

# VERB CLUSTERS, VERB RAISING, AND RESTRUCTURING

Susi Wurmbrand  
*University of Connecticut*

## Abstract:

This overview article reviews the major generalizations, trends, and theoretical findings of the research on verb clusters. The first part of the overview provides a summary of the diverse empirical distribution of verb clusters across West Germanic, including the results of several recent extensive dialectal studies. The variation within certain dialect groups is hypothesized to be the result of extensive dialectal bilingualism and a descriptive rule system is given which reflects certain dialectal subset relations. The second part of the overview surveys the merits and limits of a range of syntactic approaches to the verb cluster phenomenon and concludes with some speculations about one of the core open issues—the question of why verbs cluster and what the deep motivation for the verb clustering phenomenon is. Among others, the following empirical issues and theoretical questions are discussed in this overview: i) whether there is a direct or indirect causal relation between word order and morphology (in particular the *Infinitivus Pro Participio* ‘Infinitive for participle’ and *Participium Pro Infinitivo* ‘participle for infinitive’ effects); ii) whether in verb clusters with three verbal elements the 2-1-3 order exists as a genuine verb cluster order; iii) whether verb cluster orders are derived by syntactic movement or other linearization mechanisms; iv) whether and how verb clusters contribute to debates about the existence of directionality in syntax, the nature and motivation of movement, or the featural make-up of the verbal domain; and v) whether syntax can involve optionality.

## 1. INTRODUCTION

The term *verb cluster* refers to constructions involving more than one verbal element such as the examples in (1). The notation in this article will follow the common practice used in works on verb clusters which is to label the different verbal elements with numbers representing the hierarchical (i.e., deep structure or selectional) order of the elements. In particular, ascending numbering will be used such that the structurally highest verb (i.e., the verb that scopes over all other verbs) is assigned 1, the next highest 2 etc.<sup>1</sup> As the examples in (1a-c) show, in a language like English, the linear order of verbs in a verb cluster strictly follows the hierarchical order in that a verbal complement follows the selecting verb. To express for instance a sense that John is required to have the desire to leave, a sentence like (1a) would be used, whereas the sense that

---

<sup>1</sup> In addition to the numbering used for verbs to indicate their hierarchical relations, I also use V1, V2... to refer to the verbal elements in a verb cluster irrespective of whether the elements are main verbs, modals, or auxiliaries. Furthermore, the following abbreviations are used: AF(rikaans), AUX(iliary), DU(tch), FIN(ite), INF(itive), HMC (Head Movement Constraint), IPP (*Infinitivus Pro Participio* ‘Infinitive for participle’), LCA (*Linear Correspondence Axiom*) MOD(al), PART(icipie), PTCL (participle), PPI (*Participium Pro Infinitivo* ‘participle for infinitive’), SAND (*Syntactische Atlas van de Nederlandse Dialecten/Syntactic Atlas of the Dutch Dialects*), SG (Swiss German), WF (West Flemish).

John has a desire to be required to leave would be expressed by a sentence like (1b). In a language like German, on the other hand, the linear order of verbs in a verb cluster corresponds to the opposite hierarchical order: as can be seen in (1d), the verbs appear in a descending ...3-2-1 order—i.e., hierarchically lower verbs precede hierarchically higher verbs.

(1) English vs. German

- |    |      |          |         |         |        |       |      |                 |
|----|------|----------|---------|---------|--------|-------|------|-----------------|
| a. | John | has to   | (1)     | want to | (2)    | leave | (3)  |                 |
| b. | John | wants to | (1)     | have to | (2)    | leave | (3)  |                 |
| c. | that | John     | must    | (1)     | have   | (2)   | been | (3) elected (4) |
| d. | daß  | Hans     | gewählt |         | worden | sein  | muß  |                 |
|    | that | John     | elected | (4)     | become | (3)   | be   | (2) must (1)    |

The contrast between (1c) and (1d) might not appear particularly surprising, given that the order between heads and complements is often inverted in the two languages. However, looking in more detail at the distribution of verb clusters in German and other West Germanic languages and dialects, it has been observed that the situation is in fact far more complex in that not only ascending and descending word orders are possible, but also (in at least certain languages and constructions) various other combinations. To give an illustration of (some aspect of) the verb cluster phenomenon, consider Swiss German. As is shown in (2) (cf. Schönenberger 1995:382), constructions involving two modal verbs can (at least for certain Swiss speakers) be expressed with a fully ascending order of the verbs (cf., the ‘English’ order in (2a)), with a fully descending order as in (2b), or with mixed orders such as the 1-3-2 order in (2c) and the 3-1-2 order in (2d). Importantly, all versions of (2) have the same meaning—i.e., *want* takes scope over *can* and not vice versa.

- |     |    |   |    |     |          |     |          |     |          |     |              |
|-----|----|---|----|-----|----------|-----|----------|-----|----------|-----|--------------|
| (2) | a. | das                                     | er | ... | wil      |     | chöne    |     | vorsinge |     | Swiss German |
|     |    | that                                    | he | ... | wants    | (1) | can      | (2) | sing     | (3) |              |
|     | b. | das                                     | er | ... | vorsinge |     | chöne    |     | wil      |     |              |
|     |    | that                                    | he | ... | sing     | (3) | can      | (2) | wants    | (1) |              |
|     | c. | das                                     | er | ... | wil      |     | vorsinge |     | chöne    |     |              |
|     |    | that                                    | he | ... | wants    | (1) | sing     | (3) | can      | (2) |              |
|     | d. | das                                     | er | ... | vorsinge |     | wil      |     | chöne    |     |              |
|     |    | that                                    | he | ... | sing     | (3) | wants    | (1) | can      | (2) |              |
|     |    | all: ‘that he wants to be able to sing’ |    |     |          |     |          |     |          |     |              |

Besides differences among languages or dialects in the order of the verbal elements of a cluster, the distribution of verb clusters is also subject to various language internal properties. As is illustrated by the contrast between the examples in (2) vs. the ones in (3) (all from Swiss German), different types of constructions do not necessarily allow the same orders. While double modal constructions can be represented by the four orders in (2), a modal(1)-auxiliary(2)-participle(3) construction is best in the 1-3-2 order, accepted by some speakers in the 3-2-1 order and the 3-1-2 order, and excluded in the 1-2-3 order.

- |     |    |       |    |     |      |     |      |     |      |     |              |
|-----|----|-------|----|-----|------|-----|------|-----|------|-----|--------------|
| (3) | a. | wil   | er | si  | mues |     | gsee |     | ha   |     | Swiss German |
|     |    | since | he | her | must | (1) | seen | (3) | have | (2) |              |
|     | b. | %wil  | er | si  | gsee |     | ha   |     | mues |     |              |
|     |    | since | he | her | seen | (3) | have | (2) | must | (1) |              |

- c. %wil er si gsee mues ha  
 since he her seen (3) must (1) have (2)
- d. \*wil er si mues ha gsee  
 since he her must (1) have (2) seen (3)
- all: ‘that he must have seen her’

This article summarizes some of the many interesting empirical and theoretical contributions that have emerged from the research on verb clusters. Beginning with the empirical characterization—i.e., the determination of the (im)possible orders of verbal elements in different constructions and different languages and dialects—a phenomenon one is immediately struck by is the notable speaker, language, and dialect variation found in this area. It has long been known that among the West Germanic languages, closely related languages (in particular, languages with comparable syntactic properties) diverge significantly in the distribution of verb clusters. For instance, while Dutch and German are quite similar regarding the general layout of their syntactic structures, they show the exact opposite word orders in many verb cluster constructions. Apart from these well-known differences among languages, dialect studies conducted by several research groups have revealed that the distribution of verb clusters also shows significant micro-parametric variation and that in addition to the ‘standard’ word orders reported in grammars, many other orders are possible in certain dialects as well. Some aspects of this complex distribution and the variation attested in verb clusters will be presented in section 2.

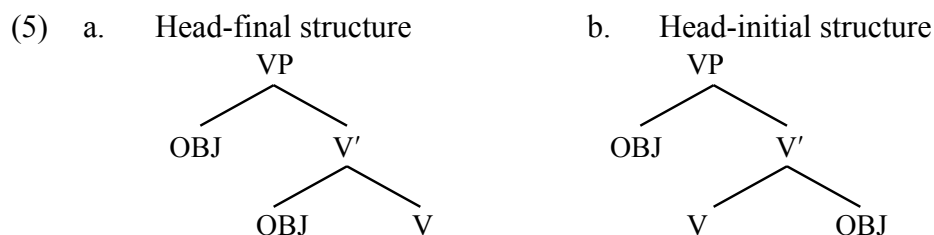
While this overview focuses on syntactic approaches to verb clusters, the challenges and limits of a syntactic treatment of the phenomenon will also be noted throughout the discussion. One of the major obstacles posed by verb clusters is that there is significant word order variation, both within and across languages, yet the meaning of the constructions are constant across speakers and languages. Accounting for the diverse distribution of word orders has proven challenging for many syntactic approaches, in particular for theories that require motivation for syntactic processes which is expressible in morphological and/or semantic terms. More generally, a question which has led to broader views of where to locate the reasons for verb cluster effects is the question of why verbs cluster and what the (deep) motivation for the phenomenon is. As is briefly summarized in section 4, this issue is addressed most exhaustively in multi-causal approaches according to which, in addition to certain syntactic properties, prosodic and pragmatic factors, and in some approaches also functional and processing factors play a role in regulating the distribution of word orders. Although it may be difficult to cover the entire spectrum of the verb cluster variation solely by syntactic considerations, this overview, among others, aims to highlight crucial generalizations pointing to the significance of syntactic factors in the distribution of verb clusters.

To see what some of the basic syntactic questions are that have been addressed in the works on verb clusters, we begin with a very simple example—an auxiliary-participle construction (i.e., a verb cluster involving two verbal elements). As is shown in (4), Dutch allows both of the two possible orders between these elements.

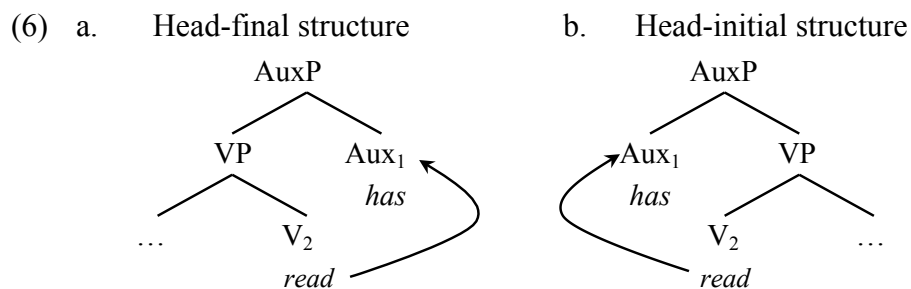
- (4) a. dat Jan het boek gelezen heeft Dutch  
 that Jan the book read (2) has (1)
- b. dat Jan het boek heeft gelezen  
 that Jan the book has (1) read (2)

A central syntactic question raised by the distribution in (4) is whether the two orders are both basic orders (whether the two verbs are freely generated in any order) or whether there is a derivational relation between them. The perhaps predominant (but not the only) view in the works on verb clusters is that only one of the orders in (4) is the basic order and that the other order is derived. The details of what the basic order is and how the reordering is derived, however, are far from being agreed upon and have lent to very engaged discussions among researchers interested in the topic.

One such issue is the question of whether languages like German and Dutch involve a head-initial or a head-final base structure. The traditional view is that the verb is base-generated to the right of its complement in Dutch and related languages (cf. (5a)), since objects generally precede the verbs in these languages. A different view, on the other hand, which has been inspired by Kayne's (1994) *Linear Correspondence Axiom* (LCA), is that all languages involve a basic head-initial structure (cf. (5b)) and that the final position of the verb is derived by leftward movement of the complement.



Assuming the different orders in verb clusters are not base-generated but derivationally related (although we will see that this is not necessarily the case), one obviously has to pick one of the structures as the basic structure. A question that has turned out to be rather challenging then is whether the distribution of verb clusters provides any indication for the superiority of one of the approaches in (5). As we will see in the course of this article, the mechanical derivation of the possible word orders is fairly trivial under both a head-final and a head-initial structure given generous assumptions about movement possibilities. To give a basic illustration, consider again the two orders in (4). Assuming a basic head-final structure as in (6a), nothing is required to derive the 2-1 order whereas the 1-2 order would require reordering such as rightward movement of the lower verb or verb phrase. Assuming a basic head-initial structure as in (6b), on the other hand, nothing is required to derive the 1-2 order, whereas the 2-1 order would require reordering such as leftward movement of the lower verb or verb phrase. Thus, as will be shown in detail in this overview, the question of whether one of the two approaches in (6) is to be chosen cannot be determined on purely empirical grounds.



Several questions also arise regarding the actual derivation of the different orders in verb clusters and the technical and theoretical nature of the arrows in (6). A very basic issue is whether the reorderings involve syntactic movement or some other form of linearization. Driven in part by the lack of motivation for different verb cluster orders, several recent approaches have shifted away from syntactic movement and tried to find other mechanisms to arrive at the different word order outputs (see section 3.1). If, on the other hand, reordering is considered to be syntactic movement, questions arising are whether the movement is head or phrasal movement and what the target positions, triggers, and motivation for the movements are.

Last but not least, an important empirical issue that is raised by verb clusters is the question of what elements are involved in what kind of clustering phenomena. As we have seen in the Swiss examples in (2) vs. (3), verb clusters are not homogenous constructs (in one and the same language), but rather, different orders depend crucially on the category of the elements involved. In general, clustering phenomena are found with auxiliaries and modal verbs. In some languages, other so-called *restructuring verbs* (which include for instance *try*, *begin*, *dare*) also participate in certain reordering phenomena. Since in many approaches, auxiliaries and modal verbs are distinguished from full main verbs, the distribution of verb clusters also raises interesting questions regarding the general architecture of clauses, questions about the classification of verbal elements (auxiliaries vs. main verbs, functional vs. lexical categories, thematic vs. non-thematic verbs), and the issue of restructuring.

This chapter presents the progress that has been made in answering these and many other questions. The works on verb clusters have brought to light a wealth of fascinating facts, and among the numerous proposals on the topic, many insightful explanations have been offered as answers to the issues mentioned. This chapter is organized as follows. Section 2 presents a summary of the empirical situation of verb clusters as attested so far in some West Germanic languages and dialects, and provides a descriptive account along the lines of the analysis developed in Haegeman and van Riemsdijk (1986). Section 3 is dedicated to the question of where and how verb clusters are derived. The main questions addressed are whether verb cluster orders are derived by syntactic movement or non-syntactic processes (section 3.1), whether verb cluster reordering involves head or phrasal movement (section 3.2.1), whether verb clusters enlighten the directionality debate (section 3.2.2), what non-verbal elements can interrupt a verb cluster (section 3.2.3), and what the motivation of verb cluster reordering is (3.3). Finally, the concluding section 4 speculates about one of the main open questions—why verbs cluster.

## 2. EMPIRICAL OVERVIEW

### 2.1 Verb cluster languages

The languages (including different varieties of these languages) which have been proposed to involve some form of verb clustering (though to very different degrees) are: Afrikaans, Dutch, Frisian, German, Hungarian, Swiss German, West Flemish, and possibly Yiddish. To define the set of languages that display verb cluster phenomena, various criteria have been invoked depending on the theoretical background assumptions. The most common criterion is word order: A language is a verb cluster language if it does not display a rigid word order pattern in multiple verb constructions—i.e., if the unmarked order of verbal elements is different from the underlying

ing order in at least one construction.<sup>2</sup> Languages that fall into this category (independently of whether one takes the underlying order to be a head-final or a head-initial order) are Afrikaans, Dutch, German, Swiss German, West Flemish (including all the dialects of these languages), as well as Hungarian. Setting aside Hungarian which differs from the other languages in many respects, the remaining languages have all traditionally been characterized as head-final. In fact, Haider (2003, et seq.) notes that verb clustering is only found in OV languages, and he relates this to an unwelcome computational complexity created by head-final structures, which can be resolved by reordering. While there is a strong tendency of verb clusters being restricted to OV languages, it is not clear whether there is an implicational relation between the two properties and the correlation between the head-final nature of these languages and verb cluster formation might also be challenged. First, apart from Hungarian, one could see Yiddish as a counterexample to the generalization that all head-initial Germanic languages display a rigid 1-2-3 order. Since Yiddish, which many researchers treat as a head-initial language, allows optional re-ordering of passive participles in multiple verb constructions (cf. den Besten and Moed-van Walraven 1986), Yiddish could be classified as a head-initial verb cluster language. Second, an OV order can of course not be seen as a sufficient condition for verb clustering, given that verb clusters are basically only found in West-Germania. But even within Germanic, it can be debated whether all head-final Germanic languages indeed display verb cluster phenomena. Frisian, for instance, features a rigid 3-2-1 order in modal and auxiliary constructions (i.e., configurations that typically show clustering effects in other languages; see sections 2.3 and 2.4). Assuming an underlying OV-structure, Frisian thus would not involve any verb cluster re-ordering. If, however, the empirical domain is extended to include *te/to/zu*-infinitives, certain re-orderings are obligatory (cf. de Haan 1992, 1993, 1996), and hence Frisian would qualify as a verb cluster language, allowing to maintain the generalization that all head-final Germanic languages are verb cluster languages.

In light of more recent approaches, in particular, approaches that are based on the *Universal Base Hypothesis* (cf. Kayne 1994), the ‘head-initial’ vs. ‘head-final’ distinction is not a grammatical notion since all languages are considered to be organized along a head-initial schema. Hence, the generalization noted above cannot be stated by referring to the directionality setting of the base structure. Thus, in these approaches, verb cluster languages do not share an underlying typological property but are rather defined by the overt/covert settings of various movement operations or the factors that trigger verb cluster formation itself.

## 2.2 Infinitivus Pro Participio (‘Infinitive for participle’; IPP)

A further commonality of the (Germanic) verb cluster languages that has been noted in the works on verb clusters is the so-called *Infinitivus Pro Participio* (‘Infinitive for participle’; IPP) effect. As illustrated in (7), Dutch modal verbs that occur in a perfective construction (i.e., under the auxiliary *have*) do not show up as participles but rather as infinitives (the same is the case for causative and perception verbs, as well as other restructuring verbs in Dutch).

- (7) a.    dat    Jan    het boek    heeft    kunnen    lezen    Dutch  
           that    Jan    the book    has (1)    can.IPP (2)    read (3)  
           ‘that Jan has been able to read the book’

<sup>2</sup> The notion of ‘unmarked’ word order is crucial to distinguish verb clusters from constructions like stylistic fronting in Icelandic or VP-topicalization.

- b. \*dat Jan het boek heeft gekund lezen  
 that Jan the book has (1) can.PART (2) read (3)  
 ‘that Jan has been able to read the book’

The IPP effect is found in Dutch, German, and West Flemish—i.e., languages that are also classified as verb cluster languages.<sup>3</sup> Thus, an interesting question arising from this first generalization is whether there is a strong or weak correlation between the IPP effect and verb cluster formation (i.e., whether it is the case that all languages displaying the IPP effect also involve verb cluster reordering and/or whether it is the case that all languages involving clustering also display the IPP effect).

Other languages that have to be considered in this respect are Afrikaans, Frisian, and Swiss German. Note first that Afrikaans and Swiss German which are other clustering languages (again independently of the base structure) appear to be irrelevant for the present issue, since these languages do not distinguish between infinitives and participles, and hence, it is not testable whether these languages display the IPP effect (cf. Robbers 1997 for Afrikaans and Schönenberger 1995 for Swiss). Turning to Frisian, it has been noted that Frisian only permits rigid descending (i.e., 3-2-1) orders in multiple verb constructions (except—as pointed out above—in constructions with *to*-infinitives); thus, under a head-final base structure, no reordering takes place in these constructions in this dialect. If the IPP effect is only found when re-ordering takes place we would expect that Frisian should not display the IPP effect. According to what is reported in the literature, this seems to be correct. It has been pointed out by a number of researchers (cf. for instance de Haan 1992, Hoekstra and Taanman 1996, Ijbema 1997) that Frisian lacks the IPP effect (in the descending order).<sup>4</sup> Thus, we can state a one-way generalization between languages that involve verb cluster re-ordering and languages that display the IPP effect: IPP is only found in languages that divert from the strict descending 3-2-1 order.

Although this correlation appears to be quite striking and is unlikely to be accidental, it is not clear what property of grammar it targets and what its importance is. The reason is that this correlation is an indirect generalization about languages and not about a causal relationship between the properties involved (verb cluster formation and the IPP effect). To strengthen the IPP/verb cluster correlation, it has therefore been suggested that there is in fact a direct causal relationship between the IPP effect and verb clustering (see for instance van der Meer 1990, den Dikken 1989). This step, however, has to be taken with some caution (see Ijbema 1997 for a critical overview of this issue). Although a detailed illustration of the different word orders will be postponed until section 2.4 (see in particular Table 2), it can be mentioned at this point already that i) the IPP effect is licensed in the presence and absence of verb cluster formation; and ii) verb cluster formation takes place in the presence and absence of the IPP effect.

Regarding the first point, it has been observed that the IPP effect occurs in all word orders that are (robustly) attested in verb clusters with three verbal elements (see section 2.4). More concretely, it occurs in all of the five orders: 1-2-3 (Dutch, West Flemish), 1-3-2 (German, Dutch

<sup>3</sup> As will be shown in section 2.4, these languages allow word orders that cannot be treated as basic word orders under either a head-final or a head-initial base structure. Thus, these languages clearly involve verb cluster reordering (in certain constructions) independently of what one assumes to be their base structure.

<sup>4</sup> According to a preliminary study conducted by P. Ackema, there is some potential variation among Frisian speakers regarding the possibility of the IPP effect in descending orders. However, the empirical distribution is not conclusive, since, as pointed out by G. de Haan (p.c.), there is a strong influence from Dutch which could interfere with the data.

varieties), 3-1-2 (German, Dutch varieties), 2-3-1 (West Flemish), and 3-2-1 (Dutch varieties, German varieties). It is particularly noteworthy that, contrary to what has been claimed in the traditional literature, the IPP effect also occurs in the 3-2-1 order, as documented in (8).

- (8) a. ..., omdat de Schrieverskring Achterhoek en Liemers Achterhoeks  
 ..., because the writer's circle Achterhoek and Liemers  
 in de moderspraake schrievm willn / ewild had 3-2-1  
 in the mother language write.INF want.IPP / want.PART had  
 '...because the writer's circle Achterhoek and Liemers had wanted to write in the na-  
 tive language.' [Blom and Hoekstra 1996: (8a,b); translation B. Moskal]
- b. ...daß d Anna ons bsuacha wella had Swabian  
 ...that the Anna us visit.INF want.INF has 3-2-1  
 'that Anna wanted to visit us' [Steil 1989:16, cited via Zwart 2007: 79, (12)]

Although the 3-2-1 word order is less common in IPP constructions in German and Dutch, it cannot be rejected as ungrammatical in these languages. First, in a questionnaire study on German verb clusters, the German example in (9) was judged grammatical by 25.9% of the German speakers from Germany (39.3% judged it as ?, and only 34.8% rejected it as ungrammatical) and by 45% of the German speakers from Austria (with 25% judging it as ?, and only 30% as ungrammatical).

- (9) Ich frage mich... 'I ask myself...' German  
 %warum der Kommissar den Fall nicht lösen können hat 3-2-1  
 why the inspector the case not solve.INF can.IPP has  
 'I wonder why the inspector has not been able to solve the case.'  
 [Wurmbrand 2004b: 55-56]

Similarly, Bader and Schmid (2009) show in a series of experiments on auxiliary-modal-infinitive constructions in German that the 3-2-1 order is judged as grammatical by up to 28% of the speakers they tested. Crucially for the discussion of the IPP effect, they also show that there is no difference in the grammaticality rate between constructions with IPP (infinitives) and constructions with participles.

Lastly, extensive investigations of Dutch dialects conducted by the *Meertens Institute* in a large-scale dialect project (yielding volumes I and II of the *Syntactische Atlas van de Nederlandse Dialecten/Syntactic Atlas of the Dutch Dialects* [SAND]) has shown that IPP constructions are allowed in the 3-2-1 order in Northern Dutch dialects, presumably due to dialect influences between German and Dutch. According to Barbiers (2005), the SAND project involved two surveys—a written pilot questionnaire in which 368 informants in 321 distinct locations were tested, and 267 follow up oral interviews involving two to three speakers in at each location. The numbers of 3-2-1 orders provided in the written questionnaires for auxiliary-aspectual *go*-infinitive constructions was 40/321 (12.5%). In the oral interviews, the number was 46/267 (17.2%). Oral interviews also produced 32/267 (12%) 3-2-1 orders in auxiliary-modal-infinitive constructions (numbers from Barbiers 2005: 236, Table 1).

The distribution of auxiliary-modal-infinitive constructions as reported in the final atlas (SANDII, Barbiers et al. 2008) is given in (10) (the numbers refer to locations in which the different orders are attested). As expected (see section 2.4), the most common order in Dutch IPP constructions is the 1-2-3 order in (10b), but as shown in (10h,g), there are dialects that (also)



allow the 3-2-1 order. Although the number of dialects that allow the 3-2-1 order in Dutch is obviously significantly smaller than the number of dialects accepting the standard 1-2-3 order, the existence of the former can nevertheless not be ignored. Importantly again, the ratio between 3-2-1 orders with and without IPP is basically the same, with the IPP version (in bold) being even slightly higher.

(10) Barbiers et al. (2008); SANDII: 20 (map 1.3.2.3.4)

a.	Vertel mij eens wie zij had kunnen/gekund roepen	Dutch
	tell me just who she had.FIN can.IPP/PART call.INF	
	‘Just tell me who she could have called.’	
b.	V1—V2.IPP—V3	had kunnen roepen 185
c.	V1—V3—V2.IPP	had roepen kunnen 7
d.	V1—V3—V2.PART	had roepen gekund 1
e.	V2.IPP—V3—V1	kunnen roepen had 18
f.	V3—V1—V2.IPP	roepen had kunnen 2
g.	V3—V1—V2.PART	roepen had gekund 3
h.	<b>V3—V2.IPP—V1</b>	<b>roepen kunnen had 19</b>
i.	V3—V2.PART—V1	roepen gekund had 17

The fact that the IPP effect arises in all verb cluster orders shows that verb cluster formation or reordering cannot be seen as a necessary condition for licensing or triggering IPP morphology. Independently of what one considers to be the base structure for the languages under consideration, there are constructions displaying the IPP effect but lacking verb cluster reordering. Thus no condition of the form ‘If IPP, then word order x’ can be formulated, and IPP licensing must be possible independent of word order.

Regarding the second point, we will see that verb cluster formation takes place in all types of constructions (e.g., double modal constructions, auxiliary-participle constructions) and is not restricted to IPP-constructions. Furthermore, Blom and Hoekstra (1996) (see also Ijbema 1997) note that in Achterhoeks reordering takes place in auxiliary-modal constructions yielding the 1-3-2 order as in (11), however, the IPP effect is optional in these constructions. Thus, the IPP effect cannot be considered to be the cause for clustering.

- (11) a. ..., omdat ik gaorne/geerne had komm wiln Achterhoeks  
 ..., because I happily had come.INF want.IPP 1-3-2  
 ‘...because I would have wanted to come’  
 [Blom and Hoekstra 1996: (7a); translation B. Moskal]
- b. ..., omdat ik gaorne/geerne had komm ewild Achterhoeks  
 ..., because I happily had come.INF want.PART 1-3-2  
 ‘...because I would have wanted to come’  
 [Blom and Hoekstra 1996: (7b); translation B. Moskal]

There is, however, one generalization pointed out in Zwart (2007) that appears to hold: The IPP effect occurs without exception in 1-2-3 clusters (Zwart 2007: 79, (14)), unless VP3 can be assumed to be extraposed. Zwart suggests to derive this generalization by the assumption that the IPP effect is obligatory whenever “a participle takes an infinitive in its complement domain”

(Zwart 2007: 84, (39)). This is paired with the assumption that the lexical VP3 can be either base-generated as a complement of the modal verb, in which case the cluster 1-2-3 is created and the IPP is obligatory, or the lexical VP could be merged outside the [1-2] complex, yielding the orders 3...-[1-2], [1-2]...-3, 3...-[2-1], and [2-1]...-3. In the latter, the IPP effect is not obligatory. While it is not entirely clear, how 1-3-2 orders (in particular, the optionality of the IPP in Achterhoeks) are derived in this system, the approach is supported by the data available so far for 2-3-1 constructions, which appear to consistently occur with the IPP effect (see (10)). In this order, the [2-3] complex presumably forms a unit, hence the infinitival VP3 would be in the complement domain of the modal/aspectual V2, which then requires the IPP effect according to the assumption above.

