Variation in verb cluster interruption

Lotte Hendriks
Meertens Institute

Except for finite verbs in main clauses, verbs in Standard Dutch cluster together in a clause-final position. In certain Dutch dialects, non-verbal material can occur within this verb cluster (Verhasselt 1961; Koelmans 1965, among many others). These dialects vary with respect to which types of elements can interrupt the verb cluster, varying from particles to various types of arguments and adverbs (Barbiers, van der Auwera, Bennis, Boef, de Vogelaer & van der Ham 2008). A study amongst forty Dutch dialect speakers reveals an ordered ranking of grammatical types, reflecting their acceptability in a verb cluster. I argue that this ranking directly follows from syntactic principles: The syntactic size and position of the intervening element affect its acceptance in a verb cluster. Potentially, these principles interact with a preference of performance dubbed 'minimize domains' (Hawkins 1994, 2003, 2004), which requires both the higher verb and the intervening element to be adjacent to the main verb, leading to two conflicting structures.

Keywords: Cross-linguistic variation, verb cluster interruption, implicational hierarchy, processing preference

1. Introduction

Except for finite verbs in main clauses, verbs in Standard Dutch cluster together in a sentence-final position.

(1) *Ik vind dat Jan een schuur moet bouwen.*
  I find that Jan a barn must build
  ‘I think that Jan should build a barn.’

In certain Dutch dialects, non-verbal material can occur within a verb cluster.1

(2) *Ik vind dat Jan moet een schuur bouwen.*
  I find that Jan must a barn build
  ‘I think that Jan should build a barn.’
Various studies have revealed several types of elements that can occur in verb clusters. Strikingly, there is much cross-dialectal variation in the use of the construction that has hitherto not been accounted for. Not only is the construction restricted to a specific part of the language area — the construction is used most frequently in dialects spoken in West and East Flanders (Vanacker 1964, among others) — even within those dialects there is variation in the types of elements that can interrupt a verb cluster, as demonstrated by Barbiers et al. (2008, map 30b). Definite objects, for instance, are much less common in a verb cluster than indefinite objects, which in their turn are less common than bare objects.

This paper reports on a novel study of verb cluster interruptions, which aimed to answer the following research question.

(3) Which factors can explain the observed variation in the use of verb cluster interruptions?

To answer this question, it was investigated which types of constituents can interrupt the cluster and which variables play a role in the acceptance of this construction. It will be demonstrated that syntactic principles can account for most of the observed variation in verb cluster interruptions.2

2. Methodology

2.1 Informants

Forty Flemish dialect speakers were recruited for the questionnaire. These informants came from places across West and East Flanders, Flemish Brabant and Flemish Limburg, where verb cluster interruptions were observed in previous research (Barbiers et al. 2008).

A number of 28 informants were recruited who had participated in the study by Barbiers et al. (2008). Other dialect speakers were inter alia recruited by contacting town halls and retirement homes. As it was difficult to find enough informants in a limited amount of time, the criteria were relaxed for these additional informants.

2.2 Design

Each oral questionnaire began with a practice question. For every test item, the interviewer read out a sentence with neutral intonation, after which the informant was asked to judge if this sentence could be used in their dialects. To check
the judgments, the informants were also asked to translate the sentence into their dialect. The speakers were not informed about the purpose of the study.

2.3 Material

On the basis of the findings in previous literature (Wurmbrand 2006; Barbiers et al. 2008), a number of types of non-verbal elements that could potentially show variation were tested. This led to a total number of 293 test items, which were offered in a randomized order in five different versions. To keep other variables constant, the test sentences were all construed according to the format in (4).

(4) Ik vind dat Jan (DP_{obj}) moet X verb
    I find that Jan (DP_{obj}) must X verb

2.4 Statistical analysis

A Pearson’s chi-square test was performed to discover whether the acceptability of a verb cluster interruption depends on the type of element that interrupts the cluster. An alpha level of .05 was used to determine the significance of the observed differences. In cases where multiple groups were compared to each other, a Bonferroni correction was applied. Additionally, Pearson’s correlational test was performed to measure the effect of the type of element on its acceptability in a verb cluster.

