yesterday).'

Further, in standard and Heerlen Dutch, adjectives always precede the verb(s) in clause-final position:

- (6) *dat Jan ziek is/ *is ziek*
 that John ill is/ is ill

Hence, the contrast in grammaticality between (5) and (7) and between (7a) and (7b) shows that if the verb in the perfective/passive cluster is an adjectival element, it has to undergo verb raising. This results in a 2-1 order, as in (7b):

- (7) a. *dat de muur (gisteren/*? al jaren) is versierd.*
 that the wall yesterday/for years is decorated
 b. *dat de muur (gisteren/al jaren) versierd is.*
 that the wall yesterday/for years decorated is

In this way, the differences between participle two-verb and three-verb clusters and infinitival ones with respect to word order can be accounted for.

3. Heerlen Dutch spontaneous speech corpus: idiolectal variability in three-verb clusters

Let us now consider the idiolectal variability in three-verb clusters in Heerlen Dutch. I focus on two types of three-verb clusters, namely the MOD-MOD-V_{Inf} and the MOD-AUX-V_{Part} as illustrated in (8a) and (8b), respectively. In doing so, a maximal comparison becomes possible with the elicited structures in the SAND project (see below, Section 4).

- (8) a. *dat we met haar moeten kunnen praten* 1-2-3
 that we to her must can talk
 b. *dat ik dat moet hebben gezegd* 1-2-3
 that I that must have said

The spontaneous speech corpus of Heerlen Dutch consists of 251 three-verb clusters of which 120 are MOD-MOD-V_{Inf} and 131 MOD-AUX-V_{Part} clusters. This spontaneous speech corpus has its limitations since the speakers prefer to produce three-verb clusters overwhelmingly more often with the finite verb in main clauses (82%) than in embedded clauses (18%).

3.1 The finite verb is in the main clause

In Dutch, the position of the highest verb =MOD_{Fin} - in main clauses is fixed; in fact, we are dealing with two-verb clusters here. If we consider these tokens first, then the modal cluster only allows one order, namely 1-2, whereas in the perfective cluster the 2-1 order occurs 71% (28/40) and the 1-2 order 29% (12/40). This ratio is a striking confirmation of the one attested already in the two-verb clusters mentioned above (see Table 1). Further, when the auxiliary is a passive one, the 2-1 order is overwhelmingly more preferred than the 1-2 order, thus, 91% (62/68) and 9% (6/68), respectively.

	Main clauses (MOD ₁ = V2)		
(MOD ₁ -)MOD-V _{Inf}	1-2	97/97	100%
(MOD ₁ -)AUX-V _{Part}	1-2	12/40	30%
	2-1	28/40	70%
(MOD ₁ -)AUX _{Pass} -V _{Part}	1-2	6/68	9%
	2-1	62/68	91%

Figure 3. The word orders in two types of two-verb clusters (with MOD_{Fin} in second position) in spontaneous Heerlen Dutch: MOD-MOD-V_{Inf} and MOD-AUX-V_{Part}

3.2 The finite verb is in the embedded clause

The attested three-verb clusters in the embedded clauses are presented in Figure 4. Theoretically, there is more variation space possible in the three-verb clusters as compared with two-verb clusters – namely six possible word orders versus two, respectively. However, the 67 speakers do not reveal any variation in the MOD–MOD–V_{Inf} cluster at the individual and community level since the ascending order is categorical in embedded clauses. Thus, the preference for a categorical 1–2 order in the modal two-verb cluster is obligatory with its three-verb counterpart. Unfortunately, the MOD–AUX–V_{Part} cluster has only two tokens revealing both a 1–3–2 and 3–1–2 order. The MOD–AUX_{Pass}–V_{Part} with its twenty-one tokens, however, shows a clear preference for the 3–1–2 order:

	order	embedded clause	total
MOD–MOD–V _{Inf}	1–2–3	23/23	100%
MOD–AUX–V _{Part}	1–3–2	1/2	50%
	3–1–2	1/2	50%
MOD–AUX _{Pass} –V _{Part}	1–2–3	1/21	5%
	1–3–2	1/21	5%
	3–1–2	19/21	90%

Figure 4. The word orders in two types of three-verb clusters in spontaneous Heerlen Dutch: MOD–MOD–V_{Inf} and MOD–AUX–V_{Part}

Figure 4 shows that the participle undergoes verb raising in the MOD–AUX–V_{Part} cluster, resulting in the 1–3–2 and 3–1–2 orders. Interestingly, Heerlen Dutch, although a regional variety, is similar to standard Dutch with respect to the different orders in the two types of clusters. In standard Dutch, the 1–2–3 order is the unmarked one in the MOD–MOD–V_{Inf} cluster whereas the orders 1–2–3, 1–3–2 and 3–1–2 are the unmarked ones in the MOD–AUX–V_{Part} cluster (Wurmbrand 2006: 324).