Lastly, Zwart (2007) also shows that the IPP effect is only found in languages with circumfixal participles (*ge-V-en*)—i.e., there is a one-way implication: if IPP then ‘*ge*’. I return to this implication and further generalizations and speculations in the concluding section 4. In the sections to follow, a summary of the verb cluster patterns as documented for the major West Germanic languages and dialects will be provided (Hungarian is set aside here; the reader is referred to Koopman and Szabolcsi 2000). While this empirical overview has the aim to present the verb cluster patterns for all major constructions in any given language, it also has to be kept in mind that it is by no means an exhaustive characterization—in particular, many interesting microparametric differences cannot be distinguished here. The overview will then form the basis for the subsequent review of the theoretical accounts of the verb cluster phenomenon.

### 2.3 Two-verb clusters

The constructions involving verb clusters with two verbs are auxiliary-participle constructions (e.g., *John has left*) and auxiliary/modal-infinitive constructions (e.g., *John will/must leave*). Table 1 summarizes the possible word orders for two-verb clusters in a number of West Germanic languages/dialects; data illustrating these orders are provided in section 2.5. This summary does not include diachronic data. The reader is referred to the detailed descriptions in Coussée (2008) for older varieties of Dutch, and Sapp (2006, 2011) for Early New High German and Middle High German.

Table 1 Verb clusters with two verbal elements

Language	AUX-PART	MOD-INF
Afrikaans	2-1	1-2 <sup>a</sup>
Dutch (1 = finite)	1-2	1-2
	2-1	2-1
Dutch (1 = non-finite)	1-2	1-2
	2-1	
Frisian	2-1	2-1
German	2-1	2-1
		[1-2]
Swiss-1	2-1	2-1
		1-2
Swiss-2	2-1	2-1
	1-2	1-2
West Flemish	2-1 <sup>b</sup>	1-2

*Notes:*

Orders in brackets are attested but very restricted in the language for which they are listed; whether these orders should be considered as possible orders for these languages will depend on further refinement and dialect separation.

<sup>a</sup> *Come* + INF allows the 2-1 order (Robbers 1997).

<sup>b</sup> 1-2 is possible if 2 is followed by an extraposed PP or CP (Haegeman 1995: 53; 1998b: 294).

*Sources:*

[References are given only for cases that might be considered to be controversial or are not well established.]

**Dutch** (SANDII): *is gestorven* in ‘She does not know that Mary has died yesterday’—1-2 (71), 2-1 (247); *heeft verteld* in ‘That is the man who has told the story’—1-2 (93), 2-1 (181); *hebben geroepen* ‘That is the man who they have called’—1-2 (9), 2-1 (254); *mag zien* as in ‘...you also may not see it’—1-2 (201), 2-1 (113).

**Frisian**: Quoted from Zwart (1996) and Ijbema (1997); confirmed by the results of a questionnaire conducted by Peter Ackema.

**German**: The predominant order is the 2-1 order; but occurrences of 1-2 orders exist as well, as is documented in Patocka (1997) for Austrian, in Sapp (2006, 2011) for Austrian and Swabian dialects, and in Dubenion-Smith (2010) for the West Central German dialects of Germany. Based on transcripts of 187 dialect recordings from the *Datenbank Gesprochenes Deutsch* ‘Spoken German Databank’ (maintained by the *Institut für Deutsche Sprache*; the transcripts were part of the Zwirner corpus compiled between 1955-1970), Dubenion-Smith 2010 notes that of the 1,326 tokens of two-verb clusters, the 1-2 order was recorded for 159 items, 12%. The rate of 1-2 orders is higher in the MOD-INF construction (25.1%) than in AUX-PART constructions (7.6%).

**Swiss-1**: Schönenberger (1995); E. Haeberli (p.c.); H. van Riemsdijk (p.c.). Sapp (2011: 138-150) showed in a detailed experiment on Zurich German that the 2-1 order is strongly preferred in perfect constructions (but some speakers also accepted the 1-2 order); the 2-1 and 1-2 orders are both possible in the MOD-INF constructions, and in Sapp’s results no preference for the 1-2 order was found (contra Lötscher 1978). However, the orders show sensitivity to focus: the 1-2 order is favoured under narrow object focus and VP focus, whereas the 2-1 order is favoured when MOD is focused.

**Swiss-2**: M. Schönenberger (p.c.; Bernese), Hsiao (1999, p.c.).

Given the distribution in Table 1, dialects can be grouped into five types: i) dialects with rigid 2-1 order (German, Frisian); ii) dialects with rigid 2-1 order in auxiliary-participle constructions, but flexible order in modal-infinitive constructions (Swiss-1); iii) dialects with flexible order in both constructions (Swiss-2, Dutch when V1 is finite); iv) dialects with flexible order in auxiliary-participle constructions and finite modal-infinitive constructions, but rigid 1-2 order in non-finite modal-infinitive constructions (Dutch when V1 is non-finite); and v) dialects with rigid 2-1 order in auxiliary-participle constructions and rigid 1-2 order in modal-infinitive constructions (West Flemish, Afrikaans).

The following three generalizations emerge from the languages/dialects investigated so far. First, the distribution of auxiliary-participle constructions does not (necessarily) coincide with the distribution of modal-infinitive constructions. In Afrikaans, Dutch (when V1 is non-finite), Swiss-1, and West Flemish, the orders vary with respect to the constructions involved—i.e., there is no general (1-2 or 2-1) word order schema for two-verb clusters in these languages. Note in particular, that in Afrikaans and West Flemish where only one order is possible in each construction, the order required in the auxiliary-participle construction is excluded in the modal-infinitive construction (and vice versa). Second, the 2-1 order is possible in all dialects for auxiliary-participle constructions. In other words, no dialect requires the 1-2 order in this construction (this is again different in the modal-infinitive construction). Third, if the 1-2 order is possible in an auxiliary-participle construction it is also possible in the modal-infinitive constructions (but not vice versa). What appears to be unattested (at least to this end) are first, dialects that display clustering effects but only allow a 1-2 order for auxiliary-participle constructions, and second, languages that allow a flexible order for auxiliary-participle constructions but a rigid 2-1 order for modal-infinitive constructions.

## 2.4 Three-verb clusters

The five major types of constructions involving verb clusters with three verbal elements are: double modal constructions, two types of auxiliary-modal constructions, modal-auxiliary constructions, and double auxiliary constructions.<sup>5</sup> Examples are given in an abstract form in (12) (English words are used to illustrate the constructions, but of course, some of the examples would be impossible in English, since modals can only be part of a multiple verb construction when they are the highest verbal element in English). Constructions involving infinitival complements with an infinitival marker (*to*, *zu*, *te* etc.) are not considered in this section (see 3.2.1.2).

(12)	a.	John	must	(1)	can	(2)	sing	(3)	MOD-MOD-V
			FIN	(1)	INF	(2)	INF	(3)	‘John must be able to sing’
	b.	John	will	(1)	must	(2)	sing	(3)	AUX-MOD-V
			FIN	(1)	INF	(2)	INF	(3)	‘John will have to sing’
	c.	John	has	(1)	must	(2)	sing	(3)	AUX-MOD-V
			FIN	(1)	IPP	(2)	INF	(3)	‘John has had to sing’
	d.	John	must	(1)	have	(2)	sung	(3)	MOD- AUX-V
		John	must	(1)	be	(2)	elected	(3)	
			FIN	(1)	INF	(2)	PART	(3)	
	e.	John	has	(1)	been	(2)	elected	(3)	AUX-AUX-V
			FIN	(1)	PART	(2)	PART	(3)	

The distribution of these constructions in various West Germanic languages and dialects is given in Table 2 (for examples the reader is referred to section 2.5). The rows labeled *Standard* provide the orders that have traditionally been assumed to be the possible orders in Dutch and German. However, detailed empirical studies conducted in the last decade have shown that this is a gross oversimplification. The results of those studies are given in the rows labeled *Dialects*. These rows refer to the superset of various dialects (i.e., there are speakers who may accept all the orders listed, most likely with preferences among the orders, but there are also speakers who only accept subsets of these orders).<sup>6</sup> If a cell contains more than one option, the word orders are given in order of preference where the highest order corresponds to the most preferred or most frequent option in a dialect group. It is very likely that the higher numbers of possible orders in the larger dialect groups (Dutch, German) are the result of extensive bilingualism—if there is widespread micro-parametric variation among several local dialects, each of which may have a particular preferred word order, the superregional possibilities could be seen as the sum of the options found in several dialects. The further discussions will largely ignore preferences or frequencies. Lastly, it should be noted that column 4 (AUX-MOD-V; FIN-IPP-INF) refers to the orders in the second construction in (12c) irrespective of whether a language overtly displays the IPP effect. In Afrikaans, for instance, infinitives and particles are non-distinct and hence it cannot be

<sup>5</sup> Not considered for this overview are causative and perception verb constructions as well as double auxiliary constructions of the form *John will have left* since these constructions are less discussed in the literature.

<sup>6</sup> See Hsiao (1999, 2000), Schmid and Vogel (2004), Wurmbrand (2004), Barbiers et al. (2008), Bader and Schmid (2009), Dubenion-Smith (2010), and Schallert (2014) for further details about a more fine-grained distribution of word orders in different dialects.

concluded that the modal verb occurs in the IPP form. However, as the different word orders show, it is still necessary to distinguish this construction from others.

Table 2 Verb clusters with three verbal elements

Language	MOD-MOD-V FIN-INF-INF	AUX-MOD-V FIN-INF-INF	AUX-MOD-V FIN-IPP-INF	MOD-AUX-V FIN-INF-PART	AUX-AUX-V FIN-PART-PART
Afrikaans	1-2-3	1-2-3	2-3-1	1-3-2 3-1-2 <sup>a</sup>	N/A
Dutch Standard	1-2-3	1-2-3	1-2-3	1-2-3 3-1-2 1-3-2	[ <sup>?</sup> 3-1-2] <sup>b</sup> [ <sup>?</sup> 1-3-2] <sup>b</sup>
Dutch Dialects	1-2-3 3-1-2 3-2-1 1-3-2		1-2-3 3-2-1 2-3-1 1-3-2 [3-1-2]	3-1-2 1-3-2 1-2-3 3-2-1	
Frisian	3-2-1	3-2-1	3-2-1 [1-2-3] <sup>c</sup>	3-2-1	3-2-1
German Standard	3-2-1	3-2-1 1-3-2	1-3-2	3-2-1	3-2-1
German, Austrian Dialects	3-2-1 1-3-2	3-2-1 1-3-2 3-1-2 [1-2-3] <sup>d</sup> [2-1-3] <sup>d</sup>	1-3-2 3-1-2 3-2-1 [1-2-3]	3-2-1 1-3-2 3-1-2 <sup>e</sup>	3-2-1 3-1-2
Swiss Dialects	1-2-3 3-2-1 1-3-2 3-1-2	N/A	1-2-3 1-3-2 3-1-2 [2-1-3] <sup>f</sup> [2-3-1] <sup>f</sup>	1-3-2 3-2-1 3-1-2	3-2-1
West Flemish	1-2-3		1-2-3 <sup>g</sup> 2-3-1 <sup>g</sup>	1-3-2 3-1-2	3-2-1 1-3-2

*Notes:*

Orders in brackets are attested but very restricted in the language for which they are listed; whether these orders should be considered as possible orders for these languages will depend on further refinement and dialect separation.

<sup>a</sup> 3-1-2 is only possible when 3 is a passive participle.

<sup>b</sup> Double-participle constructions involving *geworden* are generally considered marginal; for double-participle constructions involving *get+* PART (i.e., ‘has gotten PART’), the 1-3-2 order is mentioned in Robbers (1997: 124). Since no systematic empirical characterization is available and this construction is very marked (constructions involving the auxiliary *geworden* are rejected by most Dutch speakers), it will be ignored.

<sup>c</sup> IPP is not obligatory; some speakers do not use IPP; 1-2-3 is mentioned in Hoekstra and Taanman (1996) and Ijbema (1997) for certain constructions (perception verbs and aspectual auxiliaries) in West Frisian.

<sup>d</sup> The 1-2-3 and 2-1-3 orders are attested in the survey conducted by Schmid and Vogel (2004).

<sup>e</sup> The 3-1-2 order is the most frequent order in MOD-PASS-PART ‘*could be taken*’ constructions in the Zwirner corpus of West Central German dialects transcribed by Dubenion-Smith (2010).

<sup>f</sup> Lötscher (1978) and Salzmann (2013a, b) give examples with perception verbs, causative *la* ‘let’, *aafaa* ‘start’, *hölffe* ‘help’, and the doubling particle *go* in the 2-3-1 and 2-1-3 orders. Since constructions with these verbs could

involve a non-clustering configuration (see section 3.1) and these orders are sharply excluded with typical cluster verbs (modals), 2-1-3 and 2-3-1 are not treated as verb cluster orders here.

<sup>§</sup> 1-2-3 is obligatory when the auxiliary is in the past or has a negative marker attached; 2-3-1 is obligatory when the auxiliary is non-finite (Haegeman 1994, 1995, 1998b).

#### Sources:

**Dutch:** The 1-3-2 and 3-2-1 orders for the MOD-AUX-V construction are mentioned in Zwart (1996), Ijbema (1997), and Robbers (1997); the authors already noted variation among Dutch speakers, which is confirmed in the SANDII. The orders (top to bottom) given in the *Dutch dialects* row represent the frequency as reported in the oral interviews conducted by the SAND project (see section 2.2). The AUX-MOD-V and AUX-Asp (*go*)-V constructions show different frequencies; the order in the table represents the frequency of the former; the 3-1-2 order is more frequent in the aspectual (*go*) construction, but according to Barbiers (2005) this order does not show a clear pattern in the dialects investigated and is therefore omitted in his account. The distribution given in Barbiers et al. (2008) is as follows (numbers below 7 have been omitted; see also Barbiers 2005 for further details):

AUX-MOD-IPP-V (*had kunnen roepen* ‘has can call’): 1-2-3 (185), 3-2-1 (19), 2-3-1 (18), 1-3-2 (7) (also (10))  
 AUX-*go*-V (*is gaan swemmen* ‘is go swim’): 1-2-3 (165), 2-3-1 (128), 3-2-1 (47), 3-1-2 (18)  
 MOD-MOD-V (*moet kunnen swemmen* ‘must can swim’): 1-2-3 (242), 3-1-2 (83), 3-2-1 (37), 1-3-2 (34)  
 MOD-AUX-V (*moet hebben gemaakt* ‘must have done’): 3-1-2 (186), 1-3-2 (163), 1-2-3 (91), 3-2-1 (48).

**Frisian:** Quoted from De Haan (1993), Zwart (1996), and Ijbema (1997); confirmed by P. Ackema (p.c.).

**German:** The orders are in part the result of a questionnaire-based study conducted by the author in 1999–2000 (approx. 100 informants; see Wurmbrand 2004b for detailed results). The 1-3-2 order is documented for the MOD-MOD-V construction in Den Besten and Edmondson (1983), Broekhuis (1992: 189), and Koopman and Szabolcsi (2000), and confirmed by a subgroup of the informants of the questionnaire-based study. The 1-3-2 order is noted for the MOD-AUX-V construction in Zwart (1996) but confirmed by only very few of the informants (mainly speakers from Vorarlberg). The 3-1-2 order is reported in Patocka (1997) as the unmarked order in Austrian German, which was robustly confirmed for both AUX-MOD-V constructions by the informants of the questionnaire-based study (the 3-1-2 order has not been tested for the other modal constructions). In the Zwirner corpus of spoken West Central German dialects transcribed by Dubenion-Smith (2010), the 3-1-2 order is also the most frequent order in IPP constructions, as well as in MOD-PASS-PART constructions. The 1-2-3 order for the IPP construction was accepted by only a very small group of informants and was rejected in most non-Allemanic dialects; a systematic documentation of the variation in the IPP construction is also available in Hsiao (1999).

**Swiss:** M. Schönenberger (p.c.), E. Haeberli (p.c.), H. van Riemsdijk (p.c.), and five other consultants. The auxiliary *will* is not used in Swiss dialects and therefore this construction could not be tested. A systematic dialect characterization (and in particular, a survey of how the distinction between Swiss-1 and Swiss-2 in two-verb clusters compares to the options in three-verb clusters) is not available at this point.

Based on the distribution in Table 2, some general points can be noted. First, of the six possible combinations involving three verbal elements, five orders are robustly attested. One order, the 2-1-3 order, however, is highly infrequent, perhaps impossible, as an order in verb clusters (it is found in extraposition contexts, which display different properties from genuine verb clusters; see Rutten 1991, Robbers 1997, Wurmbrand 2001, among others). I return to the 2-1-3 order in section 3.1.

Second, the 2-3-1 order is only found in the IPP-construction. The nature of 2-3-1 orders has been subject to some debate. Several authors have noted that its distribution is marked and shows a lower frequency than other orders (Svenonius 2007, Abels 2013). While, as shown in Table 2, the 2-3-1 order is restricted to AUX-MOD-V constructions, it can also be observed that it is the only order possible in this construction in Afrikaans. Similarly, in West Flemish there are contexts in which the 2-3-1 order is obligatory (namely, when the auxiliary is non-finite; see Haegeman 1998b: 276, (25)). Furthermore, as documented in the SAND study (see the numbers below Table 2), Dutch IPP constructions are more frequent in the 2-3-1 order than in the 1-3-2 and 3-1-2 orders (Barbiers et al. 2008). Lastly, Salzmann (2013a) provides 2-3-1 constructions in Swiss German, which appear to be unmarked.

Third, as is evident from Table 2, verb cluster reordering is not restricted to a particular

type of construction (such as the IPP-construction), but is found in all types of constructions. Since the fully ascending and descending orders do not necessarily involve reordering (i.e., the 1-2-3 and 3-2-1 orders could be treated as basic orders depending on whether one takes a language to follow the head-initial or head-final schema), we restrict our attention for the moment to the 1-3-2, 3-1-2, and 2-3-1 orders, which clearly involve some sort of reordering under both the head-final and the head-initial approach.<sup>7</sup> As the table shows, at least one language displays one or more of these orders in each construction: Dutch and German dialects and Swiss German in the double modal construction; all varieties of German in the AUX-MOD-V construction; certain Dutch dialects, Afrikaans, all varieties of German, Swiss German, and West Flemish in the IPP-construction; Afrikaans, Dutch, certain German dialects, Swiss German, and West Flemish in the MOD-AUX-PART construction; and Dutch and West Flemish in the double participle construction. The fact that reordering is not restricted to one construction has important consequences for the theoretical approaches to verb clusters. As mentioned in section 2.2, since verb cluster reordering is neither restricted to IPP-environments nor necessary in IPP-constructions, it cannot be assumed that a causal relation holds between these two properties. This point is particularly important for (Standard) German. Even if we ignore the dialectal variation within German for a moment, we see that reordering cannot be restricted to the IPP constellation. The IPP-construction is special in Standard German in that it requires obligatory reordering which contrasts with the otherwise typical 3-2-1 order in that language. However, what has often been ignored (but see Kathol 1996, 1998a, b for a laudable exception) is that the special reordering is *not* restricted to IPP-constructions but also found in auxiliary-modal constructions in which the auxiliary is the future element *werden* ‘will’. Importantly, these constructions are not IPP-constructions, but involve an infinitive which is selected by the future element. The correct generalization regarding Standard German is thus that the 3-2-1 order is obligatory in all constructions except in auxiliary-modal constructions, and the 1-3-2 order cannot be attributed to the IPP property but has to be seen as a special property of an auxiliary-modal constellation.

Fourth, it is also obvious from the distribution in Table 2, that verb cluster formation cannot be seen as a simple rule or operation that arranges verbs in a cluster according to some language specific hierarchical schema (such as ‘the lowest verb precedes/follows the n-highest verb’). Rather, the distribution of verbal elements is crucially dependent on the type of construction. This is most strikingly the case in West Flemish. West Flemish allows all five orders attested in verb clusters, however, the distribution of these orders is severely restricted by the type of construction: the 1-2-3 order is only possible in the double modal construction and certain IPP-constructions; the 1-3-2 order is only possible in participle constructions; the 3-1-2 order is only possible in the MOD-AUX-PART construction; the 3-2-1 order is only possible in the double auxiliary construction; and finally, the 2-3-1 order is restricted to IPP-constructions. Thus, an account of the distribution of word orders in multiple verb constructions has to take into account the language-specific and construction-specific nature of this phenomenon.

In the next section, the generalizations regarding the distribution of verb clusters in the languages/dialects mentioned will be fleshed out. The (descriptive) generalizations will be stated

---

<sup>7</sup> This is not entirely correct. If a system allows base-generation of both head-initial and head-final VP-structures within on language, the 1-3-2 and 2-3-1 orders could also be base-generated and would thus not require any reordering. Under such a view, only the 3-1-2 and 2-1-3 (if it exists) orders would be clear cases of cluster formation. As shown in Table 2, the 3-1-2 order also exists in at least one language/dialect in all but the double participle construction, and it is also the most frequent order in certain dialects and constructions. Thus, the point made in the text still holds even for systems which allow mixed directionality in the generation of verb clusters.

from both a head-final and a head-initial perspective in the framework of the analysis of Haegeman and van Riemsdijk (1986).

## 2.5 Inversion patterns and rules

One of the first works that takes into account the variation in the distribution of verb clusters across West Germanic is the *reanalysis* approach suggested by Haegeman and van Riemsdijk (1986). The account consists of two parts. First, the authors propose that in certain constructions (namely in restructuring constructions that show verb cluster phenomena) the underlying structure can be ‘reanalyzed’. Reanalysis is essentially a re-bracketing procedure for syntactic structure which reduces the distance between the elements in a cluster and unifies elements that are further apart in the basic structure. The second part of the analysis is the postulation of PF-inversion rules that specify which elements can, cannot or have to invert in a reanalyzed structure. As we will see in more detail in section 3.1, in contrast to syntactic movement approaches to the verb cluster phenomenon, Haegeman and van Riemsdijk suggest that reordering operations are not operations of syntax proper; rather, reordering operations are post-syntactic morphophonological operations (see also Wurmbrand 2004a, b, Salzmann 2013a, b). The inversion parameters suggested by Haegeman and van Riemsdijk (1986) are summarized in (13). As can be seen, the rules consist of specifications regarding the complexity of the inverting elements (cf. (13a)), the category of the inverting elements (such as modal, auxiliary etc.), a specification regarding (non-)optionality, and specifications regarding the node dominating the inverting elements (cf. (13d) which states that the node  $V\alpha$  that dominates the inverting elements has to or does not have to be the highest node in a verb cluster).

- |  |                                       |
|--|---------------------------------------|
| (13) Inversion: Main Parameters                                | Haegeman and van Riemsdijk (1986:426) |
| a. The nonhead must be (non)branching or need not be branching |                                       |
| b. The head of V must be V-AUX or V-MOD or is unrestricted     |                                       |
| c. Inversion is optional or obligatory                         |                                       |
| d. $V\alpha$ is maximal or unrestricted                        |                                       |

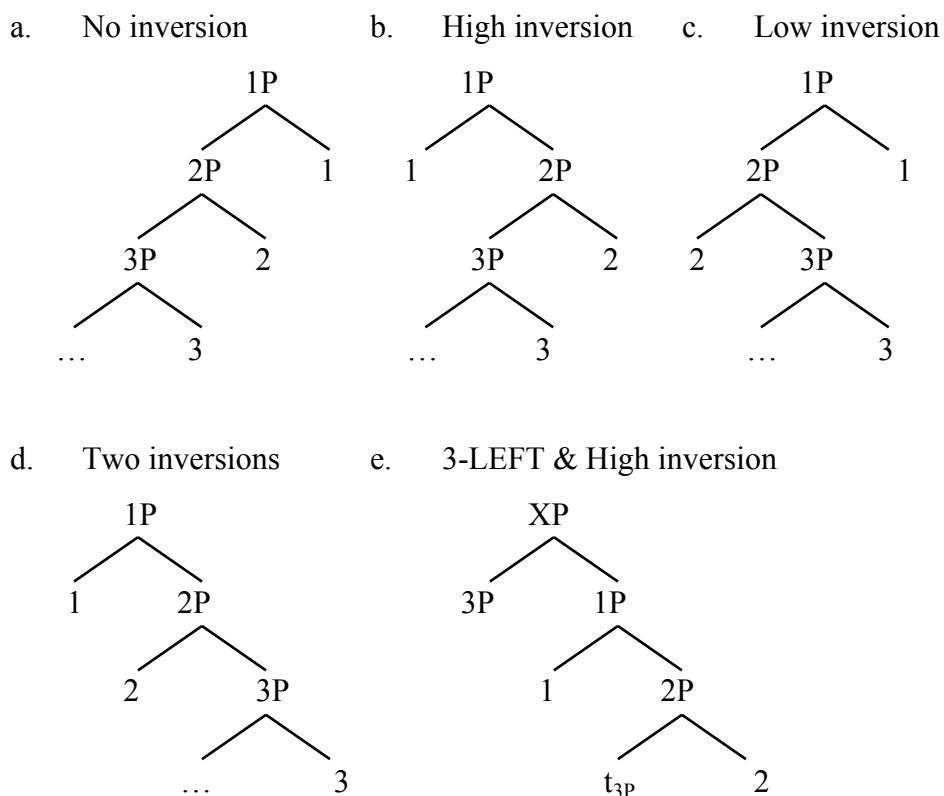
Since the Haegeman and van Riemsdijk system offers a straightforward way to formulate the conditions regulating verb cluster formation in the different languages, it will be used here (in a slightly modified way) to illustrate the specifications necessary to capture the distribution of the verb clusters in Table 1 and Table 2. Whether the first step in a verb cluster configuration—i.e., the step that creates mono-clausal structures—involves reanalysis as Haegeman and van Riemsdijk suggest or any other mechanism that has been suggested to account for clause union or restructuring constructions will not be essential for the discussion here (see Wurmbrand 2001 for an overview). Since Haegeman and van Riemsdijk’s analysis is coached in a head-final approach, the inversion rules will be provided only for a head-final system in this section, but as has been shown in Wurmbrand (2006), it is equally possible to define inversion rules in a head-initial system. Furthermore, throughout section 3, various verb cluster accounts based on a head-initial structure will be presented. The assignments and values of features needed to cover the empirical distribution under such an approach will be very similar to what would need to be stated as inversion rules under a head-initial base structure. To avoid replication, this section therefore only summarizes the head-final assumptions.

The main idea of inversion is that under certain circumstances, two sister nodes are flipped



with each other (see also Williams 2004 for a similar implementation of this idea). Inversion in this sense is reminiscent of precedence constraints as developed in the HPSG framework by, for instance, Kathol (1996, 1998a, b) and Meurers (2000) among many others. As illustrated in (14), inversion of sister nodes derives straightforwardly the orders 3-2-1, 1-2-3, 1-3-2, and 2-3-1 in verb clusters with three verbs. The 3-2-1 order in (14a) is the basic structure which does not involve any inversion; the 1-3-2 order is derived by inverting the highest verb with its sister (cf. (14b)); the 2-3-1 order is derived by inverting the middle verb with its sister (cf. (14c)); and the 1-2-3 order is derived by inverting both the highest and the middle verb with their respective sisters (cf. (14d)). Regarding the 3-1-2 order, simple inversion of sister nodes is not sufficient and recourse to an additional operation has to be made (this would also be the case in a head-initial base structure). One option would be to assume that the lowest verb phrase undergoes some form of leftward movement prior to or independent of verb cluster inversion. As illustrated in (14e), if this form of leftward movement (which I will refer to as 3-LEFT) is followed by high inversion, the 3-1-2 order is derived in a head-final structure.<sup>8</sup>

(14) Head-final derivations



Importantly, if inversion under sisterhood and leftward movement of the lowest verb phrase are the only tools available to derive verb cluster constructions (modulo object movement etc.), the lack of the 2-1-3 order can be accounted for. To derive this order, rightward movement of the

<sup>8</sup> For this basic illustration of the orders attested in verb clusters, the distribution of objects in verb clusters will be ignored. One might imagine that inversion in cases such as (14b,c,d) should result in a structure in which the object occurs between the verbs of a cluster. Although this is possible in certain languages and constructions (see 3.2.3), the more common effect is that the object is not in its base position but somewhere to the left of the cluster.

lowest verb phrase would be necessary, and the 2-1-3 order would only be possible if a form of VP-extraposition is licensed (see also section 3.1).

Armed with these tools, let us now turn to the specifications required to account for the distribution of verb clusters in the languages mentioned in Table 1 and Table 2. Assuming again a head-final approach, the inversion rule is given in (15) and the conditions under which inversion applies are listed in Table 3.