3. Results

3.1 Main results

The results of the investigation indicate that different types of elements can interrupt a verb cluster. The main results are depicted in Table 1.

When comparing the different groups, a significant association between the group and the acceptability of the verb cluster interruption was observed. In fact, the differences between all groups are significant. On first examination, the difference in the acceptability of the interrupting items seems to lie in the length and the syntactic position of the interrupting element; the elements in group 3 are larger than the elements in group 1 and 2 and the elements in group 4 are elements that might either be base-generated higher in the clausal structure, or prefer to be scrambled to a higher position. In the next section we will further investigate these differences.
Table 1. The main results

<table>
<thead>
<tr>
<th>Group 1: Particles</th>
<th>Element</th>
<th>Accepted</th>
<th>Percentage accepted</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adpositional particle</td>
<td>20/21</td>
<td>95.2</td>
<td>93.4</td>
</tr>
<tr>
<td></td>
<td>Adverbial particle</td>
<td>37/40</td>
<td>92.5</td>
<td></td>
</tr>
<tr>
<td>Group 2: Single elements</td>
<td>Predicate adjective</td>
<td>22/40</td>
<td>55.0</td>
<td>45.5</td>
</tr>
<tr>
<td></td>
<td>Bare singular object</td>
<td>19/40</td>
<td>47.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bare plural object</td>
<td>16/35</td>
<td>45.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Depictive (oriented at object)</td>
<td>18/40</td>
<td>45.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resultative</td>
<td>12/31</td>
<td>38.7</td>
<td></td>
</tr>
<tr>
<td>Group 3: Other low elements</td>
<td>VP-internal subject</td>
<td>9/37</td>
<td>24.3</td>
<td>21.8</td>
</tr>
<tr>
<td></td>
<td>Indefinite object</td>
<td>8/37</td>
<td>21.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prepositional phrase</td>
<td>7/36</td>
<td>19.4</td>
<td></td>
</tr>
<tr>
<td>Group 4: Other</td>
<td>Temporal adverb</td>
<td>3/40</td>
<td>7.5</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>Definite object</td>
<td>2/40</td>
<td>5.0</td>
<td></td>
</tr>
</tbody>
</table>

3.2 The size of the interrupting element

The acceptability of verb cluster interruption might be affected by the number of syllables, morphemes, words, or the amount of syntactic structure of the interrupting element.

The influence of the number of syllables of the interrupting constituent was tested with the sentences in (5). The results show that a larger number of syllables did not have an effect on the acceptability of the construction.

(5) a. *Ik vind dat Jan moet een pen kopen.* (Accepted: 7/37 = 18.9%)
    I find that Jan must a pen buy
    ‘I think that Jan should buy a pen.’

b. *Ik vind dat Jan moet een thermometer kopen.* (Accepted: 7/37 = 18.9%)
    I find that Jan must a thermometer buy
    ‘I think that Jan should buy a thermometer.’

The sentences in (6) indicate that the number of morphemes also does not influence the acceptability of an element in a verb cluster.

(6) a. *Ik vind dat Jan moet een pen kopen.* (Accepted: 7/37 = 18.9%)
    I find that Jan must a pen buy
    ‘I think that Jan should buy a pen.’

b. *Ik vind dat Jan moet een pen-netje kopen.* (Accepted: 7/37 = 18.9%)
    I find that Jan must a pen-DIM buy
    ‘I think that Jan should buy a little pen.’
Furthermore, as exemplified in (7), when comparing sentences in which the number of words of interrupting elements differs, the acceptability again remains similar; the small differences are not significant.