The differences between the MOD–MOD–V_{Inf} and MOD–AUX–V_{Part} regarding variability and word order are confirmed by the few speakers who produce a four-verb cluster in Heerlen Dutch. There is one token of a MOD–MOD–MOD–V_{Inf} with an ascending 1–2–3–4 order and four utterances of MOD–MOD–AUX–V_{Part} cluster. Two of them show a preference of the participle before its selecting verb 1–2–4–3 order and there are two instances of a 4–1–2–3 order.

As mentioned earlier, although there is more variation space in the three-verb than in the two-verb clusters, speakers do not reveal more idiolectal variability

in the former, not even in the MOD–AUX–V_{Part} cluster. However, it is plausible that this is due to the limitations of an empirical corpus; hence, the number of MOD–AUX–V_{Part} clusters is simply too small to conclude anything with respect to variability:

MOD–MOD–V _{Inf}	1–2–3	<i>n</i> = 43
MOD–AUX–V _{Part}	1–3–2	<i>n</i> = 1
	3–1–2	<i>n</i> = 1
MOD–AUX _{Pass} –V _{Part}	1–2–3	<i>n</i> = 1
	1–3–2	<i>n</i> = 1
	3–1–2	<i>n</i> = 19

Figure 5. Idiolectal variability in the three-verb cluster, *n* = 67

3.3 Analysis

It is clear that the degree of variability appears to be conditioned by the type of the auxiliary and the morphology of the deepest embedded verb, namely auxiliary versus modal and participle versus infinitive, respectively. First, when the deepest embedded verb is an infinitive, all speakers produce one order categorically, namely 1–2–3. Second, when the deepest embedded verb is a past participle, more variation in word order arises at the community level but there is no idiolectal variability.

In any analysis, a distinction must be made between a cluster containing a participle or an infinitive as the deepest embedded verb. Further, it should be able to account for the fact that (i) the MOD–MOD–V_{Inf} cluster does not allow other orders than 1–2–3 and (ii) in the MOD–AUX–V_{Part} cluster only a specific combination of orders arises, namely 1–2–3, 1–3–2 and 3–1–2. Thus, we have to put forward a grammatical factor causing (i) the preferred combination of 1–2–3, 1–3–2 and 3–1–2 and (ii) excluding 2–3–1, 2–1–3 and 3–2–1.

The basic head-initial structure 1–2–3 order explains the empirical results of the MOD–MOD–V_{Inf} cluster most adequately. Hence, the 1–2–3 order is categorical, so movement operations are kept zero. The co-occurrence of word orders in the MOD–AUX–V_{Part} can be accounted for by verb raising of the past participle similar to the one for the perfective two-verb cluster. Verb raising of the participle remains optional and it is free to occur anywhere in the cluster, resulting in the 1–3–2 and 3–1–2 orders. Note that verb raising of the participle and a head-initial structure exclude other theoretically possible word orders in Heerlen Dutch, namely 2–3–1, 2–1–3 and 3–2–1.

In order to see whether we are dealing with a general property of verbal clusters or a rather idiosyncratic property of the Heerlen Dutch and/or standard Dutch grammar, let us now consider the SAND corpus.

4. The empirical basis of verbal clusters: written elicited acceptability judgments corpus (Syntactic Atlas of the Dutch Dialects SAND)

4.1 Word order variation in perfective two verb-cluster in Dutch dialects in the Netherlands and Belgium

The corpus of the *Syntactic Atlas of the Dutch Dialects* is the result of a large-scale dialect syntax project carried out from 2000 through 2003 in the Netherlands and Belgium (cf. Cornips and Jongenburger 2001; Cornips and Poletto 2005; Barbiers, Cornips and Kunst 2007). In this paper, judgments elicited by means of written questionnaires will be discussed. Different types of three-verb clusters were administered and filled in by 370 native speakers of local dialects throughout the Netherlands and Flanders (Belgium) (see also Barbiers 2005). All social variables of the speakers were recognized and controlled for. The subjects were homogenized as much as possible with respect to age, mobility, functional domains in which the dialect is used, and education. As a consequence, it is very plausible that the detected variation is due to geographical factors alone. A subpart of these judgments is discussed in this paper (cf. Barbiers, Cornips and Kunst 2007).