(15) Head-final inversion rule



iff the conditions in Table 3 hold.

Table 3. Inversion rules (head-final base)

Language	Inversion rules	Optionality	Other
Afrikaans	X=MOD	Obligatory	<i>Will</i> : MOD Passive participles: optional 3-LEFT
Dutch Standard	A: Y=INF B: Y=PART	A: Obligatory B: Optional	A: optional in 2-1.FIN PART: optional 3-LEFT
Dutch Dialects	A: X is [+FIN] B: Y is [-FIN]	Optional	Optional 3-LEFT [except IPP?] IPP: underspecified for [ $\pm$ FIN]
Frisian	—	—	—
German Standard	A: X=AUX & Y=MOD	Obligatory	<i>Will</i> : MOD or AUX
German, Austrian Dialects	A: X=AUX (& Y=MOD) B: X=Mod.FIN	Optional	AUX-(X)-V: optional 3-LEFT
Swiss Dialects	A: X=AUX & Y=MOD B: X=MOD.FIN C: Y=INF	A: Obligatory B/C: Optional	Optional 3-LEFT
West Flemish	A: X=operator (MOD, TENSE, NEG) & Y=INF B: X=AUX & Y=AUX	A: Obligatory B: Optional	Present: [ $\pm$ TENSE] MOD-AUX-V: optional 3-LEFT

As mentioned above, the dialect groups given are supersets and do not distinguish micro-parametric variation within certain regional groups (which clearly does exist, as illustrated in the clustering of certain orders in various sub-regions of The Netherlands as well as throughout the German speaking areas; see Barbiers et al. 2008 for the former and Schmid and Vogel 2004, Wurmbrand 2004b, and Dubenion-Smith 2010 for the latter). The rules in Table 3 thus yield the exhaustive set of word orders for each superregional dialect group and do not distinguish between the frequencies indicated in Table 2. As suggested above, the high numbers of possible orders in certain dialect groups could arise as a result of bilingualism among different dialects, thus yielding the sum of options as given in Table 2. The way the rules have been stated is intended to reflect this, as much as possible given the empirical distribution. For instance, the Standard German inversion context requires an auxiliary-modal sister pair. Although alternative

formulations would be possible for the German, Austrian, and Swiss dialects, the same rule is used to show that these dialect groups represent a superset of the more restricted variety, allowing inversion in the same context, but also in other syntactic configurations. Similarly, one of the Standard Dutch inversion rules affects participles, but in the set of Dutch dialects, this rule is generalized to all non-finite contexts, again making the restricted Standard Dutch rule a subset of the superregional rule.

A final important note about the application of the rules in Table 3 is that the optional rules are to be understood as ‘all-or-nothing’ rules. For instance, rule B in the Dutch dialects, if it applies (note that it is given as optional) must apply to all non-finite verbs in a cluster. Thus, in a configuration V3.NON-FIN—V2.NON-FIN—V1.FIN, rule B would yield the output order 1-2-3, but crucially not 2-3-1 nor 1-3-2. In this dialect, the 1-3-2 order could still be derived in this context, however, but only via the application of rule A (since there is only one finite verb, rule A only applies to V1). In what follows, I summarize how the rules derive the possible and exclude the impossible orders in the different languages and dialects.

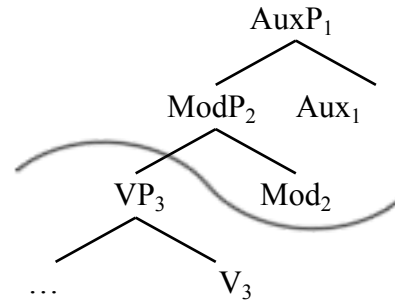
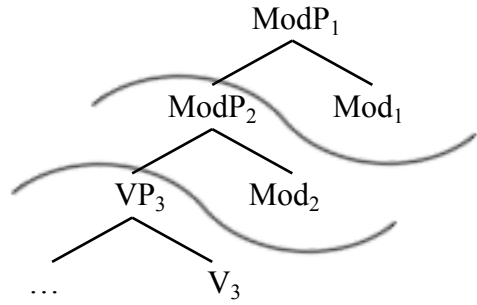
The examples and diagrams in (16) illustrate the inversion rules in Afrikaans (all data are from Robbers 1997). The Afrikaans inversion rule is stated such that it applies obligatorily to modals and their complements.<sup>9</sup> Thus, in (16b, c, d, e, f), inversion applies as indicated. Furthermore, it is assumed that the future auxiliary is a modal in Afrikaans, which has been argued for independently by for instance, Erb (2001). Thus, future constructions behave essentially like double modal constructions. Finally, passive participles allow the special leftward movement (cf. (16f)), in which case inversion between the modal and its complement does not include the lowest VP—resulting in the 3-1-2 order. Assuming that no other inversions are allowed, the rule specifications given in Table 3 derive exactly the structures attested in Afrikaans and only those structures.

- (16) a.    dat     Jan     Marie   gesien<sub>2</sub>   het<sub>1</sub>   /   \*het<sub>1</sub>   gesien<sub>2</sub>   Afrikaans  
           that     Jan     Marie   seen<sub>2</sub>   has<sub>1</sub>   /   \*has<sub>1</sub>   seen<sub>2</sub>                    2-1; \*1-2
- b.    dat     Jan     môre     kan<sub>1</sub>     werk<sub>2</sub>   /   \*werk<sub>2</sub>   kan<sub>1</sub>                    1-2; \*2-1  
           that     Jan     tomorrow   can<sub>1</sub>     work<sub>2</sub>   /   \*work<sub>2</sub>   can<sub>1</sub>
- a'.    AUX-PART: no inversion—2-1    b'.    MOD-INF: inversion—1-2
- 
- c.    dat     Jan     môre     sal<sub>1</sub>     kan<sub>2</sub>     werk<sub>3</sub>                    1-2-3; \*all others  
           that     Jan     tomorrow   will<sub>1</sub>     can<sub>2</sub>     work<sub>3</sub>                    will = MOD
- d.    dat     Jan     kon<sub>2</sub>     werk<sub>3</sub>     het<sub>1</sub>                    2-3-1; \*all others  
           that     Jan     could<sub>2</sub>     work<sub>3</sub>     has<sub>1</sub>

<sup>9</sup> The formulation of the inversion rule as in Table 3 (in particular the assumption that the rule is sensitive to modals rather than to infinitives) receives further support from the fact that modals also invert with non-verbal complements in Afrikaans (see Robbers 1997:167-171).

c'. MOD-MOD-V: 2 inversions—1-2-3

d'. AUX-MOD-V: 1 inversion—2-3-1



e. dat Jan Marie kan<sub>1</sub> gesien<sub>3</sub> het<sub>2</sub> 1-3-2; \*all others

that Jan Marie can<sub>1</sub> seen<sub>3</sub> have<sub>2</sub>

e/f. dat Jan toe kon<sub>1</sub> gevang<sub>3</sub> word<sub>2</sub> 1-3-2

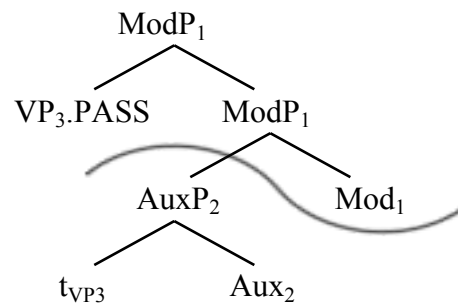
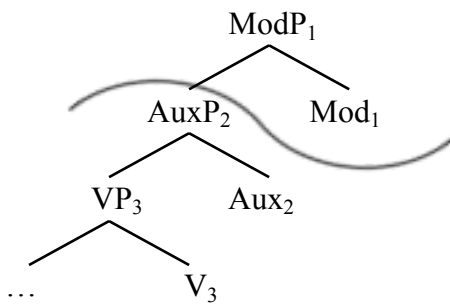
that Jan then could<sub>1</sub> caught<sub>3</sub> be<sub>2</sub>

dat Jan toe gevang<sub>3</sub> kon<sub>1</sub> word<sub>2</sub> 3-1-2

that Jan then caught<sub>3</sub> could<sub>1</sub> be<sub>2</sub>

e'. MOD-AUX-V: 1 inversion—1-3-2

f'. MOD-AUX-V: 1 inversion+3-LEFT—3-1-2



The derivations of the Standard Dutch inversion patterns are summarized in (17) (the data from Rutten 1991, Zwart 1996, Robbers 1997).

(17) a. dat Jan het boek gelezen<sub>2</sub> heeft<sub>1</sub> / heeft<sub>1</sub> gelezen<sub>2</sub> Dutch  
that Jan the book read.PART<sub>2</sub> has<sub>1</sub> / has<sub>1</sub> read.PART<sub>2</sub> 2-1; 1-2

B: optional inversion of PART

b. dat Jan het boek lezen<sub>2</sub> kan<sub>1</sub> / kan<sub>1</sub> lezen<sub>2</sub> 2-1; 1-2  
that Jan the book read.INF<sub>2</sub> can<sub>1</sub> / can<sub>1</sub> read.INF<sub>2</sub>

A: inversion of INF (usually obligatory, but optional in finite 2-verb clusters)

c. dat Jan morgen zal/moet<sub>1</sub> kunnen<sub>2</sub> werken<sub>3</sub> 1-2-3; \*all others  
that Jan tomorrow will/must<sub>1</sub> can.INF<sub>2</sub> work.INF<sub>3</sub>

A: obligatory inversions of INF

d. dat Jan het boek heeft<sub>1</sub> kunnen<sub>2</sub> lezen<sub>3</sub> 1-2-3; \*all others  
that Jan the book has<sub>1</sub> can.INF<sub>2</sub> read.INF<sub>3</sub>

A: obligatory inversions of INF (IPP=INF)

e.	dat	Jan	Marie	kan <sub>1</sub>	hebben <sub>2</sub>	gezien <sub>3</sub>	1-2-3
	that	Jan	Marie	can <sub>1</sub>	have.INF <sub>2</sub>	seen.PART <sub>3</sub>	
	dat	Jan	Marie	kan <sub>1</sub>	gezien	hebben	1-3-2
	that	Jan	Marie	can <sub>1</sub>	seen.PART <sub>3</sub>	have.INF <sub>2</sub>	
	dat	Jan	Marie	gezien	can <sub>1</sub>	hebben	3-1-2
	that	Jan	Marie	seen.PART <sub>3</sub>	can <sub>1</sub>	have.INF <sub>2</sub>	*all others

A: obligatory inversion of INF yields 1»2P (\*3-2-1)

B: optional inversion of PART or optional 3-LEFT yields {3}-1-{3}-2-{3}

The orders possible across Dutch dialects are derived as in (18) (the data are based on Barbiers 2005 and Barbiers et al. 2008; as noted under Table 2, the distribution of the 3-1-2 order in IPP constructions is somewhat unclear; possibly this order should be omitted).

(18) a.	die	het verhaal	verteld <sub>2</sub>	heeft <sub>1</sub>	/ heeft <sub>1</sub>	verteld <sub>2</sub>	2-1; 1-2
	who	the story	told.PART <sub>2</sub>	has.FIN <sub>1</sub>	/ has.FIN <sub>1</sub>	told.PART <sub>2</sub>	

b.	dat	jij	het ook	niet zien <sub>2</sub>	mag <sub>1</sub>	/ mag <sub>1</sub>	zien <sub>2</sub>	2-1; 1-2
	that	you	it also	not see.INF <sub>2</sub>	may.FIN <sub>1</sub>	/ may.FIN <sub>1</sub>	see.INF <sub>2</sub>	

A or B: optional inversion of V.FIN (A) or V.NON-FIN (B)

c.	dat	iedereen	moet <sub>1</sub>	kunnen <sub>2</sub>	zwemmen <sub>3</sub>	1-2-3
	that	everyone	must.FIN <sub>1</sub>	can.INF <sub>2</sub>	swim.INF <sub>3</sub>	
	dat	iedereen	zwemmen <sub>3</sub>	moet <sub>1</sub>	kunnen <sub>2</sub>	3-1-2
	that	everyone	swim.INF <sub>3</sub>	must.FIN <sub>1</sub>	can.INF <sub>2</sub>	
	dat	iedereen	zwemmen <sub>3</sub>	kunnen <sub>2</sub>	moet <sub>1</sub>	3-2-1
	that	everyone	swim.INF <sub>3</sub>	can.INF <sub>2</sub>	must.FIN <sub>1</sub>	
	dat	iedereen	moet <sub>1</sub>	zwemmen <sub>3</sub>	kunnen <sub>2</sub>	1-3-2
	that	everyone	must.FIN <sub>1</sub>	swim.INF <sub>3</sub>	can.INF <sub>2</sub>	*all others

A, B optional; no inversion yields 3-2-1

A: inversion of V.FIN with sister yields 1-3-2; plus 3-LEFT yields 3-1-2

B: inversion of all V.NON-FIN yields 1-2-3 (inversion of only lower INF—2-3-1—is blocked by the ‘all-or-nothing’ character of the inversion rules)

d.	wie	zij	had <sub>1</sub>	kunnen <sub>2</sub>	roepen <sub>3</sub>	1-2-3
	who	she	had.FIN <sub>1</sub>	can.— <sub>2</sub>	call.INF <sub>3</sub>	
	wie	zij	roepen <sub>3</sub>	kunnen <sub>2</sub>	had <sub>1</sub>	3-2-1
	who	she	call.INF <sub>3</sub>	can.— <sub>2</sub>	had.FIN <sub>1</sub>	
	wie	zij	kunnen <sub>2</sub>	roepen <sub>3</sub>	had <sub>1</sub>	2-3-1
	who	she	can.— <sub>2</sub>	call.INF <sub>3</sub>	had.FIN <sub>1</sub>	
	wie	zij	had <sub>1</sub>	roepen <sub>3</sub>	kunnen <sub>2</sub>	1-3-2
	who	she	had.FIN <sub>1</sub>	call.INF <sub>3</sub>	can.— <sub>2</sub>	
[	dat	hij	zwemmen <sub>3</sub>	is <sub>1</sub>	gaan <sub>2</sub>	3-1-2]
	that	he	swim.INF <sub>3</sub>	is.FIN <sub>1</sub>	go.— <sub>2</sub>	*all others

A, B optional; no inversion yields 3-2-1

A: inversion of V.FIN with sister yields 1-3-2; plus 3-LEFT yields 3-1-2

B: IPP is assumed to not be specified for [ $\pm$ FIN]; it is thus invisible for rule B; inversion of V3 [INF] yields 2-3-1

A and B: inversion of V.FIN (A) and V.INF (B) yields 1-2-3

e.	dat	hij	...	de wagen	gemaakt <sub>3</sub>	moet <sub>1</sub>	hebben <sub>2</sub>	3-1-2
	that	he	...	the car	made.PART <sub>3</sub>	must.FIN <sub>1</sub>	have.INF <sub>2</sub>	
	dat	hij	...	de wagen	moet <sub>1</sub>	gemaakt <sub>3</sub>	hebben <sub>2</sub>	1-3-2
	that	he	...	the car	must.FIN <sub>1</sub>	made.PART <sub>3</sub>	have.INF <sub>2</sub>	
	dat	hij	...	de wagen	moet <sub>1</sub>	hebben <sub>2</sub>	gemaakt <sub>3</sub>	1-2-3
	that	he	...	the car	must.FIN <sub>1</sub>	have.INF <sub>2</sub>	made.PART <sub>3</sub>	
	dat	hij	...	de wagen	gemaakt <sub>3</sub>	hebben <sub>2</sub>	moet <sub>1</sub>	3-2-1
	that	he	...	the car	made.PART <sub>3</sub>	have.INF <sub>2</sub>	must.FIN <sub>1</sub>	*all others

A, B optional; no inversion yields 3-2-1

A: inversion of V.FIN with sister yields 1-3-2; plus 3-LEFT yields 3-1-2

B: inversion of all V.NON-FIN yields 1-2-3 (inversion of only lower PART—2-3-1—is blocked by the ‘all-or-nothing’ character of the inversion rules)

Turning to German, the Standard German derivations are given in (19). Since the Standard German inversion rule only applies to Aux<sub>1</sub>-Mod<sub>2</sub> sister pairs, no inversion takes place in (19a,b,c,f,g). The only constructions that involve inversion are cases where an auxiliary selects a modal, as in (19e). Lastly, for future constructions as in (19d), which allow two word orders, it is assumed that the verb *werden* ‘will’ can be either classified as a modal or an auxiliary.

(19) a.	weil	er	das Buch	gekauft <sub>2</sub>	hat <sub>1</sub>	/ *hat <sub>1</sub>	gekauft <sub>2</sub>	Standard German
	since	he	the book	bought <sub>2</sub>	has <sub>1</sub>	/ *has <sub>1</sub>	bought <sub>2</sub>	2-1; *1-2
b.	weil	er	das Buch	kaufen <sub>2</sub>	wird/muss <sub>1</sub>	/ *wird/muss <sub>1</sub>	kaufen <sub>2</sub>	2-1; *1-2
	since	he	the book	buy <sub>2</sub>	will/must <sub>1</sub>	/ *will/must <sub>1</sub>	buy <sub>2</sub>	
c.	weil	er	es	kaufen <sub>3</sub>	können <sub>2</sub>	will <sub>1</sub>		3-2-1; *all others
	since	he	it	buy <sub>3</sub>	can <sub>2</sub>	wants <sub>1</sub>		
d.	weil	er	es	kaufen <sub>3</sub>	können <sub>2</sub>	wird <sub>1</sub>		3-2-1
	since	he	it	buy <sub>3</sub>	can.MOD <sub>2</sub>	will.MOD <sub>1</sub>		
	weil	er	es	wird <sub>1</sub>	kaufen <sub>3</sub>	können <sub>2</sub>		1-3-2
	since	he	it	will.AUX	buy <sub>3</sub>	can.MOD <sub>2</sub>		*all others
e.	weil	er	es	hat <sub>1</sub>	kaufen <sub>3</sub>	können <sub>2</sub>		1-3-2
	since	he	it	has.AUX	buy <sub>3</sub>	can.MOD <sub>2</sub>		*all others
f.	weil	er	es	gekauft <sub>3</sub>	haben <sub>2</sub>	muss <sub>1</sub>		3-2-1
	since	he	it	bought <sub>3</sub>	have.AUX <sub>2</sub>	must.MOD <sub>1</sub>		*all others
g.	weil	es	gekauft <sub>3</sub>	worden <sub>2</sub>	ist <sub>1</sub>			3-2-1
	since	it	bought <sub>3</sub>	been.AUX <sub>2</sub>	is.AUX <sub>1</sub>			*all others

The set of possible word orders in verb clusters with three verbal elements and their derivations in German and Austrian dialects are summarized in (20).<sup>10</sup>

- (20) a. weil er es kaufen<sub>3</sub> können<sub>2</sub> will<sub>1</sub> 3-2-1  
 since he it buy<sub>3</sub> can<sub>2</sub> wants.MOD<sub>1</sub>  
 weil er es will kaufen können 1-3-2  
 since he it wants.MOD<sub>1</sub> buy<sub>3</sub> can<sub>2</sub> \*all others
- B: optional inversion of MOD.FIN
- b. weil er es kaufen<sub>3</sub> können<sub>2</sub> wird<sub>1</sub> 3-2-1  
 since he it buy<sub>3</sub> can.MOD<sub>2</sub> will.MOD/AUX<sub>1</sub>  
 weil er es wird<sub>1</sub> kaufen<sub>3</sub> können<sub>2</sub> 1-3-2  
 since he it will.AUX/MOD<sub>1</sub> buy<sub>3</sub> can.MOD<sub>2</sub>  
 weil er es kaufen<sub>3</sub> wird<sub>1</sub> können<sub>2</sub> 3-1-2  
 since he it buy<sub>3</sub> will.AUX/MOD<sub>1</sub> can.MOD<sub>2</sub> \*all others
- A, B optional; no inversion yields 3-2-1  
 A or B: inversion of *will*=AUX (A) or *will*=MOD.FIN (B) yields 1-3-2, plus 3-LEFT yields 3-1-2
- c. weil er es hat<sub>1</sub> kaufen<sub>3</sub> können<sub>2</sub> 1-3-2  
 since he it has.AUX<sub>1</sub> buy<sub>3</sub> can.MOD<sub>2</sub>  
 weil er es kaufen<sub>3</sub> können<sub>2</sub> hat<sub>1</sub> 3-2-1  
 since he it buy<sub>3</sub> can.MOD<sub>2</sub> has.AUX<sub>1</sub>  
 weil er es kaufen<sub>3</sub> hat<sub>1</sub> können<sub>2</sub> 3-1-2  
 since he it buy<sub>3</sub> has.AUX<sub>1</sub> can.MOD<sub>2</sub> \*all others
- A, B optional; no inversion yields 3-2-1  
 A: inversion yields 1-3-2, plus 3-LEFT yields 3-1-2
- d. weil er es gekauft<sub>3</sub> haben<sub>2</sub> muss<sub>1</sub> 3-2-1  
 since he it bought<sub>3</sub> have<sub>2</sub> must.MOD<sub>1</sub>  
 weil er es muss<sub>1</sub> gekauft<sub>3</sub> haben<sub>2</sub> 1-3-2  
 since he it must.MOD<sub>1</sub> bought<sub>3</sub> have<sub>2</sub>  
 ob das in Betrieb genommen<sub>3</sub> könnte<sub>1</sub> werden<sub>2</sub> 3-1-2  
 whether that in service taken<sub>3</sub> could<sub>1</sub> be<sub>2</sub> \*all others  
 [Dubenion-Smith 2010: 114, (18b)]
- A, B optional; no inversion yields 3-2-1  
 B: inversion yields 1-3-2, plus 3-LEFT yields 3-1-2
- e. weil es gekauft<sub>3</sub> worden<sub>2</sub> ist<sub>1</sub> 3-2-1  
 since it bought<sub>3</sub> been.AUX<sub>2</sub> is.AUX<sub>1</sub>

<sup>10</sup> The rules as given in Table 3 predict that in simple MOD.FIN-V.INF constructions inversion, the 1-2 order, should be possible (see the references under Table 1). If further empirical investigations show that this order should be excluded, the rules would need to be modified to block inversion in two-verb clusters.

weil	es	gekauft <sub>3</sub>	ist <sub>1</sub>	worden <sub>2</sub>		3-1-2
since	it	bought <sub>3</sub>	is.AUX <sub>1</sub>	been.AUX <sub>2</sub>		*all others?

No inversion (neither A nor B apply) yields 3-2-1

A (only specified for X): inversion yields 1-3-2, plus 3-LEFT yields 3-1-2<sup>11</sup>

The data and inversion patterns of Swiss German are given in (21) (the rules yield Swiss-1 in two-verb clusters; see the comments under Table 2; the data are from Schönenberger 1995, Hsiao 1999, M. Schönenberger p.c., E. Haeberli p.c.;).

- (21) a.    das    t chatz    fisch    gässe<sub>2</sub>    hät<sub>1</sub>    /    \*hät<sub>1</sub>    gässe<sub>1</sub>    2-1; \*1-2  
           that    the cat    fish    eaten<sub>2</sub>    has.AUX<sub>1</sub>    /    \*has.AUX<sub>1</sub>    eaten<sub>2</sub>

No inversion—neither A nor B apply

- b.    das    t chatz    fisch    ässe<sub>2</sub>    mues<sub>1</sub>    /    mues<sub>1</sub>    ässe<sub>2</sub>    2-1; 1-2  
           that    the cat    fish    eat.INF<sub>2</sub>    must.MOD<sub>1</sub>    /    must.MOD<sub>1</sub>    eat.INF<sub>2</sub>

B or C: optional inversion

- c.    das    er    ...    wil<sub>1</sub>    chöne<sub>2</sub>    vorsinge<sub>3</sub>    1-2-3  
           that    he    ...    wants.MOD<sub>1</sub>    can.INF<sub>2</sub>    sing.INF<sub>3</sub>  
       das    er    ...    wil<sub>1</sub>    vorsinge<sub>3</sub>    chöne<sub>2</sub>    1-3-2  
           that    he    ...    wants.MOD<sub>1</sub>    sing.INF<sub>3</sub>    can.INF<sub>2</sub>  
       das    er    ...    vorsinge<sub>3</sub>    wil<sub>1</sub>    chöne<sub>2</sub>    3-1-2  
           that    he    ...    sing.INF<sub>3</sub>    wants.MOD<sub>1</sub>    can.INF<sub>2</sub>  
       das    er    ...    vorsinge<sub>3</sub>    chöne<sub>2</sub>    wil.MOD<sub>1</sub>    3-2-1  
           that    he    ...    sing.INF<sub>3</sub>    can.INF<sub>2</sub>    wants<sub>1</sub>

A does not apply; B or C optional; no inversion yields 3-2-1

B: inversion of MOD.FIN with sister yields 1-3-2; plus 3-LEFT yields 3-1-2

C: inversion of all V.INF yields 1-2-3 (inversion of only lower INF—2-3-1—is blocked by the ‘all-or-nothing’ character of the inversion rules)

- d.    das    de Jonas    hät<sub>1</sub>    müese<sub>2</sub>    schwimme<sub>3</sub>    1-2-3  
           that    the Jonah    has.AUX<sub>1</sub>    must.INF<sub>2</sub>    swim.INF<sub>3</sub>  
       das    de Jonas    hät<sub>1</sub>    schwimme<sub>3</sub>    müese<sub>2</sub>    1-3-2  
           that    the Jonah    has.AUX<sub>1</sub>    swim.INF<sub>3</sub>    must.INF<sub>2</sub>  
       das    de Jonas    schwimme<sub>3</sub>    hät<sub>1</sub>    müese<sub>2</sub>    3-1-2  
           that    the Jonah    swim.INF<sub>3</sub>    has.AUX<sub>1</sub>    must.INF<sub>2</sub>

A: obligatory inversion of AUX and MOD yields 1»2P (\*3-2-1)

C: optional inversion of INF or optional 3-LEFT yields {3}-1-{3}-2-{3}

<sup>11</sup> It is not clear at this point whether this construction also allows the 1-3-2 order. If indeed only the 3-2-1 and 3-1-2 orders are possible, the application of inversion rule A needs to be restricted to occur only in conjunction with the 3-LEFT rule. Since these double participle constructions are quite rare, we do not have the data to make conclusive claims about the availability of these orders.



- e. wil er si mues<sub>1</sub> gsee<sub>3</sub> ha<sub>2</sub> 1-3-2  
 since he her must.MOD<sub>1</sub> seen.PART<sub>3</sub> have.INF<sub>2</sub>  
 wil er si gsee<sub>3</sub> ha<sub>2</sub> mues<sub>1</sub> 3-2-1  
 since he her seen.PART<sub>3</sub> have.INF<sub>2</sub> must.MOD<sub>1</sub>  
 wil er si gsee<sub>3</sub> mues<sub>1</sub> ha<sub>2</sub> 3-1-2  
 since he her seen.PART<sub>3</sub> must.MOD<sub>1</sub> have.INF<sub>2</sub>
- A does not apply; B or C optional; no inversion yields 3-2-1  
 B or C: inversion of MOD.FIN and INF yields 1-3-2 plus 3-LEFT yields 3-1-2 (3»2 inversion does not apply since V3 is a PART)
- f. das de Hans gwäält<sub>3</sub> worde<sub>2</sub> isch<sub>1</sub> 3-2-1  
 that the John elected.PART<sub>3</sub> been.AUX<sub>2</sub> is.AUX<sub>1</sub>
- No inversion—neither A, B nor C apply

Lastly, (22) provides the data and summary of inversions for West Flemish (the data are from Haegeman 1994: 511/517, Haegeman 1995: 59, Haegeman 1998b, p.c.).