(7) a. *Ik vind dat Jan moet Henriette zoenen.* (Accepted: 4/36 = 11.1%)
'I find that Jan must Henriette kiss

b. *Ik vind dat Jan moet de meisjes zoenen.* (Accepted: 3/36 = 8.3%)
'I find that Jan must the girls kiss

c. *Ik vind dat Jan moet een mooi wit paard kopen.*
(Accepted: 4/37 = 10.8%)
'I find that Jan must a beautiful white horse buy

On the other hand, the following sentences demonstrate that the acceptability of a verb cluster interruption can change, even when the amount of words remains the same.

(8) a. *Ik vind dat Jan moet brood eten.* (Accepted: 19/40 = 47.5%)
'I find that Jan must bread eat

b. *Ik vind dat Jan moet Els zoenen.* (Accepted: 4/36 = 11.1%)
'I find that Jan must Els kiss

These sentences demonstrate that proper names are less acceptable in a verb cluster than bare nouns. This could be the result of the syntactic structure of the different elements, or of a definiteness effect. Note that elements that prefer to be scrambled are generally bad in a verb cluster:

(9) *Ik vind dat Jan moet de auto verkopen.* (Accepted: 2/40 = 5%)
'I find that Jan must the car sell

Further research into the interaction between scrambling and verb cluster interruption, and hence the difference between definite and indefinite objects, is required (cf. den Besten & Broekhuis 1992).

In addition to the definiteness effect, the amount of syntactic structure seems to be of influence on the acceptability of a verb cluster interruption. This becomes especially clear when comparing the various types of direct objects: Bare singular nouns are more acceptable in a verb cluster than indefinite objects ($\chi^2(1) = 5.7$, $p < .025$, $r = -.3$). Since we cannot attribute this difference to the amount of
syllables, morphemes or words, only the syntactic structure of these constituents can be argued to affect their acceptability in a verb cluster.

We can now conclude that the syntactic structure of interrupting elements affects their acceptability in a verb cluster. The effects of syntactic structure also become evident when considering particles. As became clear in Section 3.1, these elements, which have been argued to be non-projecting words (see Blom 2005 and references cited therein), are more acceptable than any other element and even obligatorily interrupt in some languages (see Wurmbrand 2006).

3.3 The position of the interrupting elements

In the previous section it became clear that adverbs are less acceptable in a verb cluster. Potentially, this difference is due to the position of adverbs in the clausal spine. As exemplified in (10), some adverbs cannot be fronted with the verb, suggesting that these are external to the VP, whereas all elements in group 2 are VP-internal and can be fronted with the verb.

(10) a. *[Waarschijnlijk winnen] zal Jan niet.
   probably win will Jan not
   ‘Jan will probably not win.’

b. [Brood eten] wil Jan niet.
   Bread eat want Jan not
   ‘Jan doesn’t want to eat bread.’

VP-internal material hence seems to be more acceptable in a verb cluster than VP-external material. The question that arises concerns the acceptability of adverbs that occupy different syntactic positions. According to Cinque (1999, 2006), there is a fixed universal base hierarchy of adverb types. On the basis of this hierarchy, I tested six different types of adverbs in six different contexts. The results, following the order in Cinque’s hierarchy, are depicted in (11) and demonstrate that the acceptability level slightly improves when lower adverbs are used; the difference between the highest adverb and the lowest adverb is significant ($\chi^2(1) = 11.0, p < .01, r = -.2$).\(^9\)

(11) \[\text{MoodP evaluative} \quad helaas ‘unfortunately’ \quad (Accepted: 0.5%)\]
   \[\text{ModP epistemic} \quad waarschijnlijk ‘probably’ \quad (Accepted: 0.9%)\]
   \[\text{TP} \quad toen ‘then’ \quad (Accepted: 1.3%)\]
   \[\text{Mod Palethic} \quad noodzakelijk/mogelijk ‘necessarily/possibly’ \quad (Accepted: 3.2%)\]
   \[\text{AspP continuative} \quad voortdurend ‘continually’ \quad (Accepted: 3.2%)\]
   \[\text{VoiceP} \quad goed ‘well’ \quad (Accepted: 6.5%)\]
4. The source of the variation