With respect to the two-verb cluster, only the perfective one was offered in the written questionnaire. The subjects had to translate the following perfective cluster in the 2-1 order from the standard into their local dialect variety:

- (9) *Weten jullie dat we Jan op de markt gezien_{Part} hebben_{AUX}?*
know you that we Jan on the market seen have

Forty subjects out of 355 (11%) have translated (9) in the 1-2 order instead of the 2-1 order in the input, including three subjects who translate (9) in the two possible orders:

- (10) *Wete gellie nog dâ we Jan op de mert hemme*
know you PART. that we Jan on the market have
gezien, [...] gezien hemme
seen seen have (Nijmegen)

Since all literature reports that standard Dutch allows two orders, as in (10), it is plausible that the fact that only 11% of the subjects reports a 1-2 order is due to a repetition-effect; that is, the word order in the instruction is just copied into the local dialect (cf. Cornips and Poletto 2005). Thus, a translation task is not a felicitous one in eliciting variability.

4.2 Word order variation in three-verb clusters in Dutch dialects in the Netherlands and Belgium

A more appropriate syntactic elicitation methodology than the translation task was used to collect data about variable word orders in several types of three-verb clusters, namely a written indirect relative judgment task. In this paper, the same types of three-verb clusters as in Heerlen Dutch will be discussed, namely MOD-MOD-V_{Inf} and MOD-AUX-V_{Part}, as in (11a) and (11b), respectively:

- (11) a. *Ik weet dat Jan hard moet kunnen werken*
I know that Jan hard must can work
b. *Jan weet dat hij voor drie uur de wagen moet*
Jan knows that he before three o'clock the car must
hebben gemaakt
have repaired

The task required the 370 subjects to judge all six possible orders in the three-verb cluster from most to least common in their dialect on a five-point scale (representing *, ?*, ??, ?, ok). The subjects were first asked: 'do you encounter the following sentences in the local dialect?' and, if so, 'which sentence(s) do you consider to be the most (highest value = 5) or the least common (highest value = 1) one in your local dialect?', as illustrated in Figure 6.

	<i>encounter</i>	<i>uncommon- common</i>
a. <i>Ik weet dat Jan hard moet kunnen werken.</i>	yes/no	1-2-3-4-5
b. <i>Ik weet dat Jan hard moet werken kunnen.</i>	yes/no	1-2-3-4-5
c. <i>Ik weet dat Jan hard kunnen moet werken.</i>	yes/no	1-2-3-4-5
d. <i>Ik weet dat Jan hard kunnen werken moet.</i>	yes/no	1-2-3-4-5
e. <i>Ik weet dat Jan hard werken kunnen moet.</i>	yes/no	1-2-3-4-5
f. <i>Ik weet dat Jan hard werken moet kunnen.</i>	yes/no	1-2-3-4-5

Figure 6. The six possible word orders in the verbal cluster MOD-MOD-V_{Inf} (cf. (11a))

This judgment task leaves room for native-speaker introspection in reflecting a considerable degree of individual variability (cf. Cornips and Poletto 2005). This is in accordance with the more recent plea that linguists should strive towards a more systematic collection strategy for intuitions in experimental elicitation settings, particularly given the fact that when research proceeds data has become increasingly subtle and in many ways more challenging (Labov 1996; Schütze 1996; Gervain and Zemplén 2005). The elicited 'intuitions' in Figure 6 provide all possible options regarding word order in the three-verb cluster. Heterogeneity is assumed by providing all alternative variants and the task is designed to reveal competence in bidialectal, heterogenous varieties rather than in a single, homogeneous variety (Cornips and Corrigan 2005, Rickford 1987: 159).