- (22) a. da Valère dienen boek gelezen<sub>2</sub> oat<sub>1</sub> / \*oa<sub>1</sub> gelezen 2-1; \*1-2  
 that Valère that book read.PART<sub>2</sub> had<sub>1</sub> / \*had<sub>1</sub> read.PART<sub>2</sub>
- No inversion—neither A nor B apply
- b. da Valère dienen boek wilt<sub>1</sub> kuopen<sub>2</sub> / \*kuopen<sub>2</sub> wilt<sub>1</sub> 1-2; \*2-1  
 that Valère this book wants.MOD<sub>1</sub> buy.INF<sub>2</sub> / \*buy.INF<sub>2</sub> wants.MOD<sub>1</sub>
- A: obligatory inversion between MOD (operator) and INF
- c. dan ze dienen boek kosten<sub>1</sub> willen<sub>2</sub> kupen<sub>3</sub> 1-2-3  
 that they this book could.MOD<sub>1</sub> want.MOD<sub>2</sub> buy.INF<sub>3</sub>
- A: obligatory inversion between both MOD (operator) and INF
- d. da Valère willen<sub>2</sub> Marie dienen boek geven<sub>3</sub> eet<sub>1</sub> 2-3-1  
 that Valère want.MOD<sub>2</sub> Marie that book give.INF<sub>3</sub> has.-TNS<sub>1</sub>
- da Valère ee<sub>1</sub> willen<sub>2</sub> Marie dienen boek geven<sub>3</sub> 1-2-3  
 that Valère has.+TNS<sub>1</sub> want.MOD<sub>2</sub> Marie that book give.INF<sub>3</sub>
- da Valère oa<sub>1</sub> willen<sub>2</sub> Marie dienen boek geven<sub>3</sub> 1-2-3  
 that Valère had.+TNS<sub>1</sub> want.MOD<sub>2</sub> Marie that book give.INF<sub>3</sub>
- da Valère nooit en-ee<sub>1</sub> willen<sub>2</sub> Marie dienen boek geven<sub>3</sub> 1-2-3  
 that Valère never NEG-has<sub>1</sub> want.MOD<sub>2</sub> Marie that book give.INF<sub>3</sub>
- ?\*da Valère willen<sub>2</sub> Marie dienen boek geven<sub>3</sub> oat<sub>1</sub> ?\*2-3-1  
 that Valère want.MOD<sub>2</sub> Marie that book give.INF<sub>3</sub> had.+TNS<sub>1</sub>
- ?\*da Valère nooit willen<sub>2</sub> Marie dienen boek geven<sub>3</sub> en-eet<sub>1</sub> ?\*2-3-1  
 that Valère never want.MOD<sub>2</sub> Marie that book give.INF<sub>3</sub> NEG-has<sub>1</sub>
- A: obligatory inversion of MOD<sub>2</sub> and INF<sub>3</sub> yields 2-3-1 (\*3»2); inversion of AUX and

e.	dat	Jan	...	moe <sub>1</sub>	gezien <sub>3</sub>	een <sub>2</sub>	1-3-2
	that	Jan	...	must.MOD <sub>1</sub>	seen.PART <sub>3</sub>	have.INF <sub>2</sub>	
	da	Jan	...	gezien <sub>3</sub>	moet <sub>1</sub>	een <sub>2</sub>	3-1-2
	that	Jan	...	seen.PART <sub>3</sub>	must.MOD <sub>1</sub>	have.INF <sub>2</sub>	
	A: obligatory inversion of MOD and INF yields 1-3-2, plus 3-LEFT yields 3-1-2						
f.	da	Jan	gekozen <sub>3</sub>		geworden <sub>2</sub>	is <sub>1</sub>	3-2-1
	that	Jan	elected.PART <sub>3</sub>		been.AUX.PART <sub>2</sub>	is.AUX <sub>1</sub>	
	dat	Jan	is		gekozen	geworden	1-3-2
	that	Jan	is.AUX <sub>1</sub>		elected.PART <sub>3</sub>	been.AUX.PART <sub>2</sub>	
	B: optional inversion of AUX and AUX yields 3-2-1 (no inversion), 1-3-2 (inversion)						

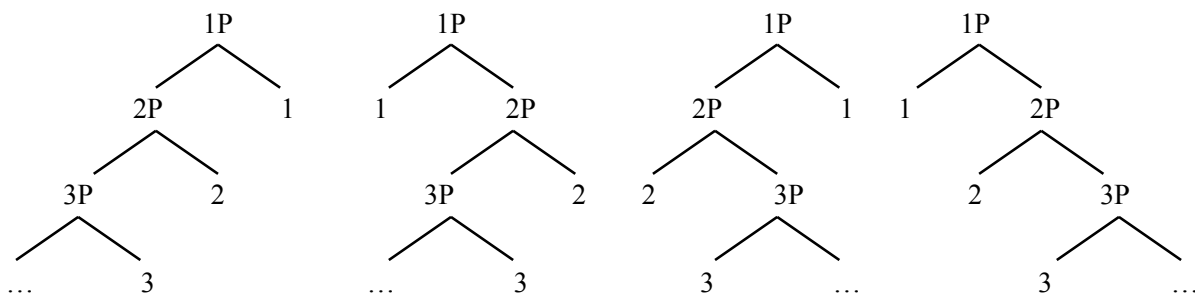
### 3. DERIVING VERB CLUSTERS

### 3.1 Movement vs. no movement

26

fixed within a single language, in either of these systems one order in a verb cluster (the 1-2-3 or the 3-2-1 order) would be the basic order, and all other orders would be derived by movement or inversion (as for instance laid out in section 2.5). There is, however, another option. If, within a language, directionality is either free or determined for each projection as part of the mapping from syntax to PF, more than one word order could be generated without inversion. Wurmbrand (2004a) and Abels (2011) propose such ‘free directionality’ models. In these accounts, certain verbal projections are generated or linearized in a head-initial, others in a head-final order, based on language and construction specific instructions. As shown in Table 2, of the six possible word orders in verb clusters with three verbal elements, five orders are robustly attested: 1-2-3, 1-3-2, 3-1-2, 3-2-1, and 2-3-1. Free directionality systems then generate four of the six logical word orders without any syntactic movement or inversion, as in (23) (for the 3-1-2 and 2-1-3 orders see below).<sup>12</sup>

(23) Free head-final or head-initial generation or linearization



According to Wurmbrand (2004a), verb clusters are generated in their hierarchical (=semantic) order in the syntax (only c-command, not precedence, is relevant in syntax). During the mapping from syntax to PF, sister nodes are linearized, at which point information about language-specific ordering restrictions comes into play (for other accounts involving post-syntactic reordering see von Stechow and Sternefeld 1988, den Besten and Broekhuis 1992, Broekhuis 1993). The main arguments for verb cluster re-ordering taking place post-syntactically rather than via syntactic movement are that reordering has no semantic effect (all constructions feed as 1»2»3 orders into the semantics) but that the IPP effect feeds verb cluster reordering in certain cases. It is hypothesized that the IPP effect, which changes a participle into an infinitive, is a post-syntactic operation. Thus, if IPP constructions are linearized like infinitives (rather than like participles), verb cluster linearization has to follow the IPP conversion, situating it in the post-syntactic component rather than in syntax.

Assuming that linearization only targets sister nodes, as shown in (23), the 3-1-2 and 2-1-3 orders cannot be generated by simple post-syntactic linearization, but must involve an additional process, such as syntactic movement. This has been a standard approach for configurations such as (24a), which on the surface involve a 2-1-3 order, however, on closer inspection, differ from verb clusters in several ways. As argued in Wurmbrand (2001), for instance, in German this con-

<sup>12</sup> Biberauer et al. (2014) propose the universal *Final-over-Final Constraint* according to which it is impossible to merge a head-initial projection under a head-final projection. In this approach, the configuration given for the 2-3-1 order in (23) could thus not be (base-)generated. One way to make these orders compatible with the *Final-over-Final Constraint* is provided in Salzmann (2013a, b). It is suggested there that these orders involve a cliticization structure [V2=V3], which would allow treating V2+V3 as a complex head and consequently 2-3-1 orders as simple two-verb clusters. It remains to be seen whether this approach can be extended to all attested 2-3-1 orders.

struction is only possible when the middle verb is a lexical restructuring verb (such as *try*, *dare*, *forget*) but not when it is a modal, auxiliary, or causative verb (see (24b)). These extraposition configurations (often also referred to as the *Third construction*—see section 3.2.1.2) always involve a *zu* ‘to’ infinitive and in German do not allow IPP effects (though IPP is possible with *proberen* ‘try’ in Dutch).

- (24) a. weil Hans den Ball versucht<sub>2</sub> hat<sub>1</sub> zu fangen<sub>3</sub> German  
 since John the ball tried<sub>2</sub> has<sub>1</sub> to catch<sub>3</sub> extraposition: 2-1-3  
 ‘since John tried to catch the ball’
- b. \*weil Hans den Ball wollen<sub>2</sub> hat<sub>1</sub> fangen<sub>3</sub> verb cluster: \*2-1-3  
 since John the ball want<sub>2</sub> has<sub>1</sub> catch<sub>3</sub>  
 ‘since John wanted to catch the ball’

In Wurmbrand (2012a, b) the claim that the 3-1-2 order involves syntactic movement is further substantiated for German, based on a special form of a parasitic participle construction illustrated in (25a).<sup>13</sup> In these constructions the main verb *verhindern* ‘prevent’ is realized as a participle instead of an infinitive, despite the fact that this verb is selected by *lassen* ‘let’ and not the auxiliary, which otherwise requires its complement to be realized as an infinitive (in a sense parasitic participles are the opposite of the IPP effect, which is why this phenomenon is also referred to as *Participium Pro Infinitivo* ‘participle for infinitive’ [PPI] effect). The morphological PPI change is only possible in the 3-1-2 order and not, as shown in (25b) in the 1-3-2 order, which is a possible verb cluster order in German when no parasitic morphology is used (cf. (25c). In Wurmbrand (2012a, b) it is argued that this form of parasitic participle construction is only possible when the participle c-commands the auxiliary, thus providing indirect evidence for syntactic movement in 3-1-2 constructions.

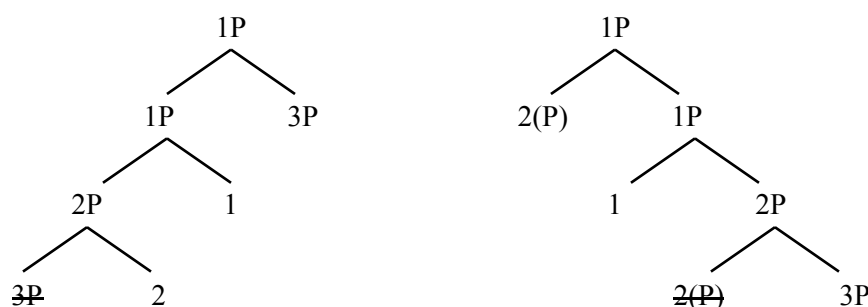
- (25) a. ohne es verhindert<sub>3</sub> haben<sub>1</sub> zu können<sub>2</sub> German  
 without it prevent.PART<sub>3</sub> have.INF<sub>1</sub> to can.IPP<sub>2</sub> parasitic participle: 3-1-2  
 ‘without having been able to prevent it’ [Vogel 2009]
- b. \*ohne es haben<sub>1</sub> verhindert<sub>3</sub> zu können<sub>2</sub> \*parasitic participle: 1-3-2  
 without it have.INF<sub>1</sub> prevent.PART<sub>3</sub> to can.IPP<sub>2</sub>
- c. ohne es haben<sub>1</sub> verhindern<sub>3</sub> zu können<sub>2</sub> 1-3-2  
 without it have.INF<sub>1</sub> prevent.INF<sub>3</sub> to can.IPP<sub>2</sub>

The account in Abels (2011) is built on the typological observation that the neutral orders of elements in a verb cluster mirror (almost exactly) the neutral word orders found among nominal elements in the noun phrase. Abels also considers verb clusters with four verbal elements, and notes (see also Bobaljik 2004) that ascending word orders (1-2-3...) show rigid ordering whereas descending word orders (3-2-1...) give rise to flexible ordering. The account proposed runs par-

<sup>13</sup> A similar ‘upwards’ selection of participles, i.e., a construction in which an auxiliary appears to license a non-selected participle to its left also exists in Frisian and certain Dutch dialect (see den Dikken and Hoekstra 1997 for Frisian and Bloemhoff 1979, Zwart 1995 for the Stellingwerf dialect; the German construction has also been discussed in Merkes 1895, Reis 1979, Meurers 2000, Vogel 2009, Haider 2011, Salzmann 2014). A ‘downward’ parasitic participle construction in which an auxiliary licenses a non-selected participle to its right also exists in Scandinavian (see Wiklund 2005, 2007). For a comparison of the two types of parasitic participle constructions see Wurmbrand (2010, 2012a, b).

allel to Abels and Neeleman's (2012) recasting of Cinque (2005) theory of Greenberg's Universal 20. While sister nodes can be (base) merged in either order, deriving the options in (23), movement that derives unmarked word orders (in both the noun phrase and a verb cluster) is restricted in that only leftward movement of XPs including the most deeply embedded noun (noun phrases) or verb (verb cluster) is possible. For verb clusters with three elements, the consequence is that 3-1-2 orders can be derived by movement (since VP3 is/contains the main verb and movement is to the left), but, crucially, the 2-1-3 order cannot be derived. To derive a 2-1-3 configuration, either movement would have to go to the right as in (26a), or strand the main verb (3P) as in (26b) (it is irrelevant whether such leftward movement would involve head movement of V2 or remnant movement of 2P—either option would leave the lowest VP behind, which is excluded by Abels and Neeleman's 2012 movement algorithm).

(26) a. Movement to the right    b. Movement without the main V (V3)



Both, Wurmbrand's and Abels' accounts are based (although with different force) on the assumption that the 2-1-3 order does not exist as a genuine verb cluster order but is only possible when a different mechanism (e.g., extraposition) applies. Similarly, Zwart (2007) suggests that surface occurrences of (apparent) 2-1-3 orders may be cases of simple two-verb clusters, which can occur in the 2-1 order, with the third verb being independent of the verb cluster. The general distribution of word orders supports this claim: the 2-1-3 order is extremely rare in the SANDII corpus (see Barbiers et al. 2008), which is the most extensive sampling of Dutch dialects, and it has not been robustly attested in the German dialects (although it is reported in the German study conducted by Schmid and Vogel 2004 for the Swiss German dialect of St. Gallen, the Low German dialect Rheiderländer Platt, an East Frisia dialect, and the South Tyrolean dialect of Meran, Italy).<sup>14</sup>

The claim that the 2-1-3 order does not exist as an unmarked verb cluster order has recently been challenged in Salzmann (2013a, b) where it is argued that in Swiss German perception verb, benefactive, and causative constructions, the 2-1-3 order is possible, and thus all of the logical word orders exist in verb clusters. Salzmann also proposes a post-syntactic account for verb cluster reordering, however, in contrast to the linearization approaches mentioned above, the word order alternations are not derived by inverting (syntactic) sister nodes but rather via a PF rebracketing operation plus inversion under adjacency. The PF procedures proposed by Salzmann are summarized in (27). The 1-2-3 order does not involve inversion, (27a); the 1-3-2 and 2-1-3 orders involve inversion after rebracketing as in (27b,c); the 3-2-1 order involves two steps of inversion, (27d); and lastly the 2-3-1 and 3-1-2 orders involve rebracketing without inversion

<sup>14</sup> Schmid and Vogel (2004) note an effect of focus on the availability of these orders (for focus sensitivity on different orders see also Sapp 2011, as summarized under Table 1).

in the first step, and then inversion after the second step of rebracketing as in (27e,f).

- (27) a. 123  
 b. 123  $\Rightarrow$  1[23]  $\Rightarrow$  132  
 c. 123  $\Rightarrow$  [12]3  $\Rightarrow$  213  
 d. 123  $\Rightarrow$  1[23]  $\Rightarrow$  1[32]  $\Rightarrow$  [1[32]]  $\Rightarrow$  321  
 e. 123  $\Rightarrow$  1[23]  $\Rightarrow$  —  $\Rightarrow$  [1[23]]  $\Rightarrow$  231  
 f. 123  $\Rightarrow$  [12]3  $\Rightarrow$  —  $\Rightarrow$  [[12]3]  $\Rightarrow$  312

One important claim Salzmann makes is that all six word orders exist as unmarked and true verb cluster orders, and in particular that the 2-1-3 order does not always involve a (non-clustering) extraposition configuration. Nevertheless the distribution of the 2-1-3 order is very restricted, and the fact that it is barely found with the core verb cluster elements (modals and auxiliaries) may still indicate a special status of this ordering. An important future task for the verb cluster research is hence to determine whether the 2-1-3 order exists as a genuine order of verb clusters and to develop further tests and diagnostics to distinguish verb clusters from extraposition configurations.

### 3.2 Syntactic approaches

As noted above, the majority of analyses of the verb cluster phenomenon are based on the idea that verb cluster reordering is a form of syntactic movement. The following sections address, among others, the following questions:

- i) Are verb cluster reordering operations best described as head movement, as phrasal movement, or both?
- ii) What is the relation between verb clusters and restructuring?
- iii) Do verb clusters shed light on the directionality debate (i.e., does verb cluster formation provide evidence for or against either a head-final or a head-initial approach)?
- iv) What kind of material can occur interspersed in a verb cluster (i.e., what is the distribution of so-called *verb projection raising* constructions)?

#### 3.2.1 Head movement vs. phrasal movement

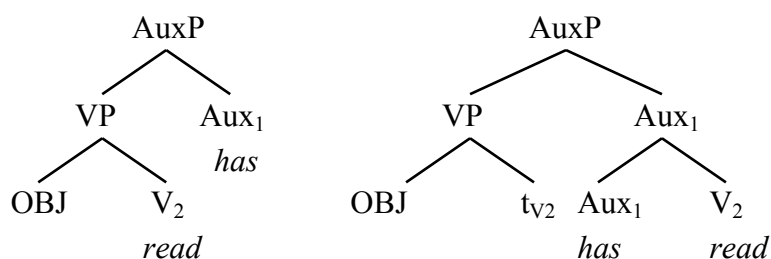
A basic theoretical question addressed directly or indirectly in many syntactic analyses of verb clusters is the question of whether verb cluster reordering involves head movement or phrasal movement. To illustrate the issue, consider again a simple two-verb cluster such as the Dutch auxiliary-participle construction in (28) (repeated from (4)).

- (28) a. dat Jan het boek gelezen heeft Dutch  
 that Jan the book read (2) has (1)  
 b. dat Jan het boek heeft gelezen  
 that Jan the book has (1) read (2)

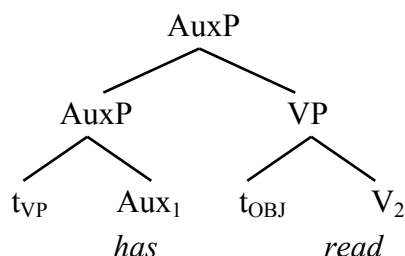
In models with rigid directionality, (28a) would count as the derived version under a head-initial base structure, whereas (28b) would be the derived version under a head-final base structure. Besides the issue of directionality, there are two options (for both the head-final and the head-initial structures) concerning the actual derivation of the non-basic examples in (28)—head movement

vs. phrasal movement. The two options are illustrated in (29) and (30) for both approaches. The diagrams in (29a) and (30a) display the basic orders and indicate what categories can undergo movement to derive the non-basic orders. As shown, movement can apply either to the head V2 or the whole VP2 (or a functional projection containing VP2 but excluding V1 if one were to assume a more refined structure for these examples). The diagrams in (29b) and (30b) show the derivations with head movement; the diagrams in (29c) and (30c) show the derivations with phrasal movement. Note that the object occurs to the left of the auxiliary in both examples in (28). To accommodate this fact, movement of the verb or the VP has to be preceded by leftward movement of the object (i.e., scrambling or object shift) in the head-initial structures in (30b,c) as well as the head-final structure involving VP-movement in (29c). Since reordering operations such as (29c) and (30b,c) apply to incomplete constituents, movement operations of this type have been dubbed *remnant movement*.

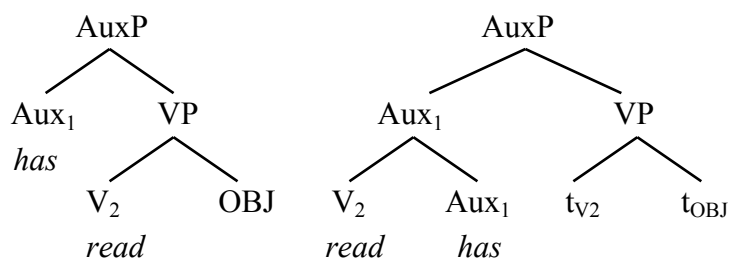
- (29) a. Head-final structure      b. Head movement (head-final)



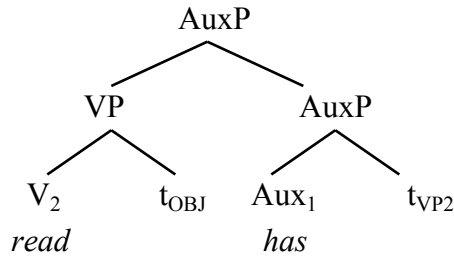
- c. XP-movement (head-final)



- (30) a. Head-initial structure      b. Head movement (head-initial)



c. XP-movement (head-initial)



In light of these options, the following questions arise: How can head movement vs. phrasal movement structures be distinguished? Can verb clusters be derived solely by head movement? Can verb clusters be derived solely by phrasal movement? The following subsections attempt to shed light on these questions.

### 3.2.1.1 Head movement derivations and issues

In this section, we will see that verb clusters cannot be seen as complex heads derived (exclusively) by head incorporation, but that (at least) certain constructions require a derivation involving phrasal movement or a non-incorporation analysis of head movement (see below). Importantly, this conclusion will be based entirely on the mechanics of the derivations and will hold for both head-initial as well as head-final base structures.

To illustrate the basic movement operations necessary to derive verb clusters by head incorporation, we will look at the distribution of IPP-constructions (repeated here in Table 4). As pointed out in section 2.4, of the six possible orders, five are attested in the IPP-construction.

Table 4. West Germanic IPP-construction

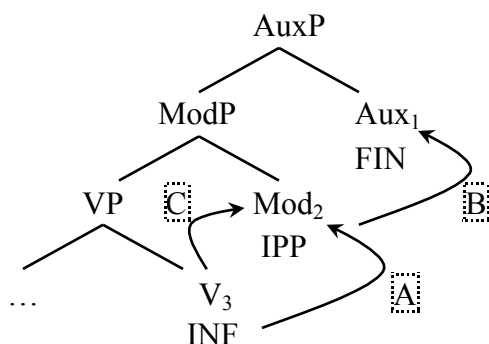
Language	3-2-1	3-1-2	1-2-3	1-3-2	2-3-1	2-1-3
Afrikaans	*	*	*	*	OK	*
Dutch Standard	*	*	OK	*	*	*
Dutch Dialects	OK	OK	OK	OK	OK	*
Frisian	OK	*	*	*	*	*
German Standard	*	*	*	OK	*	*
German, Austrian dialects	OK	OK	*	OK	*	*
Swiss dialects	*	OK	OK	OK	*	*
West Flemish	*	*	OK	*	OK	*

Let us start with the head-final structure. To derive the five orders attested in West Germanic by head incorporation (again, the purpose of this section is merely to mechanically derive the different word orders, conditions attached to different orders will be ignored for this basic illustration), three movement operations are necessary, labeled A, B, C in the diagram in (31a): movement of the lowest verb to the right of the higher verb (A), movement of the lowest verb to the left of the higher verb (C), and movement of the middle verb to the right of the highest verb (B). The reader can verify that the various combinations in (31b) yield exactly the five word orders attested (see for instance Evers 1975b, den Besten and Edmondson 1983 for accounts



similar to the one sketched).<sup>15</sup>

(31) a. Head-final structure



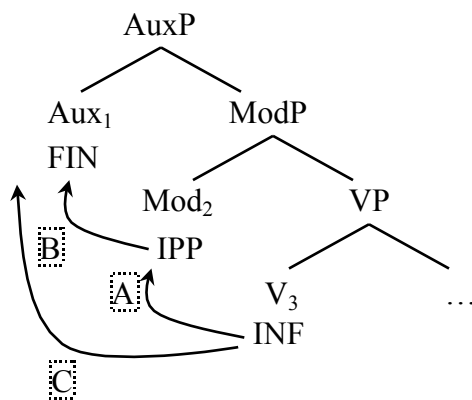
b. Possible derivations

No overt movement:	3-2-1
Only A:	2-3-1
Only B:	3-1-2
Only C (vacuous):	3-2-1
A and B:	1-2-3
C and B:	1-3-2

The sixth order—the unattested 2-1-3 order—can be excluded by the assumption of some version of Travis’s (1984) *Head Movement Constraint* (HMC): in order to derive 2-1-3, the lowest V3 would have to move and adjoin to V1, skipping over V2 and hence causing a HMC violation.

Turning to the head-initial approach, again three movement operations are necessary: movement of the lowest verb to the left of the higher verb (A), movement of the lowest verb to the left of the highest verb (C; potentially with a stop-over in V2), and movement of the middle verb to the left of the highest verb (B). These movement operations are illustrated in (32a) and the possible combinations are listed in (32b). Note that to derive the 2-3-1 order either movement C applies before movement B (i.e., movement C creates the complex head [3-1] to which V2 attaches afterwards), or movement B applies first (creating the complex 2-1) followed by movement C, however V3 does not left-adjoin to the complex 2-1 but tucks in between the two verbs (cf. Hsiao 1999 following Richards 1997 for an account along these lines).

(32) a. Head-initial structure



b. Possible derivations

No overt movement:	1-2-3
Only A:	1-3-2
Only C:	3-1-2
A and B:	3-2-1
C and B:	2-3-1
B and C:	3-2-1
B and C (tucking in):	2-3-1

Assuming a system that allows the derivations in (32), it is again straightforward to derive the five word orders found in West Germanic IPP-constructions. To exclude the unattested 2-1-3 or-

<sup>15</sup> In both (31) and (32), the orders corresponding the basic line-up of verbs (i.e., the 3-2-1 and 1-2-3 orders, respectively) are given as involving no movement. This claim is not uncontroversial and depends on one’s view of restructuring, which may require complex head formation even when the verbs already appear in their surface order (see section 3.2.1.2).

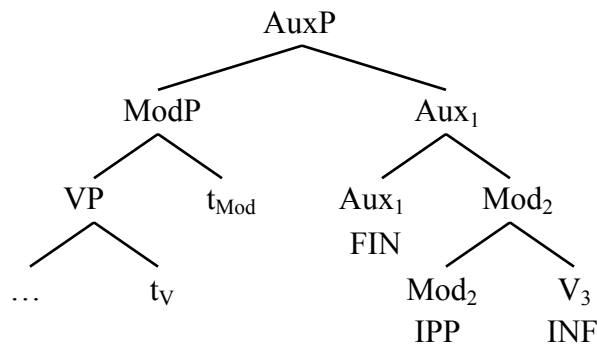
der under this approach the following options come to mind. Since movement B has to be possible in principle (i.e., to derive the orders in which V2 precedes V1), one has to ask why a derivation that *only* involves movement B is prohibited. A way to approach this questions would be to assume following Abels (2011) and Abels and Neeleman (2012) (see section 3.1) that movement from the middle of a verb cluster is excluded, or in other words, that movement of higher verbs is only possible when the lowest verb also moves (i.e., movement B would only be licensed in conjunction with movement A or C). Another option (M. den Dikken, p.c.) would be to see the impossibility of the 2-1-3 order as the result of a licensing failure of the lowest verb. Under the assumption that the lowest verb (V3) has to move (either overtly or covertly) to the next higher verb (V2), overt dislocation of V2 could be seen as an obstruction to the licensing of V3. That is, one could develop an account according to which i) adjunction to traces is prohibited (i.e., V3 cannot adjoin to the trace of V2) and ii) some locality condition excludes covert tucking in or covert adjunction of V3 to V2 when the latter is in a dislocated position (i.e., adjoined to V1).

The systems outlined in (31) and (32) naturally raise a number of questions depending on the type of syntactic theory one assumes and the principles and conditions that hold in one's favorite framework. For example, if a syntactic theory has no room for head movement (cf. Sportiche 1996, Koopman 1999, Koopman and Szabolcsi 2000), a proponent of such a system would obviously not consider any of the above derivations. If a system involves an assumption such as Kayne's (1994) *Linear Correspondence Axiom* (LCA), derivations involving a head-final structure and rightward movement (such as (31)) would not be an option in that framework. If (head) adjunction has to be uniformly in one direction, movement C in (31) would be excluded. If a system does not tolerate a violation of the HMC or involves a notion such as *Attract closest*, movement C as depicted in (32) might be problematic.