4.1 The role of syntax

As the syntactic structure and base position of the interrupting element affects the acceptability of a verb cluster interruption, the cross-dialectal variation in verb cluster interruption can be hypothesized to follow from syntactic principles. A syntactic theory is logically dependent on the assumed underlying word order in the language. If one assumes a head-initial or a head-final word order in Dutch, verb cluster interruptions would require a movement process, as illustrated in the simplified Figures 1 and 2.10

![Figure 1. SOV (Den Besten & Edmondson 1983 a.o.)](image1)

![Figure 2. SVO (Zwart 1994, 1996 a.o.)](image2)

On the other hand, if one assumes that complements with a verbal core are generated after the head, whereas nominal complements are generated before the head, following Barbiers (2000, 2008), a verb cluster interruption is the base-generated order. In that case, verb cluster interruption comes for free by simply stating that these verb clusters lack movement, or involve a movement after spell-out; this is illustrated in Figure 3.

![Figure 3. S-AUX-O-V (Barbiers 2000, 2008)](image3)

In all these theories, independent considerations are required to account for the observed variation between various types of elements. A syntactic approach would, for instance, not predict a distinction in verb cluster interruptions by different elements of the same syntactic category that occupy the same structural position.
However, the results clearly indicated that speakers of Flemish dialects find bare singular nouns more acceptable in a verb cluster than indefinite objects. In the head-initial and head-final approaches, one would need to assume an elementspecific movement process to account for this variation. Conversely, following Barbiers’ analysis, one would need to account for the fact that some elements do not move to a higher position, at least before spell-out.

4.2 The role of preferences of performance

The cross-linguistic differences in elements that can interrupt a verb cluster could also have a functional, rather than a syntactic, ground. A preference of performance that has been dubbed ‘minimize domains’ (Hawkins 2003, 2004) could potentially explain variation in verb cluster interruptions. This principle entails that the distance between items that are semantically or syntactically related to each other immediately affects the efficiency and complexity of processing.

Hawkins (1994, 2003, 2004) argues that there is a clear link between grammar and performance. According to him, syntactic structures that are preferred in performance will be conventionalized in human languages. He suggests that, as the distance between interrelated items affects the efficiency and complexity of a sentence, surface structures in which interrelated elements are adjacent to each other will often be conventionalized in languages.

Sentences (12a and 12b) make clear that the verb cluster interruption environments involve contradictory requirements, as there are at least two pairs of interrelated items: The auxiliary selects the main verb and the object is assigned a thematic role by the main verb.

(12) a. …dat Jan moet de auto maken.
   …that Jan must the car make
b. …dat Jan de auto moet maken.
   …that Jan the car must make
   ‘…that Jan should fix the car.’

Linguistic variation could accordingly be due to the option to select between multiple available structures. This idea is in line with Hawkins’ hypothesis that grammatical variation arises when processing preferences are in competition with each other (Hawkins 2004). Crucially, processing preferences could account for the observed optionality in the use of this construction: All speakers allowed both non-interrupted and interrupted verb clusters. This cannot easily be accounted for by variation in the grammar, such as parameter settings.

The optimal solution for processing is one in which the main verb is positioned between the finite verb and the interrupting element: X-V2-V1 or V1-V2-X as is
found in English. However, in the Dutch dialects, both these orders are excluded for syntactic reasons and are indeed rarely allowed; sentences in which the main verb is positioned between the bare noun and the modal verb are significantly less acceptable than sentences in which a bare noun interrupts the verb cluster ($\chi^2(1) = 18.7 \ p < 0.001$, $r = -0.5$). This again demonstrates that syntax plays an active role in the acceptability of verb cluster interruptions. Adding to the effects of syntax discussed in Section 3, we can conclude that variation in verb cluster interruption is not caused merely by processing preferences: There is a clear influence of the syntactic position and structure of the interrupting element.