Table 6 presents the quantitative results of the relative judgments task for the whole area, i.e., the Netherlands and Dutch-speaking Belgium.⁵

It shows the absolute numbers for the five-point 'yes' scale and 'no'-answers:

Table 6. Word orders in three-verb clusters in dialects of Dutch according to acceptability scale 1 through 5 ($n = 370$, 100%) (cf. Bolk 2004)

MOD-AUX-V _{Part} <i>moet hebben gemaakt</i> 'must have repaired'							
	yes (total)	y = 5	y = 4	y = 3	y = 2	y = 1	no, no answer
1-2-3	223 = 60.3%	107	47	37	20	12	147 = 39.7%
1-3-2	160 = 43.2%	78	32	33	9	8	210 = 56.8%
2-1-3	9 = 2.4%	4	2	0	1	2	361 = 97.6%
2-3-1	6 = 1.6%	2	1	0	1	2	364 = 98.4%
3-1-2	305 = 82.4%	210	49	30	9	7	65 = 17.6%
3-2-1	79 = 21.3%	44	14	15	5	1	291 = 78.7%

MOD-MOD-V _{Inf} <i>moet kunnen werken</i> 'must can work'							
	yes (total)	y = 5	y = 4	y = 3	y = 2	y = 1	no, no answer
1-2-3	315 = 85.1%	249	40	17	3	6	55 = 14.9%
1-3-2	34 = 9.2%	8	5	6	11	4	336 = 90.8%
2-1-3	2 = 0.5%	0	0	1	0	1	368 = 99.5%
2-3-1	12 = 3.2%	1	0	3	3	5	358 = 96.8%
3-1-2	75 = 20.3%	10	10	21	17	17	295 = 79.7%
3-2-1	23 = 6.2%	15	1	5	1	1	347 = 93.8%

When we consider the total of all 'yes'-answers in Table 6, it appears that the MOD-AUX-V_{Part} cluster allows high percentages for four and the MOD-MOD-V_{Inf} for two different orders (20% or more). These results confirm the findings of the Heerlen Dutch spontaneous speech corpus: A cluster containing an AUX that selects a participle as its complement allows for more variability than a MOD selecting an infinitive. Further, the word orders attested in Heerlen Dutch are the ones that have the highest percentages of acceptability throughout

the whole area, that is MOD-MOD-V_{Inf} 1-2-3 (85.1%) and MOD-AUX-V_{Part} 3-1-2 (82.4%). From a standard/Heerlen Dutch perspective, the word order in MOD-MOD-V_{Inf} 3-1-2 is remarkable, as is 3-2-1 in MOD-AUX-V_{Part}. The former appears quite common in the Netherlands, in particular, but not restricted to the eastern part and is very rare in Belgium, whereas the latter is mainly found in northern dialects (Barbiers 2005: 241 and 243).

Since this paper focuses on the issue of idiolectal variability, the first question to be raised is whether a speaker accepts several orders within a type of verbal cluster. Table 7 provides the results for subjects accepting several orders regardless the numerical value (yes = 1-5). Clearly, the MOD-AUX-V cluster allows more idiolectal variation than the MOD-MOD-V_{Inf} cluster. If we establish a threshold of 10%, the former allows for three orders per subject whereas the latter allows for two orders per subject. Moreover, in the former, the percentages accepting one, two and three orders are roughly similar whereas the percentage accepting just one order in the MOD-MOD-V_{Inf} cluster is clearly the highest one:

Table 7. The amount of orders accepted by subjects regardless the numerical value (taken from Bolk 2004)

Type cluster	0 orders	1 orders	2 orders	3 orders	4 orders	5 orders	6 orders	total number of orders per speaker/mean
MOD-AUX-V	17 = 4.6%	87 = 24.6%	139 = 39.4%	97 = 27.5%	26 = 7.4%	2 = 0.6%	2 = 0.6%	2.2 $n = 370$
MOD-MOD-V	39 = 10.5%	233 = 70.4%	71 = 21.5%	23 = 6.9%	3 = 0.9%	1 = 0.3%	0 = 0%	1.4 $n = 370$

A second question is whether a speaker considers the several orders within a cluster as equivalent. The group who allows different orders with the same numerical values (yes = 5, cf. Table 9) is smaller than the group of speakers who allows several orders within a cluster with different numerical values (Table 7). Table 8 reveals that most subjects choose a single 5 for one order in the two types of verb cluster. However, the cluster MOD-AUX-V_{Part} is the cluster that allows more than one order again:

Table 8. The amount of orders within a cluster type that has the highest acceptability value (5) (taken from Bolk 2004)