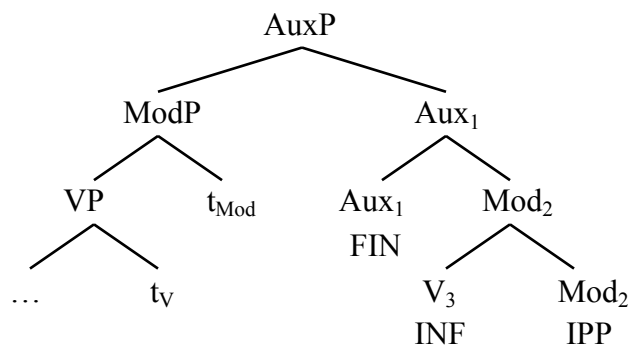
Questions of this sort are doubtlessly important to keep in mind when developing an actual analysis of verb cluster formation. However, since these types of questions are only valid in their particular frameworks, they do not seem to challenge analyses along the lines of (31) or (32) beyond the framework they are couched in, and hence they do not bear on the question of whether a (pure) head movement approach is in principle an option to account for the distribution of verb clusters. There is, however, an empirical phenomenon that goes beyond theory-internal considerations. The empirical evidence that, as we will see momentarily, argues against a (pure) head incorporation approach to verb cluster formation has become known as the *verb projection raising* phenomenon (see Haegeman and van Riemsdijk 1986 among many others). The argument against verb clusters as complexes of incorporated heads (whether head-initial or head-final) is straightforward. Under both approaches, it is predicted that in certain orders the verbs form a complex head, and hence these clusters should be impenetrable for XP-material. This prediction is not always borne out and hence head incorporation derivations as outlined above are problematic (but see below for a way to nevertheless make use of head movement in verb clusters).

To begin with the head-final structure in (31), (33) shows in detail the structures for the 1-2-3 order, (33a), the 1-3-2 order, (33b), and the 2-3-1 order, (33c). According to this approach, it should be impossible for XP-material to occur between V1 and V2, and V1 and V3 in (33a), between V1 and V3, and V3 and V2 in (33b), and between V2 and V3 in (33c).

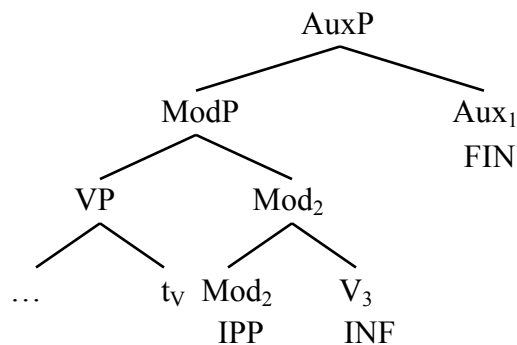
(33) a. 1-2-3 order in head-final approach



b. 1-3-2 order in head-final approach



c. 2-3-1 order in head-final approach



The following data from Swiss German, West Flemish, and German show that these predictions are not fulfilled in all but one case (see also section 3.2.3 for a more detailed discussion of verb projection raising). The examples in (34) make the point for the 1-2-3 order, the examples in (35) are problematic for the 1-3-2 order, and the example in (36) illustrates the problem for the 2-3-1 order. In more detail, in (34a) from Swiss German (SG), a definite object occurs between V1 and V2, and in the Swiss example in (34b) and the West Flemish (WF) example in (34c), various XPs occur between V2 and V3.<sup>16</sup> Since the verb clusters in (34) are split up by material which is clearly phrasal, it would be difficult to accommodate these examples under a structure like (33a).

<sup>16</sup> The Swiss data in these examples are from E. Haerberli (p.c.) and partly also from Hsiao 1999, 2000).

(34) Verb projection raising: 1-{XP}-2-{XP}-3

- a. ob si hett<sub>1</sub> **d Prüeffig** chöne<sub>2</sub> besto<sub>3</sub> SG  
 whether she had<sub>1</sub> the exam can<sub>2</sub> pass<sub>3</sub>  
 ‘[who knows] whether she would have been able to pass the exam’
- b. ob si hett<sub>1</sub> chöne<sub>2</sub> **d Prüeffig** besto<sub>3</sub> SG  
 whether she had<sub>1</sub> can<sub>2</sub> the exam pass<sub>3</sub>  
 ‘[who knows] whether she would have been able to pass the exam’
- c. da Valère oa<sub>1</sub> willen<sub>2</sub> **morgen** **no Gent** goan<sub>3</sub> WF  
 that Valère had<sub>1</sub> want.IPP<sub>2</sub> tomorrow to Gent go<sub>3</sub>  
 ‘that Valère had wanted to go to Gent tomorrow’ [Haegeman 1995:72]

The same problem arises in the 1-3-2 and 2-3-1 orders. In the German example in (35a) and the Swiss example in (35c), phrasal material appears between V1 and V3 in the 1-3-2 order (see also den Besten and Broekhuis 1992 for further examples), and in the West Flemish example in (36), both objects occur between V2 and V3 in the 2-3-1 order. The only case which indeed requires adjacency between elements of a cluster is illustrated in (35b) and (35d) (i.e., examples with elements between V3 and V2 in the 1-3-2 order). The generalization which will be discussed in more detail in section 3.2.1.5 for particles, is that XP-material is only allowed within a verb cluster (in certain languages and dialects) if the intervening XPs precede the lowest verb. Thus, in the 1-3-2 order as well as in the 3-1-2 order, phrasal material cannot occur between V3 and V2 or between V1 and V2, respectively (see also section 3.2.3).

(35) Verb projection raising: 1-XP-3-2

- a. dass er das Buch hätte<sub>1</sub> **genau** durchsehen<sub>3</sub> sollen<sub>2</sub> German  
 that he the book had<sub>1</sub> carefully through.look<sub>3</sub> shall<sub>2</sub>  
 ‘that he should have looked through the book carefully’ [Zwart 1996: 237, (18)]
- b. \*dass er das Buch hätte<sub>1</sub> durchsehen<sub>3</sub> **genau** sollen<sub>2</sub> German  
 that he the book had<sub>1</sub> through-look<sub>3</sub> carefully shall<sub>2</sub>  
 ‘that he should have looked through the book carefully’
- c. ob si hett<sub>1</sub> **d Prüeffig** besto<sub>3</sub> chöne<sub>2</sub> SG  
 whether she had<sub>1</sub> the exam pass<sub>3</sub> can<sub>2</sub>  
 ‘[who knows] whether she would have been able to pass the exam’
- d. \*ob si hett<sub>1</sub> besto<sub>3</sub> **d Prüeffig** chöne<sub>2</sub> SG  
 whether she had<sub>1</sub> pass<sub>3</sub> the exam can<sub>2</sub>  
 ‘[who knows] whether she would have been able to pass the exam’

(36) Verb projection raising: 2-XP-3-1

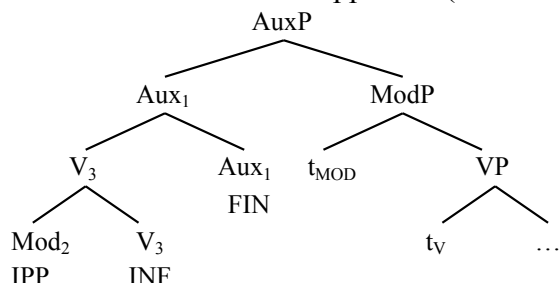
- da Valère willen<sub>2</sub> [ **Marie dienen boek** geven<sub>3</sub> ] eet<sub>1</sub> WF  
 that Valère want.IPP<sub>2</sub> [ Marie that book give<sub>3</sub> ] has<sub>1</sub>  
 ‘that Valère has wanted to give Marie that book’ [Haegeman 1998b: 260]

To conclude, it seems that the only way to maintain an incorporation analysis for the (grammatical) constructions in (34) through (36) would be to assume that the intervening material is head-incorporated in the verb clusters. This claim might be conceivable for certain small adverbs such

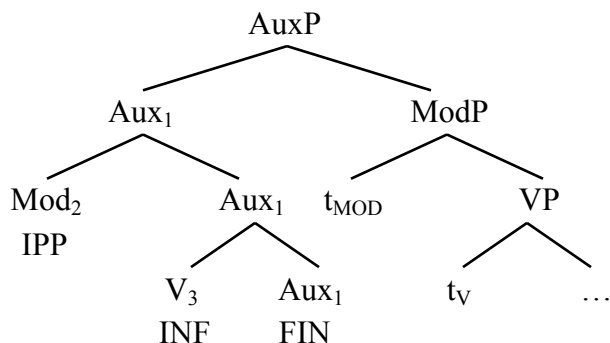
as the one in (35a), however, it appears highly implausible for the definite objects and prepositional phrases. It can thus be concluded that in light of verb projection raising constructions, a pure incorporation account is insufficient to derive the full range of verb cluster constructions in the head-final approach.

The same situation holds for the head-initial approach in (32), however, there is only one case which allows us to make the argument—the 2-3-1 order. To derive this order under a head-initial base structure by solely making use of head incorporation, two derivations have been suggested above (the two options are fleshed out in more detail in (37)). For the point to be made here, we do not have to decide between these two options since under both structures, the same prediction is made: V2 and V3 are part of a complex head, and hence, it should not be possible for phrasal material to occur between these two verbs. As we have seen above, this prediction is not borne out: examples such as the West Flemish (36), show that V2 and V3 do not form a complex head and hence a head movement derivation as in (37) has to be excluded.

(37) a. 2-3-1 order in head-initial approach (movement C before movement B)



b. 2-3-1 order in head-initial approach (tucking in)



In sum, the phenomenon of verb projection raising which is found in a subset of West Germanic languages and dialects poses a serious challenge for accounts that aim to derive verb clusters solely by head incorporation. This conclusion, however, does not force us to give up head movement as a means to derive verb clusters altogether.<sup>17</sup> Firstly, as many authors have suggested, verb cluster formation can involve both types of movement—certain configurations are derived by head movement others by phrasal movement (see Bennis and Hoekstra 1989, Haegeman 1990, 1992, 1994, 1998, Rutten 1991, Broekhuis et al. 1995, Zwart 1995, 1996, den Dikken and Hoekstra 1997 for accounts that involve phrasal and head movement derivations). Secondly, for

<sup>17</sup> But see den Dikken (1989), den Besten and Broekhuis (1992), Coppen and Klein (1992), Ijbema (1997), and Koopman and Szabolcsi (2000), among others, for exactly that position—i.e., a pure phrasal movement approach.

both the head-final and the head-initial approach, there is in principle an option to maintain a pure head movement account, coupled with several steps of independent movement of material interspersing a verb cluster. The assumption necessary would be that the verbs of a cluster do not form a complex head with each other (i.e., the verbs do not incorporate into each other), but that head movement targets different (non-overt) heads. That is, if for example in (33c), the lowest verb does not attach to the next higher verb but to a functional head between V2 and V1; or similarly, if in (37a), the modal (V2) does not attach to the higher verb but to a functional head above the AuxP, verb cluster movement could again be analyzed as head movement. One might object that these derivations involve violations of the HMC. However, as has been argued by den Dikken and Hoekstra (1997), the problem only arises under the strict version of the HMC. If the locality conditions for head movement—like A or A'-movement—are *relativized* (i.e., head movement is not blocked by intervening heads in general, but only by heads of the same type), skipping over certain heads becomes possible. Thus, while a general verb-to-verb incorporation account seems inadequate to capture the verb cluster phenomenon, a more abstract form of incorporation cannot be excluded.

### 3.2.1.2 *Excursus: The verb raising—restructuring connection*

The research on verb clusters has shown an interesting spin-off in the works on *restructuring* (a.k.a. *clause union*, *coherence*, or *complex predicates*). Restructuring is a topic on its own, and this overview can only highlight some aspects of the phenomenon in its connection to verb raising and verb clusters. In short, in many accounts of restructuring, some form of verb raising plays a crucial role. However, although verb clusters are necessarily restructuring constructions, restructuring is a much broader phenomenon, both within and across languages. There are numerous languages that show restructuring properties but do not allow verb clusters. Similarly, in verb cluster languages, we find constructions that allow restructuring but do not involve any verb cluster formation. Nevertheless, a connection between the two phenomena may exist if a more abstract form of verb raising is assumed to take place in restructuring configurations.

In an extensive study of infinitives, Gunnar Bech (1955) developed one of the first characterizations of infinitival complements in German. Bech showed that infinitives fall into two classes: infinitives that form an independent clausal domain (i.e., non-restructuring infinitives, which Bech called *kohärente Infinitive* 'coherent infinitives') and infinitives that do not exhibit clausal behavior (i.e., restructuring or *inkohärente* 'incoherent' infinitives). The first study of the two classes of infinitives in a generative framework was provided by Evers (1975a, b). Evers observed that the split among infinitival constructions in Dutch and German correlates with a reordering process of the verbal elements in an infinitival construction. He proposed that this reordering is the result of a process of *verb raising* which applies in certain infinitival clauses but not in others, thus establishing the connection between verb clusters and restructuring. Evers' analysis—which set the groundwork for most later analyses of clause structure in German and Dutch—is the first work that builds on the correlation between mono-clausality and verb movement. In particular, he suggests a *structure pruning principle* which applies to the S-node of infinitives that have lost their head (i.e., by movement of the embedded verb to the higher clause). The formulation of Evers' *Guillotine principle* is given in (38).<sup>18</sup>

---

<sup>18</sup> Evers attributes this principle to a proposal made by Kuroda in an unpublished paper (cited in Ross 1967:56).

Thus, in Evers' account, verb raising—i.e., the formation of a complex verb consisting of the matrix verb and the infinitive—creates a sentence that has lost its head and therefore the S-node cannot survive and is deleted, thereby creating a structure in which the matrix predicate and the infinitive form a mono-clausal domain.

(39) a. Lo verrò subito a scrivere Italian  
it.CL I.will.come at.once to write  
'I will come write it immediately.' [Rizzi 1982:38, (144a)]

b. Gli stessi errori si continuano stupidamente a commettere  
the same errors SI continue.3PL stupidly to make  
'People continue to make the same errors in a stupid way.' [Rizzi 1982:38, (144b)]

c. Maria è dovuta immediatamente tornare a casa  
Maria is must immediately return at home  
'Maria had to return home immediately' [Rizzi 1982:38, (144c)]

To account for the discontinuity of the two verbs found in Romance restructuring infinitives (while keeping Rizzi's assumption that the infinitive and the matrix verb form a complex head) two types of analyses have been proposed: excorporation and covert movement. The first approach (see den Dikken 1990, Roberts 1991, Guasti 1992, 1993, 1996, 1997) is based on exam-

39

ples such as the one in (40) from Italian in which a matrix causative and an embedded verb precede a floating quantifier associated with the matrix subject. Under the assumption that the floating quantifier originates in a position above the base position of the causative, the example in (40a) provides evidence for overt movement of the infinitive to the matrix predicate and further overt movement of both verbs to the left of the floating quantifier. Assuming the infinitive incorporates into the higher verb and the complex verb then undergoes further movement to the left, (40b) can then be taken to show that the matrix verb excorporates from the complex head.<sup>20</sup>

- (40) a. I professori facevano commentare tutti quel libro a Ugo  
The professors make comment all that book to Ugo  
'The professors all made Ugo comment on that book.'  
[Guasti 1997:136]
- b. I professori non fanno più commentare tutti quel libro a Ugo  
The professors NEG make anymore comment all that book to Ugo  
'All the professors do not make Ugo comment on that book anymore.'

The second type of analysis is based on the idea that restructuring constructions involve some form of covert head movement. The main reason for the assumption of covert verb raising is theory-internal. In approaches where it is assumed that restructuring infinitives start out as full clauses (so called bi-clausal approaches), a mechanism is required to unite the matrix and the embedded clauses in order to create the mono-clausal configuration. Since head movement is a standard tool to deactivate projections or barriers (see for instance Chomsky 1986, Baker 1988) this process has thus proven useful in the context of restructuring constructions to achieve the lack of clause-boundedness effects. A further reason that has been given repeatedly to motivate verb raising in restructuring constructions is the observation that the infinitival tense is deficient in these constructions (tense deficiency can be observed for instance by the impossibility of temporal modification of the embedded event in a restructuring infinitive or the fact that the embedded event has to be interpreted as simultaneous with the time of the matrix event). Based on this observation, Guéron and Hoekstra (1988) and Bennis and Hoekstra (1989) suggest that verbs have to be in a local relation with a tense head, which is generally achieved via movement of the verb to the closest tense head. What is special about restructuring constructions is that the infinitival tense head is deficient and thus not capable of licensing the infinitival verb. Hence, the verb (or the infinitival tense) has to move to the matrix clause.

There is a variety of different kinds of covert head movement approaches. One type of approach assumes that head movement in restructuring infinitives does not apply in overt syntax but takes place in the covert component (i.e., as LF-incorporation of the infinitive; see, among others, Sternefeld 1990, Grewendorf and Sabel 1994, Gonçalves 1998). Since restructuring phenomena such as clitic climbing, long object movement, or auxiliary switch are overt operations, a question for this approach (noted also in Roberts 1997) is how covert movement could license the overt restructuring properties. The second set of approaches assumes that head movement in restructuring infinitives does take place in overt syntax, but that for various reasons, this movement is not reflected in the (phonological) output of the structure. Adopting a version of the copy theory of movement in which either copy in a movement chain may be chosen at PF and LF,

<sup>20</sup> Note however, that this conclusion only holds under the particular assumptions stated in the text. One question of debate for instance is the claim that floating quantifiers mark the trace position of the subject (see Bobaljik 1998 and references therein for arguments against this position). Furthermore, excorporation could be dispensed with assuming the relativized HMC approach suggested in den Dikken and Hoekstra (1997).



Roberts (1997) suggests that head movement applies overtly, but that in the phonological component, the tail rather than the head of the chain is pronounced. Furthermore, many authors assume that the head that undergoes movement in a restructuring configuration is not the actual infinitival verb but rather an (abstract) infinitival tense, agreement, or voice head (cf. Kayne 1989, 1990, 1991, Roberts 1993, Bok-Bennema and Kampers-Manhe 1994, Rooryck 1994, Sabel 1996, Terzi 1996, Wurmbrand 2013, To appear). This form of overt (yet abstract) head movement thus creates a complex head in restructuring constructions but at the same time leaves behind the infinitival verb.

As a last point, it should be noted that, although there is an overlap between verb clusters and restructuring—verb clusters are a subset of the restructuring constructions—the two phenomena have to be kept separate. Cross-linguistically, restructuring is a common phenomenon (see Wurmbrand To appear), but verb cluster formation is very rare, essentially being restricted to the West Germanic languages (possibly also Hungarian; see É. Kiss and Riemsdijk 2004). Even within the languages allowing both restructuring and verb clusters, it has been observed that verb cluster formation cannot be seen as a necessary condition for restructuring. This has been noted very early on for Dutch. In Dutch, clausal non-restructuring infinitives do not license the IPP effect and the full embedded clause appears in extraposed position. Verb raising constructions as in (41a), on the other hand, involve the IPP effect and all but the embedded verb appears to the left of the matrix verb. As noted in den Besten and Rutten (1989), there is however another type of construction, which they therefore referred to as the *Third construction*, a term which has survived in the verb cluster research throughout the years. In the *Third construction*, as illustrated in (41b,c), no IPP effect is found, but parts or all of the embedded clause, except again the verb, appear to the left of the matrix verb. Matrix predicates differ regarding whether they allow the *Third construction* alongside a verb raising configuration with the IPP (such as the *try* infinitive in (41b)) or whether they only allow the *Third construction* and prohibit verb raising and the IPP entirely (such as the *decide* infinitive in (41c)). Den Besten and Rutten (1989) propose to analyze the *Third construction* as a form of long scrambling plus remnant extraposition of the embedded infinitives as in (41d). Crucially, this form of long scrambling is only possible from restructuring infinitives. Thus, restructuring exists independently of verb raising.

- d. .... SCR *try/decide*.PART [XP=restructuring infinitive t<sub>SCR</sub> V.INF]

The same conclusion is reached for examples such as (42) from German.

- (42) [Zu reparieren ]<sub>VP</sub> wurden nur blaue Autos vergessen German  
 [to repair ]<sub>VP</sub> were only blue cars forgotten  
 ‘They only forgot to repair blue cars.’ [Wurmbrand 2007: 264]

The example in (42) involves long object movement, a property only possible in restructuring contexts. Since the embedded verb clearly does not form a complex head with the matrix verb but rather heads its own VP (it appears in matrix initial position, which can only be occupied by a maximal projection in German), yet restructuring is licensed, restructuring cannot require (overt) verb raising or verb cluster formation. However, as mentioned above, it is possible that restructuring involves a form of abstract complex predicate formation.

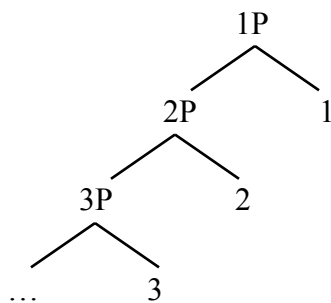
### 3.2.1.3 Phrasal movement derivations

In this section I address the question of whether verb clusters can be derived solely by phrasal movement or whether certain constructions necessitate head movement. Before laying out arguments that have been presented for head movement in verb clusters, a quick overview of the different phrasal movement operations necessary to derive verb clusters in the head-initial and head-final approach will be provided. The purpose of this illustration of the basic derivations is to show that none of the orders poses an a priori challenge for phrasal movement accounts.

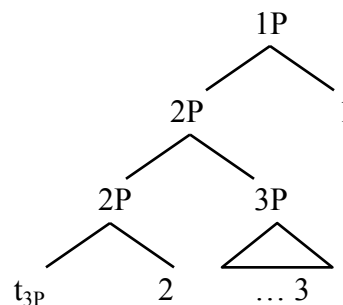
To begin with the head-final approach, the 3-2-1 order in (43a) is the underlying order and hence no movement is necessary (or all movements are covert). The 2-3-1 order can be derived by movement of the lowest VP to the right of the middle VP as in (43b). Movement of the entire middle VP to the right of the highest VP yields the 1-3-2 order as in (43c) (see den Besten and Broekhuis 1992 for one of the first suggestions of this derivation). Movement of both lower VPs as in (43d) results in the 1-2-3 order. Lastly, the 3-1-2 order can be derived as in (43e) by leftward movement of the lowest VP (and adjunction to either 2P or 1P), followed by rightward movement of the (now remnant) middle VP (in the first option, only the lower segment of 2P moves to strand behind 3P).

#### (43) Head-final derivations

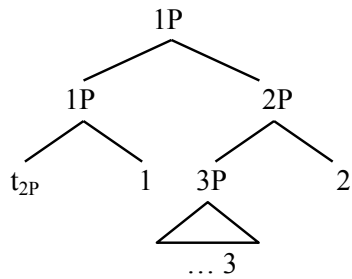
- a. 3-2-1: no movement



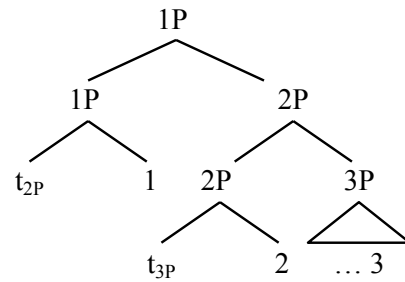
- b. 2-3-1: movement of 3P



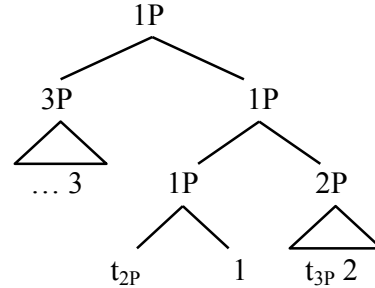
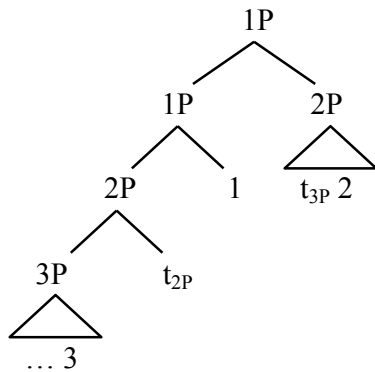
c. 1-3-2: movement of 2P



d. 1-2-3: movement of 3P and 2P



e. 3-1-2: 3-LEFT and movement of 2P

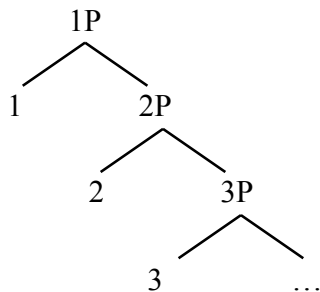


To exclude the unattested 2-1-3 order of the verb clusters, one can follow Haegeman (1992) who suggests that verb clusters of this type are subject to some form of *Relativized Minimality*. If it is assumed that verb cluster movement has to target the next higher VP and that intervening VPs (except adjunction segments) cannot be skipped, movement of the lowest VP and adjunction to the highest XP by skipping the middle XP would be prohibited, and hence the 2-1-3 order could not be derived. If this approach is adopted, then only the first derivation in (43e) is possible.

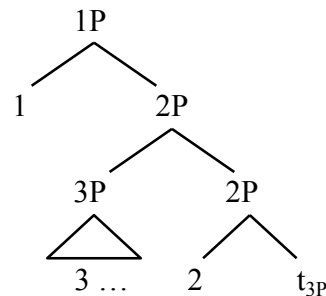
Under the head-initial approach in (44a), the 1-2-3 order is the basic order, which does not require any movement. To derive the 1-3-2 order, the lowest VP undergoes movement to left of the middle VP as in (44b). Movement of the middle VP to the left of the highest VP yields the 2-3-1 order in (44c). Movement of both lower VPs results in the 3-2-1 order in (44d). Finally, the 3-1-2 order is derived by moving the lowest VP to the left of the highest VP as in (44e) (whether it moves in one step or stops in the specifier of the middle VP is not crucial at this point).

#### (44) Head-initial derivations

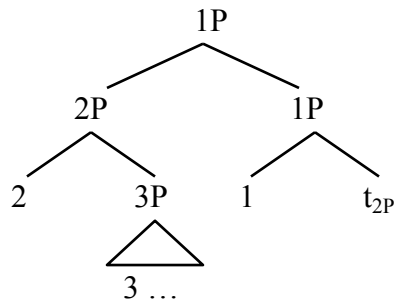
a. 1-2-3: no movement



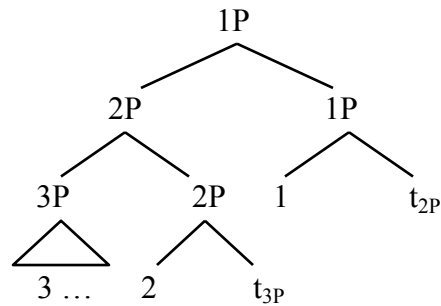
b. 1-3-2: movement of 3P



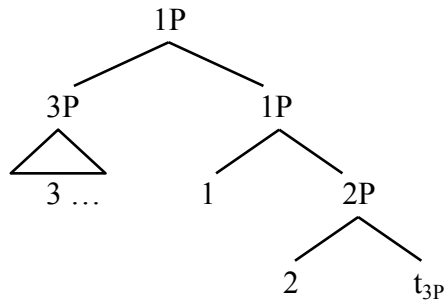
c. 2-3-1: movement of 2P



d. 3-2-1: movement of 3P and 2P



e. 3-1-2: 3-LEFT (possibly stopping over in 2P)



In order to block the 2-1-3 order, one could assume, for instance, that movement of the lowest VP targets the specifier position of the middle VP (and not for instance the specifier position of a functional projection between 1P and 2P). Since the lowest VP would then be part of the middle VP, movement of the latter would always carry along VP3 and hence 2-1-3 could not be generated. Alternatively, one could assume that traces must be c-commanded by their antecedents (at surface structure). This would exclude a derivation in which 3P vacates 2P by adjoining to a projection between 1P and 2P, followed by (remnant) movement of 2P above 1P (this option would obviously be unavailable in approaches built on such remnant movement such as Koopman and Szabolcsi 2000). A final option is, as discussed in section 3.1, to follow Abels (2011) and Abels and Neeleman's (2012) who propose that verb cluster movement must always include the main verb.

The illustrations in (43) and (44) show that both head-final and head-initial approaches have the capacity to derive the possible verb cluster orders via phrasal movement operations. In contrast to (pure) head movement approaches, the challenge for phrasal movement approaches is generally not that they are too restrictive but rather that they are too permissive. Deriving verb projection raising configurations is trivial in the structures in (43) and (44). What is less clear, however, is how verb projection raising configurations can be excluded in languages that do not allow phrasal material between the elements of a cluster (see section 3.2.3). The way to achieve these restrictions in phrasal movement structures is by employing additional (leftward) movement operations that (fully or partially) pre-empty the VPs that undergo verb cluster movement (i.e., remnant movement instead of full phrasal movement has to become obligatory in certain languages and constructions).