4.3 Interaction effects

The results cannot straightforwardly be accounted for by a purely syntactic theory, nor by processing preferences. In fact, the results seem to follow directly from syntactic principles, interacting with preferences of performance. The processing requirement of elements to be close to the main verb is dependent on their syntactic structure as well as their structural position. These results are in line with Hawkins (2004), who states that complexity increases with the number of linguistic forms and the number of conventionally associated (syntactic and semantic) properties that are assigned to them.

We can now account for differences between elements of the same syntactic category that occupy the same syntactic position, such as the sentences in (13).\textsuperscript{14}

\begin{enumerate}
\item\hspace{1em} Ik vind dat Jan moet brood eten. (Accepted: 19/40 = 47.5%)
I find that Jan must bread eat
\textquoteleft I think that Jan should eat bread.	extquoteright
\item\hspace{1em} Ik vind dat Jan moet brood met pinda-kaas eten.
(Accepted: 6/40 = 15.0%)
I find that Jan must bread with peanut-butter eat
\textquoteleft I think that Jan should eat bread with peanut butter.	extquoteright
\end{enumerate}

The direct objects in both these sentences are base-generated in the same position below the modal of obligation. Furthermore, the entire constituents are of the same syntactic category. The difference between these sentences can only be accounted for by the syntactic structure of the intervening element, which potentially affects the minimal domains between the verbs.\textsuperscript{15}

We can now account for the scale in Table 1: Group 1 consists of particles, which are the least complex elements, see also Section 3.2; group 2 consists of syntactically simplex elements that are generated low in the syntactic structure; group 3 consists of syntactically more complex elements. The differences between these three groups are therefore expected. Following Miller & Chomsky (1963),
Hawkins (2004) argues that terminal elements that have a higher amount of syntactic structure associated with them require more linguistic properties to be processed and are hence more complex. The internal structure of intervening elements is hence predicted to affect their ability to occur in a verb cluster.

The general unacceptability of the elements in group 4 can be attributed to their position in the syntactic structure. On the one hand, group 4 consists of adverbs that are generated higher in the syntactic structure. Elements that are base-generated syntactically close to the main verb have a stronger selectional relationship with the verb and are therefore predicted to have a stronger preference to be adjacent to the main verb in linear order. Scrambled elements, on the other hand, are base-generated in a lower position. The preference for these elements to scramble hence seems to be stronger than the preference to be close to the main verb. Further research is required to investigate the precise relation between scrambling and verb cluster interruptions.

5. Conclusion

The frequencies of the accepted verb cluster interruptions by the different types of elements have revealed an ordered ranking of grammatical types, reflecting their acceptability in a verb cluster. Variation in verb cluster interruptions is hence structured and is demonstrated to follow from syntactic principles:

1. Elements that have a higher amount of syntactic structure associated with them are less acceptable in a verb cluster.
2. Elements that are base-generated syntactically close to the main verb have a stronger preference to be adjacent to the main verb in linear order.

On the basis of these facts we can formulate a two-folded hierarchy of verb cluster interruptions.

(14) a. Syntactically simplex elements > syntactically complex elements
    b. Syntactically low elements > syntactically high elements

Crucially, the hierarchy is not conspicuous in the results of each dialect speaker. For instance, some speakers allow verb cluster interruptions by some, but not all elements from one group. This supports the suggestion that variation is influenced by preferences of performance, rather than pure syntactic factors.
Notes

* I am very grateful to Sjef Barbiers for his helpful comments, feedback and suggestions, as well as Marjo van Koppen for her comments on an earlier version of this work.

Parts of this paper were presented at my thesis defense (Utrecht, 16 October 2013), the Uiltjesdagen (Utrecht, 25 November 2013), a t-lezing (Meertens Instituut, 2 December 2013), at the LOT winter school master class (Enoch Aboh & Marjo van Koppen, 20–24 January 2014) and at the TIN-dag (1 February 2014). I thank the people in the audiences for their helpful comments.