Type cluster	0 orders	1 order, yes=5	2 orders yes=5	3 orders yes=5	4 orders yes=5	total
MOD-AUX-V	59 = 16%	211 = 57%	73 = 20%	22 = 6%	4 = 1%	369
MOD-MOD-V	103 = 28%	253 = 69%	12 = 3%	-	-	368

The final question to be addressed is whether the speakers revealing idiolectal variability have the same combination of orders. The results in Table 9 show that the two most favorite orders to combine in the two types of clusters are 1-2-3 / 3-1-2 and the three most favorite ones in the MOD-AUX-V_{Part} are 1-2-3 / 1-3-2 / 3-1-2.⁶

Table 9. Combinations of accepted orders per type of cluster (more than four subjects, taken from Bolk 2004)

	Combinations of orders	MOD-AUX _{Perf} -V	MOD-MOD-V
two orders	a) 1-2-3 / 1-3-2	4	15
	b) 1-2-3 / 3-1-2	82	47
	c) 1-2-3 / 3-2-1	-	4
	d) 1-3-2 / 3-1-2	39	-
	e) 3-1-2 / 3-2-1	14	-
three orders	a) 1-2-3 / 1-3-2 3-1-2	71	9
	b) 1-2-3 / 2-3-1 3-1-2	-	4
	c) 1-2-3 / 3-1-2 3-2-1	21	5
	four orders	a) 1-2-3 / 1-3-2 3-1-2 / 3-2-1	22

Thus, the individual preferred combination of word orders in the SAND corpus always includes the standard Dutch word orders throughout the Netherlands and Belgium, that is, the 1-2-3 order for the MOD-MOD-V_{Inf} cluster and the order 1-2-3 / 3-1-2 ($n = 82$) and 1-2-3 / 1-3-2 / 3-1-2 ($n = 71$) in the MOD-AUX-V_{Part} cluster.

In sum, the MOD-MOD-V_{Inf} cluster allows for less variability than the MOD-AUX-V_{Part} cluster at the level of the individual speaker as well as the overall range of word orders. Second, individual variation, regardless of its geographical distribution, always includes the standard Dutch word orders. Hence, the 1-2-3 order is always a subset of the MOD-MOD-V_{Inf} clusters and the favorite combination of two orders in the MOD-AUX-V_{Part} cluster is 1-2-3 and 3-1-2, whereas the most favorite one of three orders is 1-2-3, 1-3-2 and 3-1-2. So, if a local dialect speaker displays variability, one of the variants always concerns the standard Dutch order; adult dialect speakers are no longer 'true' monolinguals anyway. Indeed, most competent speakers of a dialect variety can resort to standard Dutch.

5. Conclusion

The goal of this paper is to discuss idiolectal variability concerning word order in the two-verb and three-verb cluster (right periphery) on the basis of two solid empirical studies. The first study is based on 33.5 hours of recorded spontaneous speech of 67 speakers in the city of Heerlen in the southeastern part of the Netherlands. The second study involves written relative acceptability judgments collected for the *Syntactic Atlas of the Dutch Dialects* that has resulted in a large-scale geographical investigation of Dutch dialects in the Netherlands and Belgium. This corpus contains written judgments of 370 native speakers of local dialects throughout the whole area. These corpora allow possible intra- and inter-speaker variation to be distinguished.

The subjects in both corpora reveal idiolectal variability with respect to word orders per type of two- and three-verb cluster. The degree of individual variation and word order preferences depend on the type of auxiliary and morphology of the deepest embedded verb; that is, verbal clusters in which the auxiliary (perfective/passive) selects a past participle allow for more variability than those in which it (modal) selects an infinitive. With the patterns of usage in Heerlen, in particular, the fact that only one order is categorical and that there exists a dependency relationship between possible orders provide evidence that one of the orders has to be derived by movement. In this paper, it is argued that a basic head-initial structure with optional verb raising is capable of accounting for the empirical results. Furthermore, one of the variants included in the idiolectal variability always concerns the standard Dutch word order signaling a very advanced process of standardization.

Notes

- * I would like to thank Hans Broekhuis and one anonymous reviewer for their useful comments and the following trainees of the Meertens Institute who have contributed to this paper in one or another way: Elske Bolk, Hester van den Bosch, Josine Pennings and Anne Ribbert. Remaining errors are mine.
1. Only one subject had a conversation with the interviewer.
 2. The figure appearing before the speaker's pseudonym refers to the number of the tape-recording.
 3. Sentences in which the form *be* and the participle can be interpreted as a copula and adjective, respectively, are excluded as much as possible (see also 2.4).
 4. As Wurmbrand notes (2006: 233), the choice between different approaches can only be made in conjunction with the choice of a particular syntactic framework or aspect of generative theory, and hence the decision between the head-final or head-initial structure is largely a subjective one.