#### 3.2.1.4 Mixed accounts—head-final perspective

The question addressed in this section is whether a pure phrasal movement approach is sufficient to derive verb clusters across West Germanic or whether there are reasons to assume that certain constructions must (also) involve head movement. A number of interesting arguments based on locality and the size of the moved material have been put forward for the claim that certain (parts of) verb clusters are derived by head movement. Before presenting this evidence, it should be noted that there are certain phenomena—in particular certain adjacency requirements—which will not be considered as arguments for head movement here. As discussed in section 3.2.1.1, a commonly accepted restriction for head movement configurations is that a complex head cannot include phrasal material. Thus, if phrasal material can occur between two verbs, it can be concluded that these two verbs do not form a complex head (at least not in overt syntax). However, the opposite does not hold. If a verb cluster disallows phrasal material between the verbs, it does not follow that the structure necessarily involves a head raising configuration. The non-conclusiveness of arguments for a complex head structure based on adjacency requirements between the relevant verbs has been noted in several works. For instance, in German and Dutch clause-final position, it is impossible to intersperse a ...2-1 verb cluster by extraposed material (see van Riemsdijk 1998, Haider 2003, among many others). However while it is true that an overt head movement configuration entails adjacency between the elements involved, adjacency does not entail a head movement configuration (for instance, Case adjacency in English does not entail that the verb and the object form a complex head). As for verb clusters, the extraposition puzzle has received alternative accounts, which relate the impossibility of extraposed material between two verbs of a cluster to independently observed restrictions on extraposition rather than the formation of a syntactic complex head (see Wurmbrand and Bobaljik 2005, Wurmbrand 2007; see also Salzmann 2013a, b for a recent non-syntactic complex head account). These approaches have the advantage that the elements of a cluster remain part of separate projections in syntax, which, as was shown in (42), is clearly possible.

The first type of argument for head movement in verb clusters has been provided based on locality, specifically the assumption that head and phrasal movement are subject to different locality domains (only heads intervene in head movement dependencies and only XPs intervene in phrasal movement dependencies). The lack vs. existence of certain locality effects can then be taken as evidence for/against a particular type of movement. An elaborate instantiation of this argument can be found in den Dikken and Hoekstra's (1997) investigation of the Frisian PPI effect—the occurrence of participles instead of infinitives—illustrated in (45).

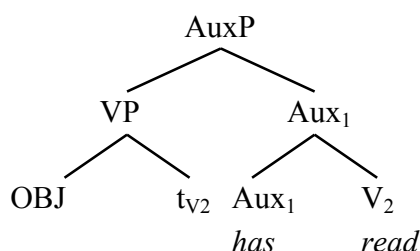
- (45) a. hy    soe    it    dien<sub>3</sub>    /   dwaan<sub>3</sub>    wollen<sub>2</sub>    ha<sub>1</sub>    Frisian  
          he    would   it    do.PART<sub>3</sub> / do.INF<sub>3</sub>    want.PART<sub>2</sub>    have.INF<sub>1</sub>  
          ‘He would have liked to do it.’    [den Dikken and Hoekstra 1997: 1058]
- b. hy    soe    it    dien<sub>3</sub>    ha<sub>2</sub>    kinne<sub>1</sub>    /   kinnen<sub>1</sub>  
          he    would   it    do.PART<sub>3</sub>    have.INF<sub>2</sub>    can.INF<sub>1</sub> / can.PART<sub>1</sub>  
          ‘He would be able to have it done.’    [den Dikken and Hoekstra 1997: 1070, 1074]

The analysis suggested in den Dikken and Hoekstra (1997)), which crucially makes use of both head movement and phrasal movement, not only derives the word orders in Frisian verb clusters but also accounts for the existence of the PPI effect and the distribution thereof. Although it is not possible to replicate the details of the analysis here, the conclusion is essentially that move-

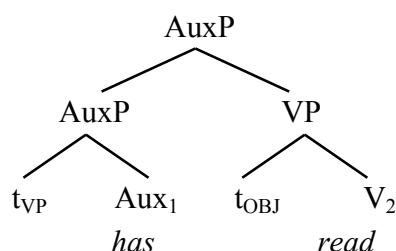
ment of participles (which is necessary to license both true and parasitic participles) can skip intervening heads of the same type (i.e., heads with the same features), whereas movement of auxiliaries is blocked by intervening heads with the same features. Assuming a restrictive theory of movement, the authors conclude that participle movement is phrasal movement (hence not sensitive to intervening heads), whereas in the same language, movement of auxiliaries is head movement (hence subject to the relativized HMC).

A second important contribution to the issue of head vs. phrasal movement has been provided in den Besten and Rutten 1989, (see also Rutten 1991, den Besten and Broekhuis 1992, Broekhuis et al. 1995 and many others). The following discussion will only apply to head-final structures. As mentioned at the beginning of this section, a crucial difference between the head movement structure in (29b) and the phrasal movement structure in (29c) (repeated below in (46a,b) is that in the latter a remnant constituent is moved—i.e., all but the verb has to have left the VP prior to verb cluster movement. This difference leads to the following predictions. If the leftward movement step is prohibited (i.e., if a construction involves VP-internal elements that are not allowed to scramble or object shift), and a construction of the form “XP-1-2” is well-formed, the derivation can only involve head movement. If, on the other hand, leftward movement of (certain) VP-internal XPs is prohibited and a construction of the form “XP-1-2” is ungrammatical in these contexts (but well-formed when the VP contains movable XPs), the derivation can only involve phrasal movement. In what follows, we will see that in Dutch under certain assumptions, both predictions find an instantiation.

(46) a. Head-movement (head-final)



b. XP-movement (head-final)



Contexts that have been taken to involve elements that cannot undergo movement (but see below) are particle [PTCL] constructions such as *to call up* and (semi-)idiomatic expressions such as *to take into consideration*. As is illustrated in (47a,b), (non-idiomatic) objects are free to move to the left of adverbs such as *toen toch* (with some subtle differences in meaning). For material that is part of an idiomatic expression, (47c), and particles, (47d), on the other hand, this movement operation is prohibited (the position of the object has no influence on the ungrammaticality).

(47) a. dat Jan {dat aanbod} toen toch {dat aanbod} in overweging nam DU  
 that Jan {that offer} then ADV {that offer} in consideration took  
 ‘that Jan then did take that offer into consideration’

b. dat Marie {haar vader} toen toch {haar vader} op belde  
 that Marie {her father} then ADV {her father} up called  
 ‘that Marie then did call her father’

c. \*dat Jan dat aanbod in overweging toen toch nam  
 that Jan that offer in consideration then ADV took  
 ‘that Jan then did take that offer into consideration’

- d. \*dat Marie haar vader op toen toch belde  
 that Marie her father up then ADV called  
 ‘that Marie then did call her father’

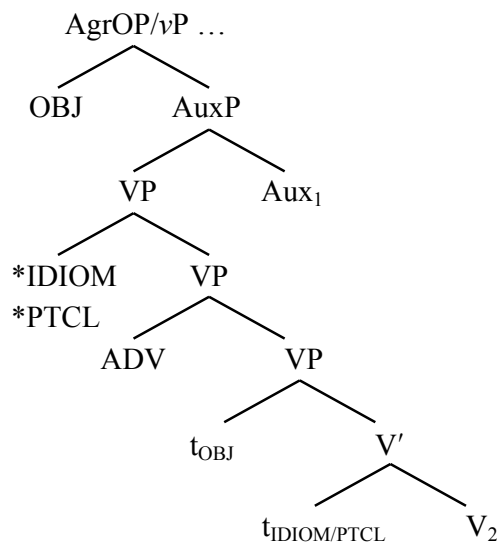
Assuming that adverbs such as *toen toch* mark the left edge of the VP, one can conclude (but see again below) that parts of idiomatic expressions and particles cannot leave the VP. These types of constructions thus offer contexts in which the two predictions mentioned above can be tested. Let us first look at a situation where the constructions in (47) are embedded in a verb cluster (in particular an auxiliary-participle construction). In the 1-2 order (which is the relevant order for the head-final approach), idiomatic material can only appear to the (immediate) left of the auxiliary as in (48a) and can neither occur to the left of the VP-adverb (cf. (48c)) nor between the auxiliary and the participle (cf. (48e)). Particles, on the other hand, are also prohibited from occurring to the left of a VP-adverb (cf. (48d)), but they can appear either to the (immediate) left of the auxiliary as in (48b) or between the auxiliary and the participle as in (48f).

- (48) a. dat Jan {dat aanbod} toen toch {dat aanbod} in overweging heeft<sub>1</sub> genomen<sub>2</sub>  
 that Jan {that offer} then ADV {that offer} in consideration has<sub>1</sub> taken<sub>2</sub>  
 ‘that Jan then HAS taken that offer into consideration’
- b. dat Marie {haar vader} toen toch {haar vader} op heeft<sub>1</sub> gebeld<sub>2</sub>  
 that Marie {her father} then ADV {her father} up has<sub>1</sub> called<sub>2</sub>  
 ‘that Marie then HAS called her father’
- c. \*dat Jan dat aanbod in overweging toen toch heeft<sub>1</sub> genomen<sub>2</sub>  
 that Jan that offer in consideration then ADV has<sub>1</sub> taken<sub>2</sub>  
 ‘that Jan then has taken that offer into consideration’
- d. \*dat Marie haar vader op toen toch heeft<sub>1</sub> gebeld<sub>2</sub>  
 that Marie her father up then ADV has<sub>1</sub> called<sub>2</sub>  
 ‘that Marie then has called her father’
- e. \*dat Jan dat aanbod toen toch heeft<sub>1</sub> in overweging genomen<sub>2</sub>  
 that Jan that offer then ADV has<sub>1</sub> in consideration taken<sub>2</sub>  
 ‘that Jan then has taken that offer into consideration’
- f. dat Marie {haar vader} toen toch {haar vader} heeft<sub>1</sub> op gebeld<sub>2</sub>  
 that Marie {her father} then ADV {her father} has<sub>1</sub> up called<sub>2</sub>  
 ‘that Marie then has called her father’

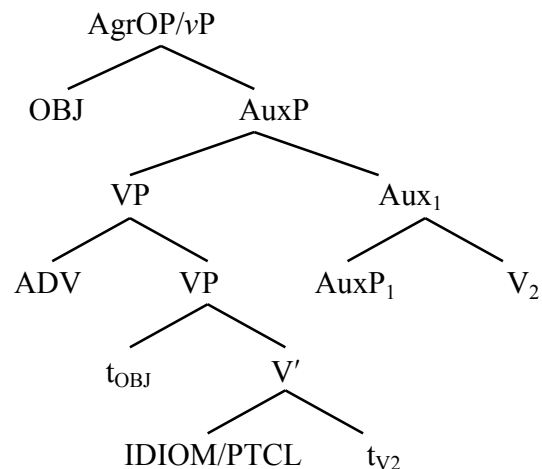
Under the assumption that parts of idiomatic expressions and particles cannot leave the VP, the distribution in (48) provides an argument for head movement and against phrasal movement in this type of construction. Let us start with (48c,d). The prohibition against movement of parts of idioms and particles as in (49a) immediately accounts for the ungrammaticality of these examples (independent of the question of whether the lower verb has undergone head or phrasal movement). Furthermore, assuming idioms and particles cannot leave the VP, remnant VP-movement cannot be an option for (48a,b). If, on the other hand, the examples in (48a,b) involve head movement such as in (49b), idioms and particles can be stranded in the VP when the lower verb (head) moves and adjoins to the higher auxiliary. Finally, assuming that (at least in this language and construction) verb cluster formation can only involve head movement and phrasal movement is prohibited, (48e) is expected to be ungrammatical since in this case more than the

verbal head (the (remnant) VP) has undergone movement. Similarly, (48f) can be accounted for if it is assumed that the particle and the verb form a complex head at the level where verb cluster formation applies. That is, either particle constructions are base generated complex verbal heads (among others, Neeleman and Weerman 1993, Neeleman 1994 and references therein) or involve incorporation of the particle into the verb (den Dikken 1992, 1995a, Koopman 1995 among many others).<sup>21</sup> Under the assumptions mentioned, particle and (semi-)idiomatic constructions can thus be seen as an instantiation of the first prediction—verb cluster formation in auxiliary-participle constructions (and the same holds for modal-infinitive constructions) can only involve head movement in Dutch.

(49) a. Illicit particle/idiom movement



b. Head movement and particle stranding



<sup>21</sup> To account for the optionality of particle stranding (cf. (48b) and (48f)) it has to be assumed that either exorporation is possible or that particle incorporation can occur before or after verb cluster formation. Since the structure of particle constructions is a topic on its own, a detailed discussion cannot be provided here. See Neeleman (1994), den Dikken (1995a), and references therein for in-depth discussions.



Before turning to an alternative analysis, cases instantiating the second prediction will be discussed. In Dutch, some form of verb cluster formation (or rather restructuring) is also found in certain infinitival constructions involving the infinitival marker *te* ‘to’. As is shown in (50), the object appears to the left of the matrix verb, indicating mono-clausality, whereas the particles and idiomatic phrases occur between the two verbs. The grammaticality of (50a,b) thus contrasts sharply with the ungrammaticality of the auxiliary-participle construction in (48e). Under a head-final base structure, (50a,b) then lead to the conclusion that remnant XP-movement is possible in infinitival constructions involving matrix verbs such as *try* or *decide*.<sup>22</sup>

- (50) a.    dat    Elsje   de prijs       probeerde<sub>1</sub>   [ in ontvangst    te nemen<sub>2</sub> ]<sub>XP</sub>  
           that    Elsje   the award    tried<sub>1</sub>       [ in acceptance   to take<sub>2</sub>    ]<sub>XP</sub>  
           ‘that Elsje tried to accept the award’
- b.    dat    Elsje   de prijs       beslot        [ in ontvangst    te nemen<sub>2</sub> ]<sub>XP</sub>  
           that    Elsje   the award    decided    [ in acceptance   to take<sub>2</sub>    ]<sub>XP</sub>  
           ‘that Elsje decided to accept the award’
- c.    dat    Elsje   haar vader   probeerde<sub>1</sub> op    te bellen<sub>2</sub>  
           that    Elsje   her father    tried<sub>1</sub>    up    to call<sub>2</sub>  
           ‘that Elsje tried to call her father’
- d.    dat    Elsje   haar vader   beslot<sub>1</sub>    op    te bellen<sub>2</sub>  
           that    Elsje   her father    decided<sub>1</sub> up    to call<sub>2</sub>  
           ‘that Elsje decided to call her father’

To determine whether phrasal movement is possible or necessary, we must look again at contexts in which phrasal movement would be excluded. As above, the immovability of particles and parts of idiomatic expressions allows us to construct such a context. If these elements can occur to the left of the higher verb in the 1-2 order, the construction can only involve head movement (again assuming that particles and idiom chunks cannot leave their base position); if these elements cannot occur to the left of the higher verb in the 1-2 order, head movement is prohibited. Changing the examples in (50) along these lines reveals an interesting contrast which is illustrated in (51).

- (51) a.    dat    Elsje   de prijs       in ontvangst    probeerde<sub>1</sub> te nemen<sub>2</sub>  
           that    Elsje   the award    in acceptance    tried<sub>1</sub>        to take<sub>2</sub>  
           ‘that Elsje tried to accept the award’
- b.    \*dat    Elsje   de prijs       in ontvangst    beslot<sub>1</sub>        te nemen<sub>2</sub>  
           that    Elsje   the award    in acceptance    decided<sub>1</sub>    to take<sub>2</sub>  
           ‘that Elsje decided to accept the award’
- c.    dat    Elsje   haar vader   op    probeerde<sub>1</sub> te bellen<sub>2</sub>  
           that    Elsje   her father    up    tried<sub>1</sub>        to call<sub>2</sub>  
           ‘that Elsje tried to call her father’

<sup>22</sup> The label of the projection in (50) is left unspecified here. As mentioned in section 3.2.1.2, there are several approaches to restructuring, and the details of the structure depend on one’s view of restructuring.

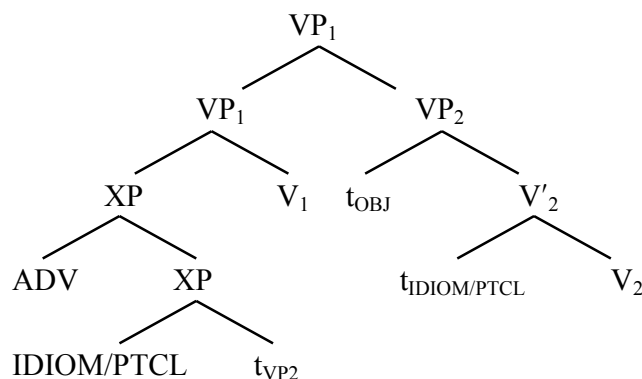
- d. \*dat    Elsie   haar vader   op   beslot<sub>1</sub>       te bellen<sub>2</sub>  
          that   Elsie   her father   up   decided<sub>1</sub>      to call<sub>2</sub>  
          ‘that Elsie decided to call her father’

While *try* constructions allow particles and idiomatic phrases to occur to the left of the higher verb, *decide* constructions prohibit stranding of these elements. Assuming again that movement of particles and idiomatic phrases is excluded, one can conclude from the examples in (51) that head movement is possible in *try* constructions but not an option in *decide* constructions. Thus, infinitives of the latter type instantiate the second prediction—verb cluster formation in infinitival constructions of the *decide* type can only involve phrasal movement in Dutch.

To recapitulate, the account sketched leads to the conclusion that there are three types of verb clusters: verb clusters that only allow head movement (auxiliary-participle constructions), verb clusters that only allow phrasal movement (*decide* type infinitives), and verb clusters that allow both head and phrasal movement (*try* type infinitives). Although this account seems advantageous in that it allows us to develop criteria to distinguish between head and phrasal movement in verb clusters, one might object that the arguments presented in (47) to (51) are built on assumptions that are not necessarily generally shared (such as a head-final base structure and the immovability of particles and idiom chunks) and that a unified account (i.e., an analysis that employs only one type of movement in verb cluster constructions) might be preferable. In the rest of this section, an outline of such an alternative analysis will be presented.

The crucial examples that led to the conclusion that certain constructions can only involve head movement are (48a,b,e), (51a,c), and under certain assumptions about the structure of particle verbs also (48f). However, (49b) is not the only derivation for these examples, and as shown in (52), a phrasal movement account is conceivable as well. To derive the correct word order, the idiomatic phrase or particle first moves out of the VP and (string-vacuously) adjoins to the VP or a functional projection between the adverb and the VP (XP in the structure). The lowest VP can then undergo remnant phrasal movement to the right, ‘standing’ the particle or idiom chunk to the left of the higher verb.

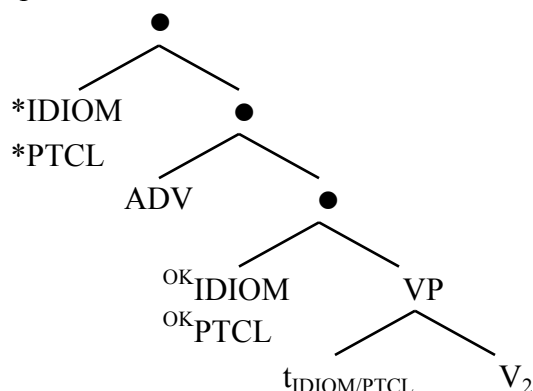
(52) Phrasal movement



A derivation such as (52) has been excluded above by the assumption that particles and parts of idioms cannot move out of the VP. The reason for this assumption was the ungrammaticality of the examples in (48c,d) and (51b,d)—i.e., examples that show that particles and parts of idioms cannot occur to the left of VP-adverbs. In the previous account, this fact was seen as a direct consequence of the assumption that idiomatic phrases and particles are immobile—i.e., they cannot undergo any kind of movement. To accommodate these facts under a phrasal movement ac-

count such as the one outlined in (52), it can then not be the case that movement of idiomatic phrases and particles is inherently blocked (otherwise the derivation in (52) would be impossible). Rather it has to be assumed that particles and idiomatic phrases are able to move, however, only as long as they do not cross any other material (see den Besten and Broekhuis 1992, Broekhuis et al. 1995 for claims along these lines).<sup>23</sup> Thus, short string-vacuous (i.e., shape preserving) movement as in (52) is licensed, however, movement across the adverb as in (48c,d) is ruled out (cf. (53); the specific labels of the projections targeted by movement are not relevant, hence simply given as ●). At this point, one has to ask why particle and idiom movement should be subject to such a restriction. However, note that the same question arises for the claim that particles and idiom chunks are immobile. A detailed motivation cannot be given here, but it seems that either assumption could be motivated by some syntactic or semantic adjacency requirement between the different parts of an idiomatic expression and between the particle and the verb or a shape preservation condition.

(53) Idiom/particle movement



While an analysis along these lines accounts for the examples in (48a-d), it raises some questions concerning the example in (48e) and examples involving *te*-infinitives. Starting with the former, the importance of (48e) (repeated as (54a)) and the examples in (54b-e) is that in auxiliary-participle constructions, no phrasal material can intervene between the auxiliary and the participle. This restriction follows straightforwardly in a head movement approach. For the phrasal movement account in (52) and (53), however, it would mean that objects and idioms (but not particles; cf. (48f)) not only can move out of the VP, but in fact must leave the VP obligatorily. Furthermore, it has to be ensured that only the lowest VP undergoes movement to the right in this construction.

- (54) a.    \*dat    Jan    dat aanbod    toen toch    heeft<sub>1</sub>    in overweging    genomen<sub>2</sub>  
           that    Jan    that offer    then ADV    has<sub>1</sub>    in consideration    taken<sub>2</sub>

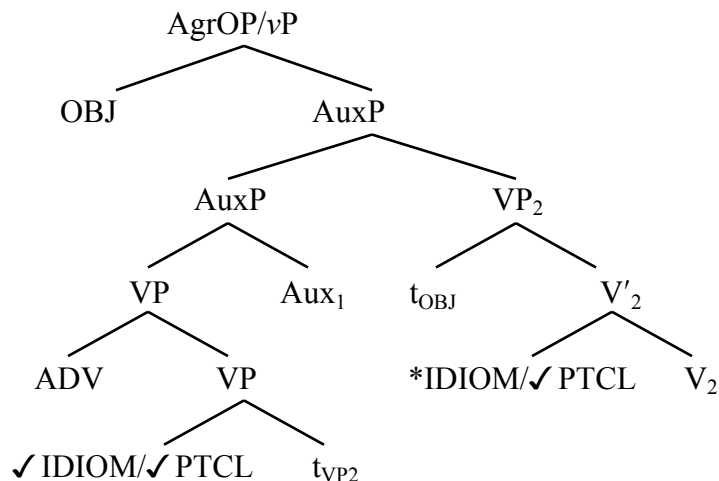
<sup>23</sup> In fact, particles can appear to the left of certain elements (namely stranded prepositions) which has been taken as support for a movement analysis of particles and against the claim that particles are immovable (cf. for instance Zwart 1997). However, since these reduced elements are the only categories that can appear between a particle and a verb, the conclusion that particles undergo leftward movement should also be taken with some caution. An alternative view would be to assume that particles and other preposition-like elements undergo syntactic or phonological cliticization; if this process occurs in the phonological component, the order between a particle and a stranded preposition could be seen as a phonological ordering effect and hence these constructions would not provide evidence for syntactic movement.

- b. \*dat Jan toen toch een aanbod heeft<sub>1</sub> in overweging genomen<sub>2</sub>  
that Jan then ADV an offer has<sub>1</sub> in consideration taken<sub>2</sub>
- c. \*dat Jan toen toch in overweging heeft<sub>1</sub> een aanbod genomen<sub>2</sub>  
that Jan then ADV in consideration has<sub>1</sub> an offer taken<sub>2</sub>
- d. \*dat Jan toen toch heeft<sub>1</sub> een aanbod in overweging genomen<sub>2</sub>  
that Jan then ADV has<sub>1</sub> an offer in consideration taken<sub>2</sub>
- e. \*dat Jan dat aanbod in overweging heeft<sub>1</sub> toen toch genomen<sub>2</sub>  
that Jan that offer in consideration has<sub>1</sub> then ADV taken<sub>2</sub>

all intended: ‘that Jan then HAS taken that/an offer into consideration’

Thus in sum, the assumptions necessary to account for the distribution of particles and idiomatic phrases in auxiliary-particle verb clusters under a phrasal movement approach are as follows: i) all VP-internal elements except particles are required to leave the VP; ii) particles can move out of the VP; iii) idiomatic phrases and particles can only move string-vacuously; and iv) only the lowest VP can move to the right (cf. (55)).

(55) XP-movement



Note that this analysis has a direct consequence for the structure of examples such as the one in (56a), where both the object and the idiomatic phrase appear below the VP-adverb. In the phrasal movement analysis just sketched, both the object and the idiomatic phrase must have left the VP and attached *below* the VP-adverb (as is evident from (56b), these movement operations can again not change the underlying order between the two phrases).

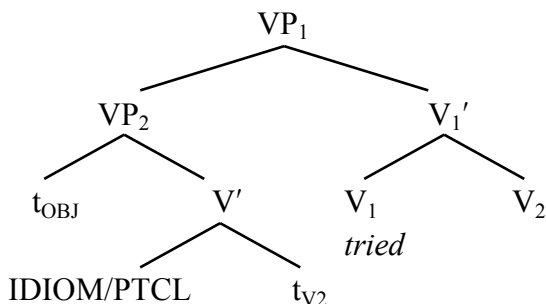
- (56) a. dat Jan toen toch een aanbod in overweging heeft<sub>1</sub> genomen<sub>2</sub>  
that Jan then ADV an offer in consideration has<sub>1</sub> taken<sub>2</sub>  
‘that Jan has taken an offer into consideration at that moment’
- b. \*dat Jan toen toch in overweging een aanbod heeft<sub>1</sub> genomen<sub>2</sub>  
that Jan then ADV in consideration an offer has<sub>1</sub> taken<sub>2</sub>  
‘that Jan has taken an offer into consideration at that moment’

The second set of examples that requires some adjustment in a pure phrasal movement approach involves the examples in (51), schematically repeated in (57).

- (57) a. *that* SUBJ OBJ *decided* IDIOM/PTCL *to* VERB  
 b. *that* SUBJ OBJ *tried* IDIOM/PTCL *to* VERB  
 c. \**that* SUBJ OBJ IDIOM/PTCL *decided* *to* VERB  
 d. *that* SUBJ OBJ IDIOM/PTCL *tried* *to* VERB

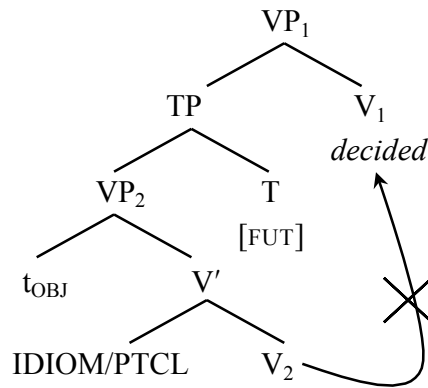
As has been laid out above, an account that makes use of head and phrasal movement offers a straightforward explanation for the distribution in (57): *try* infinitives allow both head movement and phrasal movement, whereas *decide* infinitives only permit phrasal movement. An interesting question is why only certain infinitival constructions permit head movement. Without going into details at this point, a common feature of most approaches to this problem is the claim that *try* and *decide* infinitives represent two different types of infinitival constructions. While *try* infinitives are in some sense deficient (i.e., they either lack certain functional projections or involve projections that are or become ineffective by various mechanisms; see section 3.2.1.2), *decide* infinitives are made of more ‘solid’ material (i.e., projections that make them less transparent and hence block certain operations). It has been observed that, semantically, *try* and *decide* infinitives differ in that the latter involves an embedded future interpretation (the time of the embedded event is situated after the matrix event time), whereas the former triggers a simultaneous interpretation (see Wurmbrand 2001, 2014). If this difference also shows a structural reflex, *try* infinitives would lack temporal projections as in (58a), whereas *decide* infinitives would involve a temporal projection encoding future as in (58b). Based on these structures, the prohibition of head movement in *decide* constructions can then be reduced to some version of the HMC. That is, if *decide* infinitives but not *try* infinitives include a tense projection, the intervening tense head would block head movement across it in (58b) but not in (58a). On the other hand, no problem arises for phrasal movement in *decide* infinitives.<sup>24</sup>

- (58) a. Head movement: *try* infinitive



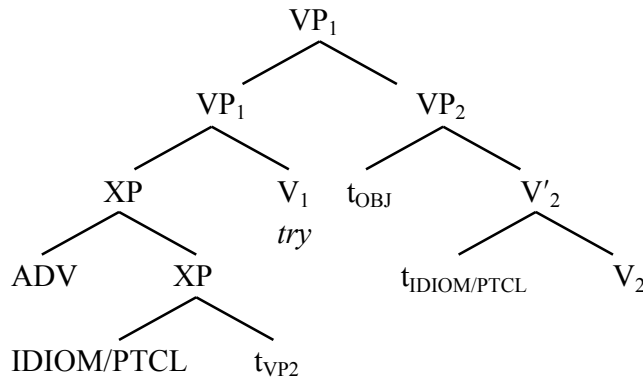
<sup>24</sup> Assuming that phrasal movement in verb clusters is some form of extraposition, it is expected that this form of movement does not have to target the closest XP on the right but can attach to a higher projection, as long as no clause boundary is crossed. Thus, movement of VP<sub>2</sub> in (58b) across TP would be unproblematic unless there is also a CP. Since not all infinitival constructions allow the remnant movement operation in (57a,b) (in particular, factive and propositional infinitives prohibit this configuration), the presence vs. absence of a CP, like the suggested presence vs. absence of the TP in (58), could be seen as the crucial property that distinguishes infinitives allowing a remnant movement construction from ones that prohibit it (see Wurmbrand 2001 for an analysis along these lines).