1. In this paper, the term verb cluster is used in a pre-theoretical sense, indicating a sequence of verbs that do not necessarily form a constituent together.

2. For reasons of space I cannot discuss all factors that determine cluster interruptions as observed in previous literature. I refer the reader to Wurmbrand (2006) for a more extensive discussion.

3. For the purpose of this study if have chosen the simplest possible acceptability judgement task. For follow-up research it would be interesting to see if the use of more fine-grained methods such as a 5- or 7-points Likert scale or magnitude estimation will yield different results.

4. For reasons of space, not all results can be discussed in this paper.

5. One factor that could have negatively affected the outcome of the investigation is known as accommodation; speakers of a language tend to adjust their language to that of their speech partner. For reasons of time, the interviewer who conducted the research was not a speaker of the dialect under investigation, but a speaker of Standard Dutch. Since many instances of verb cluster interruption are not allowed in Standard Dutch, conceivably more instances of verb cluster interruptions might have been observed if a different research method had been chosen. Fortunately, the results show a high degree of reliability, as they concur with observations made in the previous literature.

6. These numbers are not always based on the total number of informants. Certain constructions could not be used by all informants and were therefore excluded from the count. Furthermore, some items have not been investigated in all dialects, as three dialect speakers did not participate in the second part of the investigation.

7. Difference group 1 & 2: $\chi^2(1) = 43.8, p < .001, r = -.4$; difference group 2 & 3: $\chi^2(1) = 17.5, p < .001, r = -.2$; difference group 3 & 4: $\chi^2(1) = 8.7, p < .01, r = -.2$. Note that the Bonferroni correction requires an alpha level of .016.

8. These microvariational findings are strikingly parallel to those of a cross-linguistic study on Afrikaans, Dutch, German, Swiss and West-Flemish reported in Wurmbrand (2006: 275).

9. Note that celerative, repetitive, frequentative and completive adverbs, which are positioned even lower in Cinque’s hierarchy, were not included in this research, because these adverbs can all be generated in two distinct positions, one of which higher in the structure, as exemplified in (i).

(i) a. John quickly (slowly) lifted his arm.
   b. John lifted his arm quickly (*slowly).
10. I refer the reader to Wurmbrand (2006) for a more extensive discussion of the many different analyses of verb cluster interruptions.

11. More specifically, in a head-final approach, one would for instance have to argue that particular noun phrases can move out of the embedded verb phrase before verb projection raising takes place, cf. Coppen & Klein (1992), in a head-initial approach, one would have to assume various landing sites depending on the type of object.

12. A language-specific incorporation process could potentially account for part of the variation, as this usually affects terminal elements. However, verb cluster interruptions behave differently from incorporated nouns in a number of ways. Firstly, incorporation is only possible from complement position. However, adverbs and VP-internal subjects have been observed in a verb cluster. Secondly, truly incorporated elements occupy a different position.

   (i) a. …dat Jan zijn pap niet heeft proberen (op) te (*op) eten.
       …that Jan his porridge not has try (up) to (*up) eat.
   b. …dat Jan niet heeft proberen (piano) te (piano) spelen.
       …that Jan not has try (piano) to (piano) play

Finally, noun incorporations have been observed in dialects that generally disallow verb cluster interruptions (see Barbiers et al. 2008).

Nevertheless, the observed variation could partly be due to a difference between head and XP movement. For reasons of space I cannot discuss this possibility further.

13. This is inter alia supported by the morphological shape of the main verb, which is dependent on the auxiliary.

14. $\chi^2(1) = 9.8, p < 0.01, r = -0.4$

15. Another possibility is that syntactic complexity causes the variation: the amount of syntactic features could potentially lead to a different requirement of elements to move. This is refuted by the fact that there is an effect of the amount of syntactic structure, regardless of the types of syntactic elements.

References


**Author’s address**

Lotte Hendriks
Meertens Instituut
Postbus 94264
1090 GG Amsterdam, The Netherlands

Lotte.Hendriks@meertens.knaw.nl