5. See Barbiers (2005) for a more extensive overview of theoretical framework and the linguistic and geographical distribution of all administered types of the three-verb clusters in the SAND project.
6. Note that the 3–2–1 order excludes only raising of the most embedded verb (participle or infinitive). See Barbiers (2005) for a possible analysis for all attested word orders in all types of verbal clusters.

References

- Barbiers, Sjef
2005 Theoretical restrictions on word order variation in three-verb clusters. In Cornips and Corrigan (eds.) 2005, 233–264.
- Barbiers, Sjef, Leonie Cornips, and Jan-Pieter Kunst
2007 The Syntactic Atlas of the Dutch Dialects (SAND): A corpus of elicited speech as an on-line Dynamic Atlas. In *Models and Methods in the Handling of Unconventional Digital Corpora*. Volume 1: *Synchronic Corpora*, Joan C. Beal, Karen P. Corrigan and Hermann Moisl (eds.), 54–90. Hampshire: Palgrave-Macmillan.
- Barbiers, Sjef, Leonie Cornips, and Suzanne van der Kleij (eds.)
2002 *Syntactic Microvariation*. Electronic publication. Amsterdam: Meertens Instituut. <http://www.meertens.knaw.nl/projecten/sand/synmic/>
- Besten, Hans den, and Hans Broekhuis
1989 Woordvolgorde in de werkwoordelijke eindreeks. *Glott* 12: 79–137.
- Bolk, Elske
2004 Woordvolgorde in de werkwoordelijke eindreeks. Een beschrijving van de voorkomende variatie in het Nederlandstalig taalgebied. Master thesis University of Amsterdam/Meertens Institute.
- Broekhuis, Hans
1999 Adjectives and adjective phrases. *Modern Grammar of Dutch*. Occasional papers 2. Tilburg University.
- Cedergren, Henriette, and David Sankoff
1974 Variable rules: performance as a statistical reflection of competence. *Language* 50(2): 333–355.
- Cornips, Leonie
1994 Syntactische variatie in het Algemeen Nederlands van Heerlen. Unpublished dissertation, University of Amsterdam, IFOTT 6.
1998 Syntactic variation, parameters and their social distribution. *Language Variation and Change* 10(1): 1–21.
2005 Variation and formal theories of syntax, Chomskyan. In *Encyclopedia of Language and Linguistics*, Keith Brown (ed.), 330–332. Oxford: Elsevier.

- 2006 Intermediate syntactic variants in a dialect-standard speech repertoire and relative acceptability. In *Gradience in Grammar. Generative Perspectives*, Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky, and Ralf Vogel (eds.), 85–105. Oxford: Oxford University Press.
- Cornips, Leonie, and Karen P. Corrigan
2005 Convergence and divergence in grammar. In *Dialect Change: Convergence and Divergence in European Languages*, Peter Auer, Frans Hinskens, and Paul Kerswill (eds.), 96–134. Cambridge: Cambridge University Press.
- Cornips, Leonie, and Karen P. Corrigan (eds.)
2005 *Syntax and Variation. Reconciling the Biological with the Social*. (Current Issues in Linguistic Theory 265.) Amsterdam: John Benjamins.
- Cornips, Leonie, and Willy Jongenburger
2001 Elicitation techniques in a Dutch syntactic dialect atlas project. In *Linguistics in The Netherlands 18, 2001*, Hans Broekhuis and Ton van der Wouden (eds.), 57–69. Amsterdam: John Benjamins.
- Cornips, Leonie, and Cecilia Poletto
2005 On standardising syntactic elicitation techniques. Part I. *Lingua* 115(7): 939–957.
- Cornips, Leonie, and Anne Ribbert
2006 Variatie in tweeledige verbale clusters in het Heerlens Algemeen Nederlands (HAN). *Nederlandse Taalkunde* 11(1): 59–76.
- Gervain, Judit, and Gábor Zemplén
2005 Focus raising. A paradigmatic example of the treatment of syntactic microvariation. In Cornips and Corrigan (eds.) 2005, 123–145.
- Guy, Gregory
2004 Dialect unity, dialect contrast: the role of variable constraints. Talk presented at the Meertens Institute, Amsterdam, 9 August 2004.
- Henry, Alison
1995 *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.
2002 Variation and syntactic theory. In *The Handbook of Language Variation and Change*, Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), 267–282. Malden, Mass.: Blackwell.
- Kayne, Richard
1994 *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Koopman, Hilda
1995 De plaats van geïncorporeerde hoofden in de werkwoordscluster. *Tabu* 25: 174–179.
- Kortmann, Bernd (ed.)
2004 *Dialectology Meets Typology. Dialect Grammar From a Cross-Linguistic Perspective*. (Trends in Linguistics. Studies and Monographs 153.) Berlin: Mouton de Gruyter.