- b. Head movement: *decide* infinitive

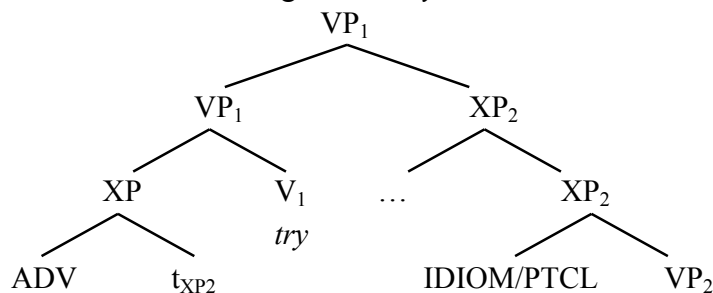


Returning to the account that only employs phrasal movement to derive verb clusters (as outlined in (52b), (53), and (55)), the derivation of (57d) (i.e., a *try* infinitive with a stranded particle or idiomatic phrase) is straightforward. As before, the idiomatic phrase leaves the VP, and the lowest VP undergoes movement to the right as in (59a). To account for the difference between auxiliary-participle constructions and *try* infinitives (i.e., the contrast between (48e) and (57b)) in the mixed account, it has been suggested that the former only allow head movement, whereas the latter allow head or phrasal movement. A similar distinction can be made in the pure phrasal movement approach. While auxiliary-participle constructions permit only movement of the lowest VP, *try* infinitives also allow movement of higher projections as in (59b).<sup>25</sup>

- (59) a. Phrasal movement lowest VP: *try* infinitive



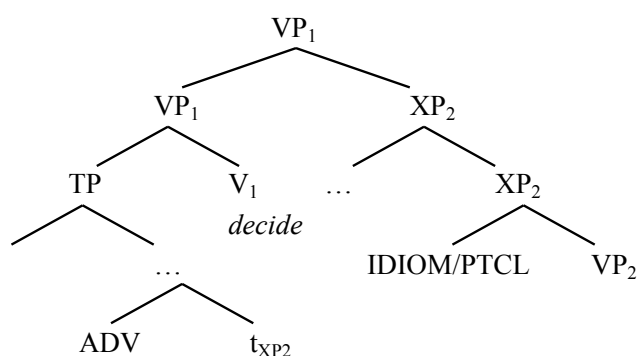
- b. Phrasal movement higher XP: *try* infinitive



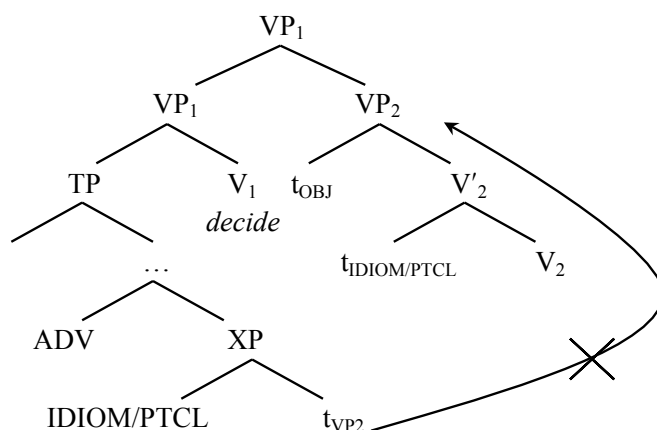
<sup>25</sup> Note that the main issue of this section is simply to determine how different structures can be derived. The question of why certain things move or do not move will be addressed in section 3.3.

Turning lastly to *decide* infinitives, the assumption necessary is that in contrast to *try* infinitives only the higher VPs/XPs can undergo rightward movement. Thus, depending again on the derived position of the object, either the XP or TP moves to the right as in (60a), but movement of the lowest VP as in (60b) is prohibited. While under the mixed account, the more restricted behavior of *decide* infinitives could be derived from the structure of these infinitives (i.e., the presence of a tense head, which blocks head movement), it is not a priori clear that the presence vs. absence of functional projections in the infinitive should have an effect on the possibility vs. impossibility of movement of the lowest VP in the pure phrasal movement approach. All else being equal, considerations of parsimony may seem to favor the mixed account sketched above, however, the ultimate choice between the two approaches evidently depends on the specific locality conditions adopted. For instance a version of *Shortest Move* could certainly exclude movement as in (60b), however, it would then need to be seen whether such a restrictive theory of extraposition can be maintained for extraposition in general.

(60) a. XP-movement: *decide* infinitive



b. \*VP-movement: *decide* infinitive



To conclude, if it is assumed that idiomatic phrases and particles are immobile, the distribution of these elements in Dutch verb clusters leads to the conclusion that both head and phrasal movement are necessary to derive verb clusters and restructuring under a head-final base structure. If it is assumed that idiomatic phrases and particles are not immobile per se but subject to a shape preservation constraint, a pure phrasal movement approach can in principle be maintained. Table 5 summarizes the constructions discussed and the assumptions necessary to account for the distribution of these constructions in the mixed account and a pure phrasal movement approach.

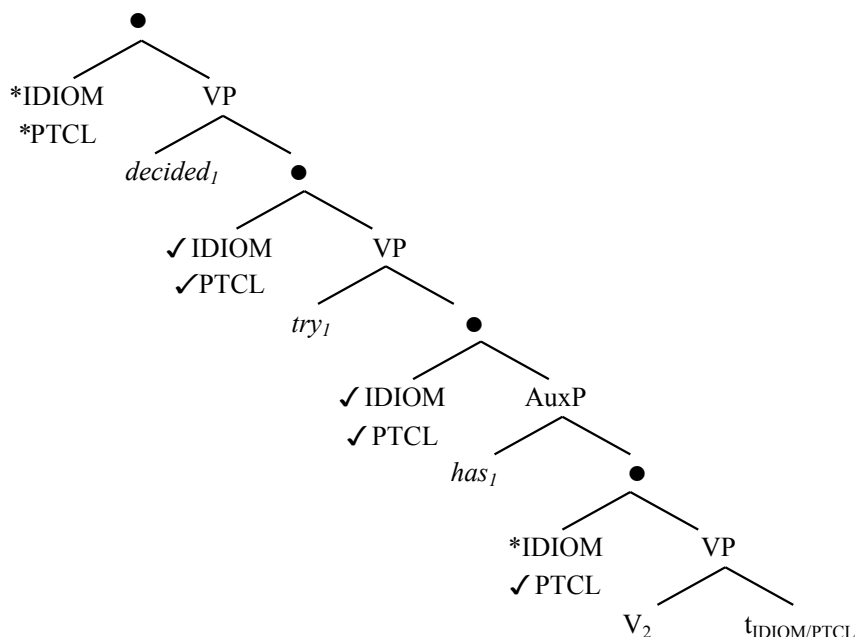
Table 5. Head and/or phrasal movement in head-final approaches

Constructions	Head or phrasal movement	Only phrasal movement
AUX-PART	V2 movement	VP2 movement; *XP movement
Tenseless infinitives ( <i>try</i> )	V2 or VP2 movement	VP2 or XP movement
Future infinitives ( <i>decide</i> )	VP2 (or larger) movement *V2 movement (HMC)	XP (or larger) movement *VP2 movement
Particles	No movement	Movement; shape preservation
Idiomatic phrases	No movement	Movement; shape preservation

### 3.2.1.5 Mixed accounts—head-initial perspective

In the previous section the distribution of idiomatic expressions and particles in verb cluster and restructuring constructions have been discussed from the perspective of a head-final base structure. Under a head-initial structure, the questions are of a different nature. The diagram in (61) illustrates again the possible and impossible word orders (for examples see the previous section). The first point to note is that in this system, all VP material (including idiomatic phrases and particles) must undergo leftward movement. Second, the height of this movement depends on the kind of verbal element above the main VP. In an auxiliary-participle construction, particles can target a position below the auxiliary, but idiomatic phrases must move above *have*. Lastly, in this approach, *try* and *decide* infinitives must also be distinguished regarding the size of their complements: the complement of a *decide* infinitive must include the highest projection that can be targeted by idiomatic expressions and particles (further movement of these elements across *decide* is impossible), whereas the complement of a *try* infinitive can be a smaller and transparent complement which allows movement of idiomatic expressions and particles across *try*.

(61) Head-initial base structure





The last property is typically taken care of by an account of restructuring and is set aside here. The other two properties (obligatory movement of VP-internal material and the height of this movement) have been addressed in the works on verb clusters, and I summarize a representative account, Zwart (1996), which offers an analysis of the distribution of verbs and particles in Dutch verb clusters.

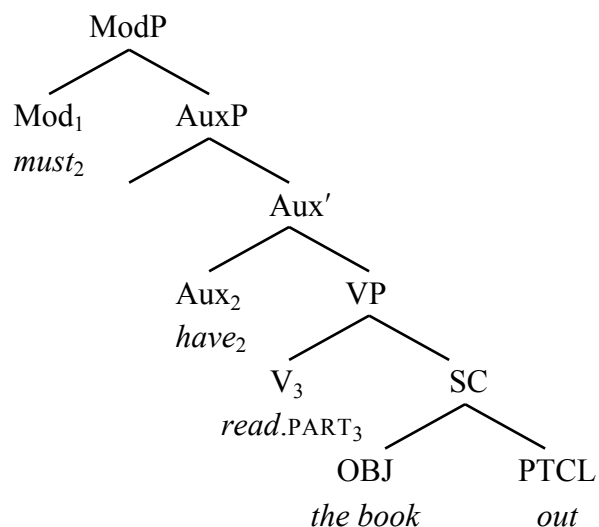
The constructions Zwart concentrates on and provides an analysis for are the auxiliary-participle construction and the modal-auxiliary-participle construction (i.e., the constructions showing the greatest variability in Dutch; see Table 2). The examples in (62) through (63) summarize the distribution of these verb clusters, including the (im)possible positions of particles. As is shown in (62), the order between the two verbs in an auxiliary-participle construction is free and particles can occur either (directly) before or after the auxiliary. In modal-auxiliary-participle constructions, the unmarked orders are 1-2-3 and 3-1-2 as in (63) (some speakers also accept the 1-3-2 order), and the generalizations regarding particle placement are: i) the particle has to follow the object; and ii) the particle has to precede the verb it is associated with.

- (62) a. dat Jan {\*uit} het boek {uit} heeft<sub>1</sub> {uit} gelezen<sub>2</sub> 1-2  
 that Jan {\*out} the book {out} has<sub>1</sub> {out} read<sub>2</sub>  
 ‘that Jan has read/finished the book’
- b. dat Jan {\*uit} het boek {uit} gelezen<sub>2</sub> {\*uit} heeft<sub>1</sub> 2-1  
 that Jan {\*out} the book {out} read<sub>2</sub> {\*out} has<sub>1</sub>  
 ‘that Jan has read/finished the book’
- (63) a. dat Jan {\*uit} het boek {uit} moet<sub>1</sub> {uit} hebben<sub>2</sub> {uit} gelezen<sub>3</sub> 1-2-3  
 that Jan {\*out} the book {out} must<sub>1</sub> {out} have<sub>2</sub> {out} read<sub>3</sub>  
 ‘that Jan must have read/finished the book’
- b. dat Jan {\*uit} het boek {uit} gelezen<sub>3</sub> {\*uit} moet<sub>1</sub> {\*uit} hebben<sub>2</sub> 3-1-2  
 that Jan {\*out} the book {out} read<sub>3</sub> {\*out} must<sub>1</sub> {\*out} have<sub>2</sub>  
 ‘that Jan must have read/finished the book’
- c. %dat Jan het boek {uit} moet<sub>1</sub> {uit} gelezen<sub>3</sub> {\*uit} hebben<sub>2</sub> 1-3-2  
 that Jan the book {out} must<sub>1</sub> {out} read<sub>3</sub> {\*out} have<sub>2</sub>  
 ‘that Jan must have read/finished the book’ [M. den Dikken p.c.]<sup>26</sup>

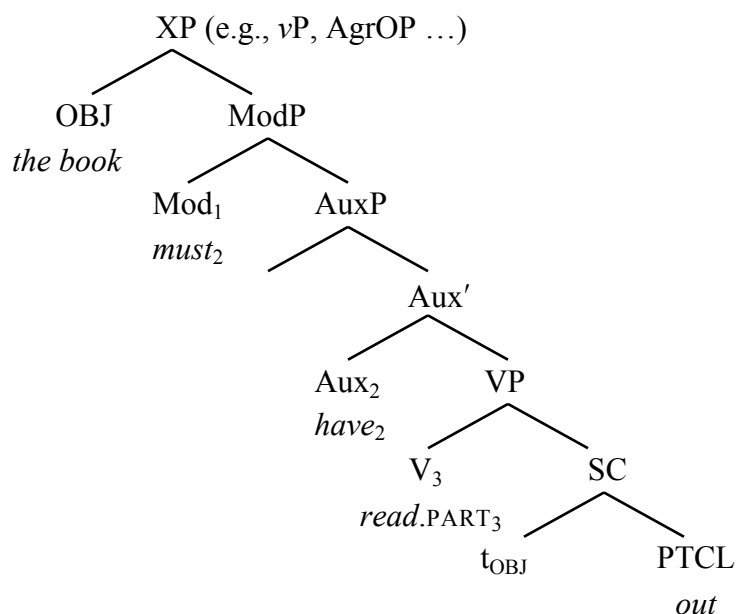
The core assumption in Zwart’s (1996) approach is that—starting from a basic structure as in (64a)—a number of licensing operations (i.e., movements to specifier positions of functional projections) have to apply (Zwart pursues a feature checking approach). First, the object obligatorily moves to an object licensing position (e.g., Spec,AgrOP), which is the highest licensing position in the cluster (i.e., it is above the licensing positions of particles, participles, but lower than the licensing position of the subject; cf. (64b)). This is a fairly straightforward assumption, except that (in contrast to what is generally assumed for object shift in Germanic) this movement has to apply obligatorily for all objects (i.e., it is not sensitive to notions such as definiteness, old vs. new information etc.).

<sup>26</sup> The 1-3-2 order is mentioned in Zwart (1996); the example in (63c) is provided by M. den Dikken. It is not clear if all speakers accept the two particle positions as given in this example.

(64) a. Base structure



b. Object movement

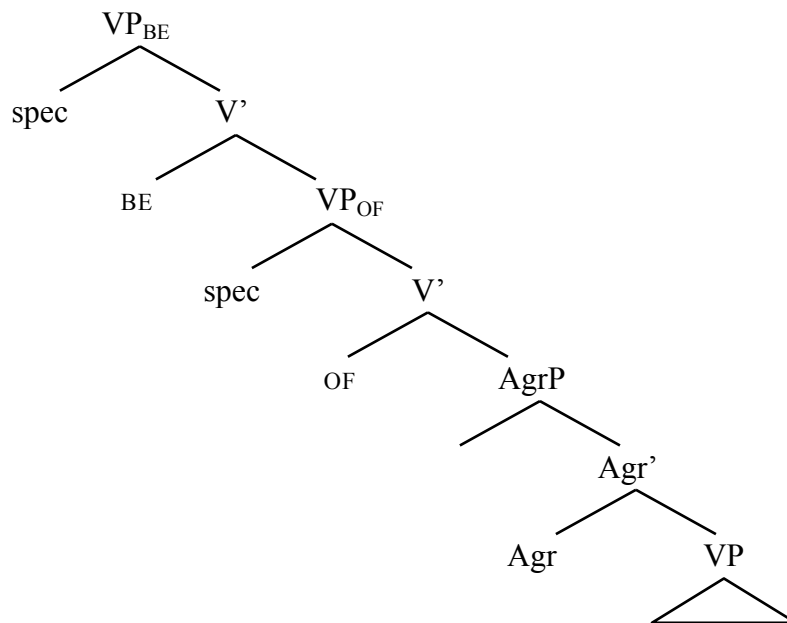


The second and most innovative licensing operation is the one that applies to participles. According to Zwart, participles also move obligatorily, however, they can move to one of two positions in Dutch. Following Kayne (1993), Zwart assumes that the auxiliary *have* is composed of two heads—OF and BE as in (65a). The complement of OF is a small clause consisting of an AGR phrase which in turn selects a lexical projection (either an NP as in *he has a book*, or a VP as in *he has read a book*). The derivation proceeds as follows. First, OF incorporates into BE and the complex BE+OF is pronounced as *have*. Second, the (remnant) VP (i.e., the participle) moves to either the specifier of OF or the specifier of BE (in Dutch).<sup>27</sup> Assuming that *have* is pronounced in

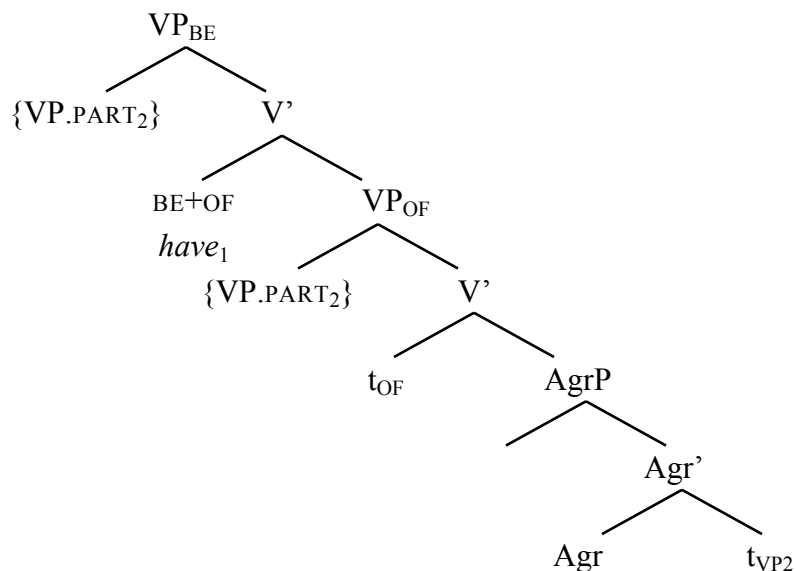
<sup>27</sup> This is not an entirely correct description of Zwart's analysis since Zwart leaves open (p. 249) the phrase structure status of the participle (i.e., whether the element that moves to Spec,OF/BE is a head or a phrase). The reason is

the position of BE, the two orders in (62) are then not the result of movement vs. non-movement of participles, but rather of the two different landing sites for participle phrases as in (65b).

(65) a. Complex *have*



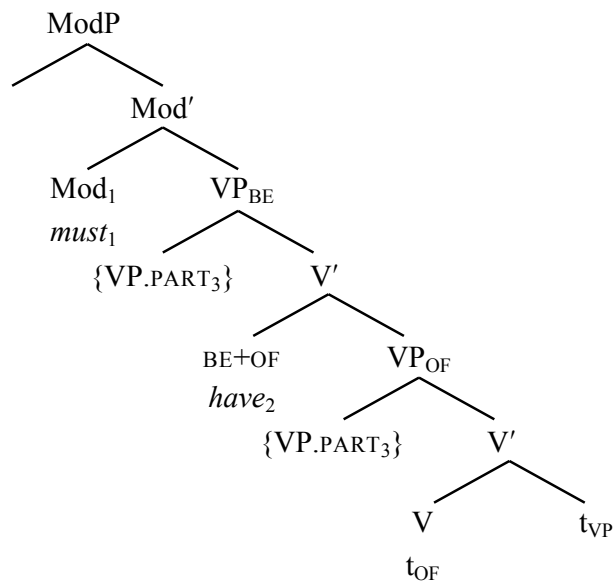
b. Participle movement



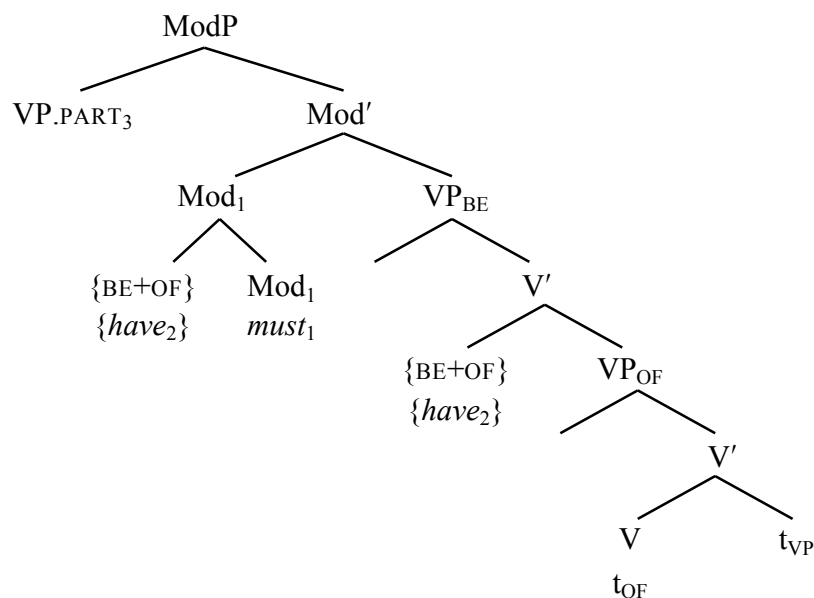
that clausal complements are not carried along when the participle moves to its licensing position. However, this problem can be avoided (by keeping the assumption that participle movement is XP-movement) if the clausal complement undergoes prior movement to a position above the VP but below the landing site of the participle (see Hinterhölzl 1997, 1998, 1999 and Koopman and Szabolcsi 2000 for extensive use of this option). Since this remnant movement approach seems to be more in the spirit of Zwart's proposal than the rather ad hoc assumption that heads move into specifier positions, the participle movement will be represented as VP movement, with the understanding that additional movements might have occurred that pre-empted the VP.

This analysis carries over straightforwardly to the 1-2-3 and 1-3-2 orders in modal-auxiliary-participle constructions illustrated in (66a) (AgrP is ignored from now on since, it does not play a crucial role in the analysis). However, additional assumptions are necessary to account for the 3-1-2 order and to exclude the 2-1-3 and 2-3-1 orders (which do not occur for this construction in any of the West Germanic languages/dialects; see Table 2). Regarding the 3-1-2 order, Zwart suggests that infinitives overtly or covertly adjoin to the selecting modal verbs, and that this mechanism of adjunction transfers the licensing ability of the adjoining element to the host head. Thus, in (66b), the infinitival auxiliary—the BE+OF complex—adjoins to the modal and hence the modal becomes a licenser for participles. In Dutch, infinitive incorporation is covert, whereas in a language like German, it is overt (but see below).

- (66) a. Modal-auxiliary-participle construction: 1-2-3 and 1-3-2

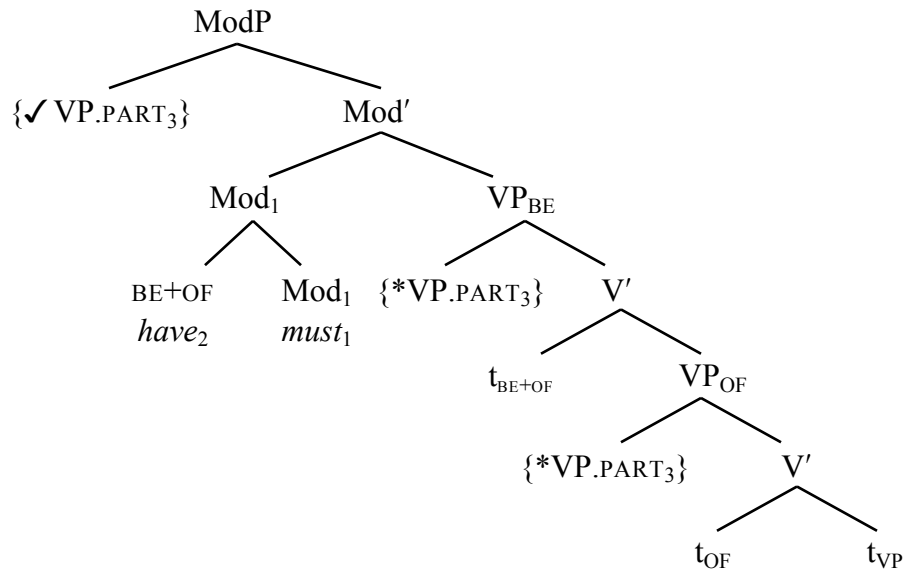


- b. Modal-auxiliary-participle construction: 3-1-2 and 3-2-1



Note that transferring the capacity to license participles to the modal head does not mean that the lower licensing positions are unavailable. Since Zwart treats incorporation of the infinitive into the modal head in (66b) as an obligatory process—i.e., it applies in all the constructions in (63) (but it is necessarily covert in Dutch), the fact that participles can occur in the lower positions in (63a) and (63c) leads to the conclusion that covert incorporation simply increases the number of licensing positions (or in other words, covert incorporation extends the licensing domain but does not force a particular position where licensing has to take place). There is one exception, however, namely as Zwart notes, the lower specifiers become unavailable as licensing positions for participles when the auxiliary incorporates overtly. Thus, in all languages, it is the case that the participle has to precede the auxiliary when the auxiliary (overtly) precedes the modal. The diagram in (67) illustrates the potential structure that is unattested across West Germanic and that is excluded in Zwart’s account by the assumption that the licensing transfer is obligatory and licensing positions are deactivated when overt adjunction takes place (but see the next section for some issues regarding this claim).

(67) Unattested 2-1-3

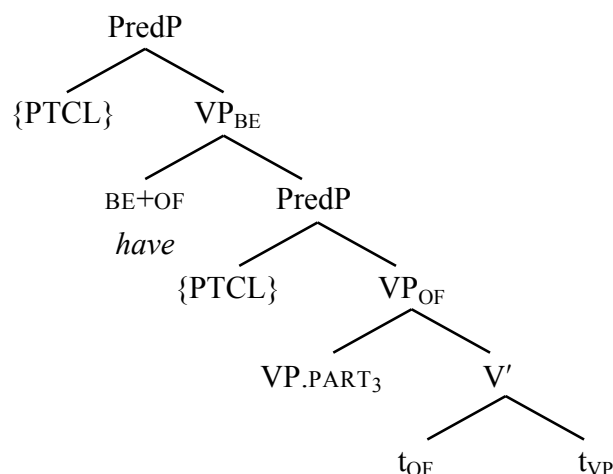


Finally, let us now turn to the third licensing operation—the licensing of particles. In the same spirit as participle licensing, Zwart assumes that particles obligatorily move to a particle licensing position—PredP. The location of PredP is in principle flexible, however, it has to obey the following two constraints: i) it is lower than the object licensing position; and ii) it is higher than the overt position of the verb with which the particle is associated. Zwart suggests that PredP is projected above the projection that hosts the licensing verb. The licensing capacity can again be transferred. If the licensing verb is in a specifier position of an XP (as e.g., the participle in (66)), the licensing capacity is transferred to the head X of that projection. The VP in Spec,VP<sub>OF</sub> in (68a) hence transfers the licensing capacity to (the trace of) OF. Further incorporation of X into higher heads then extends the domain where PredPs can be projected. Thus, incorporation of OF into BE in (68a) allows the projection of PredP above the VP<sub>BE</sub> headed by OF+BE. Similarly, if OF+BE incorporates (overtly or covertly) into a higher head (e.g., a higher modal), PredP would be licensed above the XP hosting this higher head. In contrast, if the verb associated with the particle moves overtly, PredP can only be generated above the XP hosting the (original) licens-

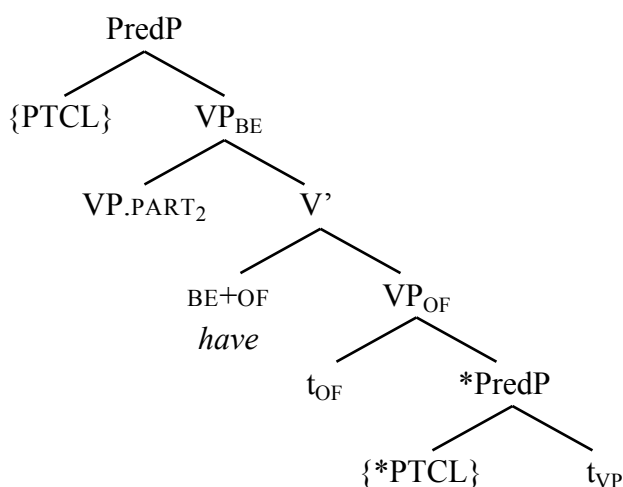
ing verb. As shown in (68b), if the participle occurs in the higher specifier, PredP can only be projected above the VP hosting the participle—the lower position becomes unavailable. These assumptions thus account for the fact that particles can be further up the tree than the (overt position of the) verb they are associated with but not further down.