- Kroch, Anthony
1989 Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3): 199–244.
- Labov, William
1972 *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
1994 *Principles of Linguistic Change. Internal Factors*. Oxford: Blackwell.
1996 When intuitions fail. *Papers from the 32nd Regional Meeting of the Chicago Linguistics Society*: 76–106.
- Meechan, Marjory, and Michele Foley
1994 On resolving disagreement: linguistic theory and variation – “There’s bridges”. *Language Variation and Change* 6(1): 83–85.
- Pintzuk, Susan
1995 Variation and change in Old English clause structure. *Language Variation and Change* 7(2): 229–260.
- Rickford, John
1987 The haves and have nots: Sociolinguistic surveys and the assessment of speaker competence. *Language in Society* 16(2): 149–177.
- Schütze, Carson T.
1996 *The Empirical Base of Linguistics. Grammaticality Judgments and Linguistic Methodology*. Chicago: The University of Chicago Press.
- Seiler, Guido
2004 On three types of dialect variation. In Kortmann (ed.) (2004), 367–399.
- Stroop, Jan
1970 Systemen in gesproken Werkwoordsgroepen. *Taal en Tongval* 22: 128–147.
- Wilson, John, and Alison Henry
1998 Parameter setting within a socially realistic linguistics. *Language in Society* 27(1): 1–21.
- Wurmbrand, Susanne
2006 Verb clusters, verb raising, and restructuring. In *The Blackwell Companion to Syntax*, vol. V, Martin Everaert and Henk van Riemsdijk (eds.), 229–343. Oxford: Blackwell.
- Zwart, C. Jan-Wouter
1993 *Dutch Syntax. A Minimalist Approach*. Groningen Dissertations in Linguistics.
1996 Verb clusters in continental West Germanic dialects. In *Microparametric Syntax and Dialect Variation*, James R. Black and Virginia Motapanyane (eds.), 229–258. Amsterdam: John Benjamins.

Towards a multivariate model of grammar: The case of word order variation in Dutch clause final verb clusters*

Gert De Sutter

Abstract. In the present paper, we argue that grammatical phenomena in general, and word order phenomena in particular, are essentially multivariate in nature, and that a systematic-empirical, quantitative, multivariate treatment is needed in order to describe and model grammatical variation adequately. This does not just involve getting the right data, or starting from articulate and appropriate linguistic models, but it crucially involves linking data to models (i.e., operationalizing linguistic hypotheses in ways that make them empirically testable) and statistically analyzing the data by means of sophisticated multivariate techniques.

These theoretical and methodological claims are demonstrated on the basis of a quantitative corpus study which tackles one of the most intriguing cases of grammatical variation in Dutch, viz. word order variability in clause final verb clusters consisting of a past participle and an auxiliary verb.

1. Introduction

Recent years have witnessed an increased interest in grammatical variation phenomena. Research projects investigating phenomena such as the dative alternation, the relative order of prepositional and/or nominal constituents, heavy noun phrase shift and other extraposition phenomena, to name just a few, have yielded a better insight into the motivating principles of grammatical optionality, such as avoiding ambiguity, marking discourse function and processing facilitation (cf., e.g., Hawkins 1994; Arnold *et al.* 2000; Wasow 2002; Gries 2003). Nevertheless, some types of grammatical variation remain a hard nut to crack. One of these phenomena is the word order variability in Dutch clause final verb clusters consisting of a finite auxiliary verb and a participle. In such constructions, the auxiliary verb can either precede or follow the participle:

- (1) a. *dat ze een goed werkstuk heeft afgeleverd* (DS)¹
that she a good paper has submitted