(68) Particle licensing

- a. Licensing transfer—licensing verb in lower position



- b. Licensing transfer—licensing verb in higher position



While the claim that “the overt position of the verb in which the capacity to license originates [...] determines the range of licensing positions” correctly derives the impossibility of particles following the verb they are associated with, an adjustment is needed for verb second contexts. As shown in (69), particles must remain in sentence final position when the main verb (i.e., the licensing verb) moves to second position. If it is indeed the overt position of the verb that counts, these facts are unexpected.

- (69) a.    dat    Jan    zijn moeder    {op}    belt  
           that    Jan    his mother    {up}    calls  
           ‘Jan calls his mother up’

- b. Jan { \*op } belt zijn moeder { op }  
 Jan { \*up } calls his mother { up }  
 ‘Jan calls his mother up’

To accommodate these facts, it could be assumed that verb second movement is a type of A'-head movement (along the lines of the relativized HMC suggested in den Dikken and Hoekstra 1997) which does not affect the licensing of particles. A different solution would be to assume that particle licensing is not (directly) dependent on the verb the particle is associated with but rather is subject to certain locality conditions on particle movement. In particular, all that seems necessary is to assume that the particle has to be able to c-command its trace in overt syntax (cf. Robbers 1997 for this suggestion). Thus, if the VP containing the trace of the particle moves to a position higher than the landing site of the particle, this condition cannot be met and the structures are excluded. If, however, only the verb moves to a position higher than the particle (such as in (69b)), the particle continues to c-command its trace and the structures are licensed.

To conclude, the attractiveness of Zwart's account is its restrictiveness, on the one hand, and flexibility, on the other hand. The account makes use of two types of movement—head and phrasal movement—and develops general licensing conditions restricting the possible word orders. Variation across languages can be derived from different positions in which certain elements can be licensed as well as overt/covert options for certain movement operations. The detailed ingredients of the account are summarized in Table 6. The table also distinguishes (as far as it is possible to determine) between assumptions or conditions that (presumably) hold universally and those that are language-specific. In the next section, we return to this summary in the context of the directionality debate and compare the head-initial approach with an alternative head-final approach.

Table 6. Standard Dutch verb clusters—Zwart 1996 (adjusted)

Properties	Zwart's approach (revised)
Basic structure	head-initial (universal)
Movement size	head and phrasal movement
Particles	<ul style="list-style-type: none"> <li>licensed in Spec,PredP (universal)</li> <li>PredP is below AgrOP (universal?)</li> <li>licensing has to be overt (West Germanic)</li> <li>Particle has to c-command its trace (universal?)</li> </ul>
Participles	<ul style="list-style-type: none"> <li>licensed in Spec,OF or Spec,BE (universal)</li> <li>licensing has to be overt (universal?)</li> <li>have to be licensed in the specifier position of the VP that includes the overt copy of the auxiliary</li> <li>exception: when <i>have</i> is pronounced in BE, the specifier of OF is still a licensing position</li> </ul>
Infinitives	<ul style="list-style-type: none"> <li>licensed by modal (universal)</li> <li>covert head movement (Dutch)</li> <li>exception: finite two-verb clusters (overt movement possible)</li> </ul>
Objects	<ul style="list-style-type: none"> <li>licensed in Spec,AgrOP (universal)</li> <li>licensing has to be overt (OV languages)</li> <li>AgrOP is higher than all VPs/ModPs etc. (Dutch)</li> </ul>

### 3.2.2 *The directionality debate*

For a significant amount of time, a central question in many works on West Germanic verb clusters has been the question of whether a head-initial or head-final approach can be considered as superior (for works that directly address the issue of directionality see for instance Kaan 1992a, b, Koster 1994, den Dikken 1994, 1995b, 1996, den Dikken and Hoekstra 1997, Zwart 1996, Robbers 1997, Haegeman 1998b, Koopman and Szabolcsi 2000, Ackema 2004). Although in recent years the questions and interests have somewhat shifted (see section 3.1), the controversy can still be considered alive, as can be seen, for instance, in the approach developed in Haider (2003 et seq.) where it is proposed that directionality plays the key role in determining whether a language allows or disallows verb cluster effects.

As the illustration of the various options to derive verb cluster configurations above has shown, the answer to the directionality question does not lie in the success vs. failure of the (mechanical) derivation of verb cluster patterns—both a head-final and a head-initial approach provide the tools to create the correct word orders by movement. Rather, one has to look for an answer by considering questions of extendibility (does either system fare better in extending to other constructions), predictability (does either system fare better in predicting which pattern are possible/impossible), and restrictiveness (does either system fare better in accounting for what is impossible). The works cited above claim to have positive answers to (some of) these questions. Den Dikken (1994, 1995b, 1996) investigates the scope properties in West Flemish verb clusters and argues for the superiority of a head-initial approach. Zwart (1996) discusses the distribution of verb clusters in a number of West Germanic languages and dialects and concludes that the variation found across West Germanic is best accounted for in a head-initial approach. The same conclusion is reached in Robbers (1997) for Afrikaans and in Haegeman (1998b) for West Flemish. Since Zwart's analysis has a broad empirical coverage which has already been outlined in the previous section, and the directionality issue is given as one of the foremost goals in his contribution, it will be used here as a 'straw man' for head-initial approaches.

Zwart's contribution is noteworthy in that it not only offers an analysis for the distribution of word orders (including particles) in various languages and constructions but also provides an explicit comparison between a head-final and a head-initial approach that takes into account the two general questions mentioned above. The critique of head-final approaches raised in Zwart (1996) is given in (70). (Note that Zwart does not present (70) as a critique of one particular analysis but rather as general weaknesses of head-final approaches. Thus, not all points necessarily apply to all head-final analyses.)

#### (70) Critique of head-final approaches

[Zwart 1996: 238f]

- a. There is no consistent direction of adjunction, neither across Continental West Germanic dialects, nor even within particular Continental West Germanic dialects
- b. There is no fixed phrase structure level of the category adjoined
- c. It is not clear what triggers the various movements, in the sense that there is no understanding how particular asymmetries are to be explained (for instance the asymmetry between infinitives and participles, the former adjoining to the left and the latter to the right in a number of dialects (e.g. Luxemburgish))
- d. The parametrization makes no reference to the *timing* of the movements (i.e. either in overt syntax or in covert syntax), which is generally considered to be a major source of parametric variation (see most recently Chomsky 1993)



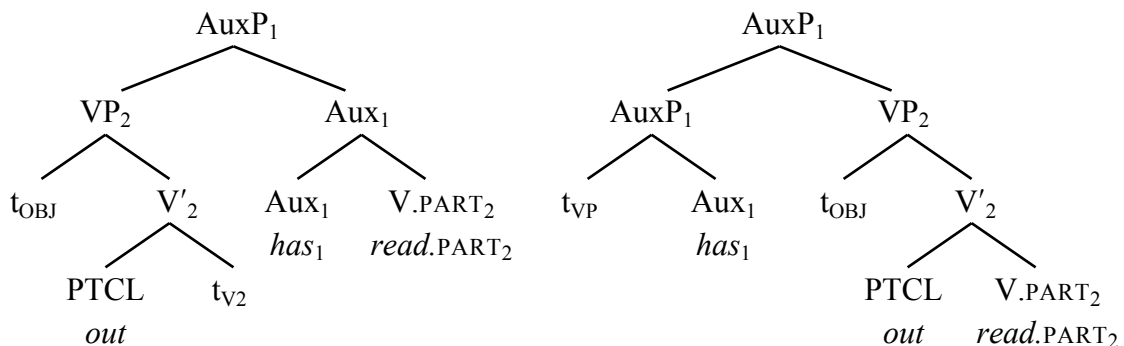
- e. It is unclear why in the Germanic SOV languages multi verb constructions show such a variety of word orders within the cluster, whereas in the Germanic SVO languages multi verb constructions invariably show strictly ‘ascending’ orders
- f. It is unclear why certain phenomena (the IPP effect, Verb Projection Raising) are sensitive to the surface order of the members of the verb cluster
- g. The analysis relies on a phrase structural split among the Germanic languages for which there is no independent empirical basis [...]

In what follows, I outline a head-final analysis for the same set of constructions as discussed in Zwart (1996) and compare the two approaches. The conclusion will be that certain questions pointed out in (70) are important challenges, however, not as Zwart concludes solely for head-final approaches but rather for accounts of verb clusters in general.

The assumptions necessary to derive the (Standard) Dutch verb clusters discussed in Zwart (1996) are: i) particles do not move (or, following the discussion in section 3.2.1.1, only move string-vacuously); ii) participles move optionally to the right; iii) infinitives move obligatorily to the right; iv) verb cluster movement can affect either minimal or maximal categories and is subject to a strict locality condition; and v) objects move obligatorily to a specifier position. Point v) assures that objects are not part of the verb cluster in Dutch (the position of objects will be ignored in the diagrams below). The examples in (62) (repeated here as (71)), are derived as follows. In (71b), no movement occurs; the elements appear in their base positions. The two options in (71a) are derived as in (72): the object leaves the VP; and then either the lower V2 (as in (72a)) or the lower VP2 (as in (72b)) moves to the right. Both movement operations yield the 1-2 order, however, in the first derivation, the particle is stranded (i.e., occurs to the left of the auxiliary), whereas it is carried along with the verb in the second derivation. The impossible positions of the particles in (71), follow from assumption i)—particles do not move (neither to the right nor to the left).

- (71) a. dat Jan {\*uit} het boek {uit} heeft<sub>1</sub> {uit} gelezen<sub>2</sub> 1-2  
           that Jan {\*out} the book {out} has<sub>1</sub> {out} read<sub>2</sub>  
           ‘that Jan has read/finished the book’
- b. dat Jan {\*uit} het boek {uit} gelezen<sub>2</sub> {\*uit} heeft<sub>1</sub> 2-1  
           that Jan {\*out} the book {out} read<sub>2</sub> {\*out} has<sub>1</sub>  
           ‘that Jan has read/finished the book’

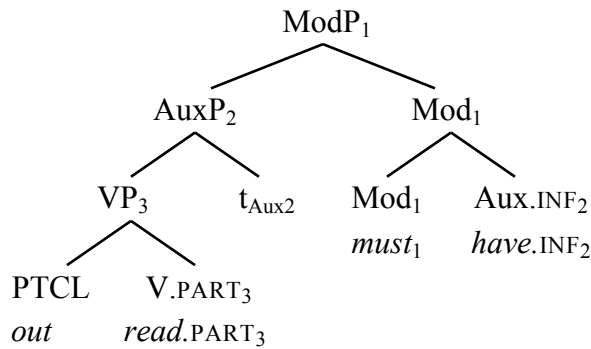
- (72) a. Optional head movement of PART      b. Optional phrasal movement of PART



The examples involving modal-auxiliary-participle constructions are repeated in (73) and (74). The 3-1-2 order is straightforwardly derived by head movement of the infinitive. Given that particles do not move (or only string-vacuously), it also follows that the only position the particle can occupy is immediately preceding the lowest verb as in (73a'). Note that movement of V2 in this case is obligatory in Dutch, since V2 is an infinitive (see also Table 3). Thus the order 3-2-1 is not possible. Phrasal movement of the infinitive yields the 1-3-2 order in (73b'), which is attested in certain dialects.<sup>28</sup>

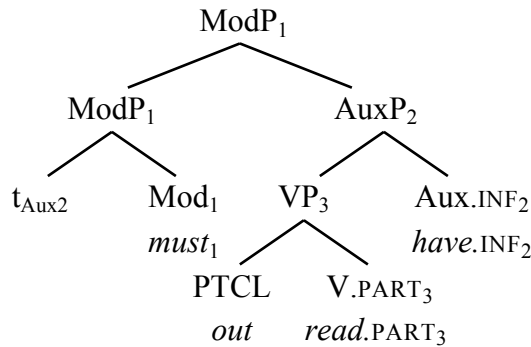
- (73) a. dat Jan {\*uit} het boek {uit} gelezen<sub>3</sub> {\*uit} moet<sub>1</sub> {\*uit} hebben<sub>2</sub> 3-1-2  
 that Jan {\*out} the book {out} read<sub>3</sub> {\*out} must<sub>1</sub> {\*out} have<sub>2</sub>  
 'that Jan must have read/finished the book'

a.' PTCL-3-1-2: head movement of INF



- b. %dat Jan het boek {uit} moet<sub>1</sub> {uit} gelezen<sub>3</sub> {\*uit} hebben<sub>2</sub> 1-3-2  
 that Jan the book {out} must<sub>1</sub> {out} read<sub>3</sub> {\*out} have<sub>2</sub>  
 'that Jan must have read/finished the book' [M. den Dikken p.c.]

b.' 1-PTCL-3-2: phrasal movement of INF



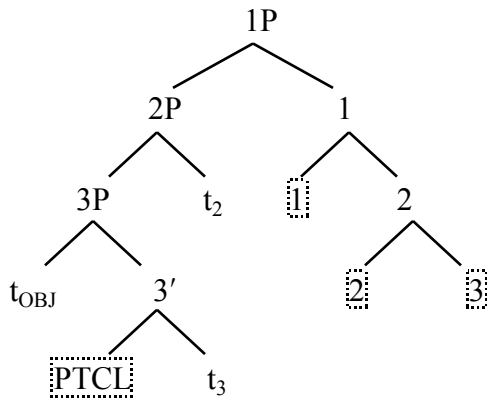
Finally, the 1-2-3 construction in (74a) allows three options for particles. There are four possible derivations (I assume only cyclic derivations, i.e., the lowest V(P)3 moves before the higher V(P)2): i) two applications of head movement, which yields the order PTCL-1-2-3 in (74b); ii)

<sup>28</sup> As noted in fn. 26, it is not entirely clear where particles can go in this order. If speakers of this dialect share M. den Dikken's intuition that particles can occur to the right or to the left of V1 in the 1-3-2 order, one of the options for particles as discussed in the text would have to be eliminated—i.e., particles could then not be considered as immobile but one would have to pursue an analysis according to which particles can undergo string-vacuous movement (see section 3.2.1.1). I leave this issue open here.

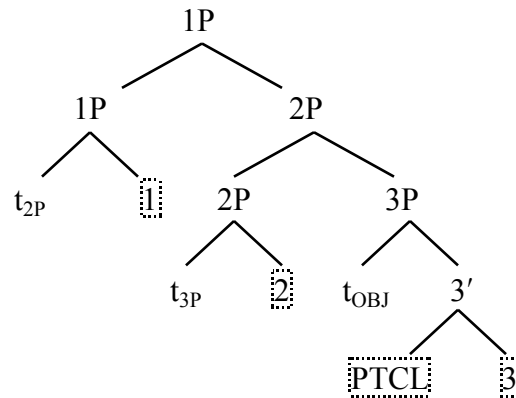
two applications of phrasal movement, which yields the order 1-2-PTCL-3 in (74c); iii) head movement of the lower V3 and phrasal movement of the higher VP2, which yields the order 1-PTCL-2-3 in (74d); and iv) phrasal movement of the lower VP3 and head movement of the higher V2, which yields the order PTCL-3-1-2 in (74e). Thus, the 3-1-2 order in (73a) can be derived in two ways, which may be supported by the fact that in this construction, the 3-1-2 order is the most frequent order (see the frequencies reported by the SAND project, noted under Table 2).<sup>29</sup>

- (74) a. dat Jan {\*uit} het boek {uit} moet<sub>1</sub> {uit} hebben<sub>2</sub> {uit} gelezen<sub>3</sub> 1-2-3  
 that Jan {\*out} the book {out} must<sub>1</sub> {out} have<sub>2</sub> {out} read<sub>3</sub>  
 ‘that Jan must have read/finished the book’

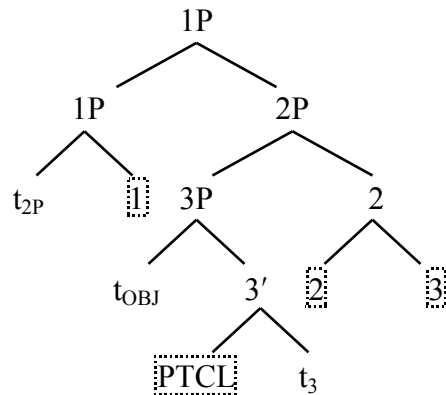
b. 2x head movement



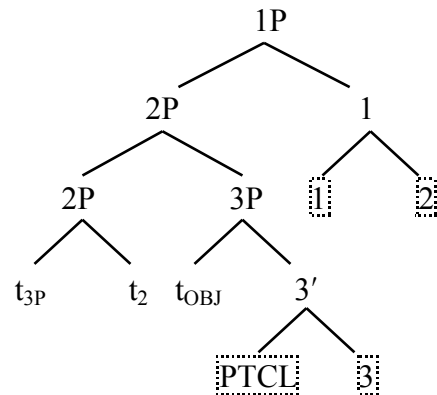
c. 2x phrasal movement



d. head, then phrasal movement



e. phrasal, then head movement



Importantly, under this analysis, the direction of adjunction is uniform (i.e., to the right). Thus, Zwart's (1996) first point of criticism in (70a) does not seem to be justified (furthermore, as we have seen in 3.2.1, all other orders can also be derived by uniform right-adjunction in a head-final approach if reordering involves phrasal movement). To discuss the other points in (70), let

<sup>29</sup> Since the 3-1-2 order is excluded in Standard Dutch in double infinitive constructions (see Table 2), mixed movements as in (74d,e) should presumably be restricted to constructions with two different types of movement (participle and infinitive movement involve different operations, but the application of two infinitive movements involve the same operation).

us now compare the two approaches. The assumptions of the head-final approach are summarized in Table 7, together with the assumptions of Zwart’s account, repeated from Table 6.

Table 7. Dutch verb clusters—head-final vs. head-initial approach

Properties	Zwart’s approach (revised)	Head-final approach
Basic structure	head-initial (universal)	directionality
Movement size	head and phrasal movement	head and phrasal movement
Particles	<ul style="list-style-type: none"> <li>licensed in Spec,PredP (universal)</li> <li>PredP is below AgrOP (universal?)</li> <li>licensing has to be overt (West Germanic)</li> <li>Particle has to c-command its trace (universal?)</li> </ul>	do not move or only move string-vacuously
Participles	<ul style="list-style-type: none"> <li>licensed in Spec,OF or Spec,BE (universal)</li> <li>licensing has to be overt (universal?)</li> <li>have to be licensed in the specifier position of the VP that includes the overt copy of the auxiliary</li> <li>exception: when <i>have</i> is pronounced in BE, the specifier of OF is still a licensing position</li> </ul>	move optionally to the right
Infinitives	<ul style="list-style-type: none"> <li>licensed by modal (universal)</li> <li>covert head movement (Dutch)</li> <li>exception: finite two-verb clusters (overt movement possible)</li> </ul>	<ul style="list-style-type: none"> <li>move obligatorily to the right</li> <li>exception: finite two-verb clusters</li> </ul>
Objects	<ul style="list-style-type: none"> <li>licensed in Spec,AgrOP (universal)</li> <li>licensing has to be overt (OV languages)</li> <li>AgrOP is higher than all VPs/ModPs etc. (Dutch)</li> </ul>	<ul style="list-style-type: none"> <li>move obligatorily (e.g., Spec,AgrOP)</li> </ul>
*2-1-3 (MOD-AUX-V)	<ul style="list-style-type: none"> <li>licensing conditions of participles</li> </ul>	<ul style="list-style-type: none"> <li>locality: closest target for 3(P)</li> </ul>
*2-1-3 (universal?)	<ul style="list-style-type: none"> <li>all movement within a cluster either overt or covert (no overt/covert mix)</li> </ul>	<ul style="list-style-type: none"> <li>locality: closest target for 3(P)</li> </ul>
*2-3-1 (Standard Dutch)	<ul style="list-style-type: none"> <li>no XP-movement of INF (Dutch); IPP = INF</li> </ul>	<ul style="list-style-type: none"> <li>movement of INF is obligatory</li> </ul>

As can be seen in the table, the two approaches both make use of head and phrasal movement,<sup>30</sup> and both involve obligatory object movement. Hence, on these points, the accounts do not differ. Similarly, both accounts require a stipulation regarding finite two-verb clusters and hence this point also does not enter into the comparison (however, see Koopman and Szabolcsi 2000 for an analysis that incorporates this fact). Regarding the other assumptions, a comparison is not an easy task (if possible at all), in particular, since the analyses are based on underlying assumptions that are not necessarily shared by the proponents of the different approaches. One such assumption is the claim that structures have to be compatible with Kayne (1994’s) LCA, which would prohibit head-final structures and rightward movement. The fact that Zwart’s analysis conforms to the LCA, whereas head-final approaches obviously conflict with this view, does not force the conclusion that head-final approaches should be rejected. What Zwart’s account has demonstrat-

<sup>30</sup> One of Zwart’s critiques of the head-final approach which is mentioned repeatedly in the article is that both head and phrasal movement are necessary to derive verb clusters. Although the account presented in the text indeed involves both head and phrasal movement, it is important to note that this is not necessarily the only way to derive Dutch verb clusters under a head-final approach. It has been chosen here since it appears to be the most straightforward option. However, as we have seen in the preceding sections, it is possible—under certain assumptions about particles—to derive the word order patterns exclusively by phrasal movement in the head-final approach. Thus, Zwart’s claim is not entirely correct. Furthermore, it should be noted that Zwart’s own analysis crucially involves both head and phrasal movement. Thus, it seems that this point of criticism is not justified.

ed is that the facts can be described within a system invoking the LCA. If one is independently committed to that framework, then Zwart's approach shows the range of ancillary assumptions necessary to derive the facts. However, if one is not independently persuaded by the LCA (see for instance the works on scope by Büring and Hartmann 1997a, b and Fox and Nissenbaum 1999 which suggest that the LCA might be too strong), nothing in the account of verb clusters bears on the matter. This is most obvious in the revival of directionality proposed in Haider (2003 et seq.). The goal of Haider's approach is to predict which languages allow verb clusters, and the head-final nature of the West Germanic languages is assumed to be the driving factor. Thus, if Haider's approach can be maintained, it is exactly the directionality that provides an answer to Zwart's question (70e), which, incidentally, neither Zwart's own nor the alternative head-final approach I outlined above have anything to say about.

Similarly, the claim that the uniformity of structures presents an advantage under a head-initial approach has to be balanced against the non-uniformity of overt vs. covert settings (cf. questions (70d,g)). Under an approach involving directionality, languages have to set the directionality parameter but do not have to involve any settings for instance for particles (particles do not move, and hence they occur to the right of the verb in English, whereas they occur to the left of the verb in German and Dutch). Under an approach involving a uniform base, languages do not have to set any directionality parameter but they have to involve an overt vs. covert setting for particles, objects and many more. Since the overt/covert settings are arbitrary, the directionality differences found between head-final and head-initial languages are as stipulative in the head-initial approach as they are in the head-final approach (i.e., there is no deep reason for the claim that movement is overt in German and Dutch, whereas it is covert in English—as there is no reason for the claim that German and Dutch are head-final). Thus, while the idea of a uniform base structure is certainly attractive, it does not solve the problem that languages are different, it simply shifts the burden of explanation to a different area. In other words, at our current stage of understanding, under both the head-initial and the head-final approach, directionality differences among languages have to be represented as irreducible facts about languages that cannot be further derived but simply must be learned.

Let us finally turn to the criteria mentioned at the beginning of this section—extendibility, predictability, and restrictiveness. Although Zwart's (1996) system is focused on certain verb clusters in Standard Dutch, it is easy to see how the system can be extended to other varieties of Dutch and other languages. For instance, to derive the 2-3-1 order in IPP constructions, all that is needed is to assume that an IPP complement can or must (depending on the dialect) qualify as a participle, which would correctly trigger movement of the ModP (2P) to the specifier of OF+BE. Similarly, to derive the variation found in Dutch dialects in infinitival constructions it would be sufficient to assume that movement of infinitives can either be overt or covert. Head-final approaches, too, have the capacity to derive the variation, as we have seen already in section 2.5 where inversion rules were given for all languages, which could be encoded as part of the movement rules.

Regarding predictability and restrictiveness, let us consider the observation that verb clusters, cross-linguistically, do not occur in the 2-1-3 order (as we have seen in section 3.1, this seems to be correct for standard cluster elements, modals and auxiliaries, but requires setting aside certain other constructions in Swiss German; see Salzmann 2013a, b). The options presented by the two accounts have been added in Table 7. As discussed in the previous section and summarized in (75a), Zwart (1996) assumes that infinitives undergo covert head movement to MOD and participles undergo phrasal movement to Sepc, OFP, Spec, BEP or Spec, MODP. The latter,

can be pronounced in either position, which derives the three possible orders in (75b) for Standard Dutch. If, on the other hand, infinitive movement is overt as in (75c), only the highest copy of the participle phrase can be pronounced.

- (75) a.  $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  Movement of 3P and V2
- b.  $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  V2 movement covert in Dutch  
 $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  Highest copy of 3P: 3-1-2  
 $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  Middle copy of 3P: 1-3-2  
 $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  Lowest copy of 3P: 1-2-3
- c.  $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  V2 movement overt  
 $[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  Highest copy of 3P: 3-2-1  
 $*[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  \*Middle copy of 3P: 2-1-3  
 $*[_{1P} \{3P\} \{2\} -1 \ [_{2P,BE} \{3P\} \{2\} \ [_{2P,OF} \{3P\} \ t_{3P} ]]]$  \*Lowest copy of 3P: 2-1-3
- d.  $[_{2P,BE} \{3P\} \{2=OF+BE\} \ [_{2P,OF} \{3P\} \{2=OF\} \ t_{3P} ]]$  Higher or lower copy of 3P

The way Zwart excludes the impossible linearizations in (75c) is via the assumption that the lower positions lose their licensing capacity when the head OF+BE undergoes overt movement. Note, however, that this crucially only applies to the auxiliary OF+BE and not to heads in general. As detailed in (75d), to derive AUX-PARTICIPLE orders, the low position (the specifier of OF) is crucially available, even when OF overtly moves to BE. Thus the suspension of the lower licensing positions in (75c) contrasts with the retention of the licensing position in (75d)—in both cases the head of the licensing position has moved away via overt movement, yet one position remains a valid licensing position but the other doesn't. It is not entirely clear whether this difference is motivated beyond the goal to exclude the 2-1-3 order. One may also wonder how the choice of the higher copy of OF+BE in (75c) (arguably a PF property) can affect the licensing potential of heads in syntax.<sup>31</sup> Moreover, the general ban on the 2-1-3 order in verb clusters may also pose a problem in contexts, such as the different varieties of Dutch, where infinitive movement can be overt or covert. What must be excluded, for instance in double infinitive constructions, is covert movement of V3, followed by overt movement of V2. As indicated in Table 7, an additional 'consistency' assumption (either all movement is overt or all movement is covert) is needed, at least for movements of the same type (since, as we have seen, infinitive and participle movement can differ regarding the overt/covert application).

Under the head-final approach sketched above, on the other hand, 2-1-3 orders can be excluded by a simple assumption of locality: movement of 3 must obey the HMC, and movement of 3P could be assumed to be strictly local in that it must target the closest XP above it and cannot skip a projection (see, among others, Haegeman 1992). Note that this approach would also allow to distinguish between verb cluster movement (\*2-1-3) and extraposition (✓2-1-3), if the latter is not subject to this strict requirement (which would be advantageous in light of the differences between VPR and extraposition noted in Salzmann 2013a, b, 2014).

<sup>31</sup> As will be shown in section 3.3.2, Barbiers (2005) offers an analysis of the impossibility of the 2-1-3 order under a head-initial approach. However, crucially, Barbiers assumes that all verb cluster movements are optional, essentially giving up the goal to syntactically derive the word order variation, by relegating the differences in the cross-linguistic distribution of verb cluster orders to language/construction-specific assumptions about whether a particular movement applies or doesn't apply (but syntactically, movement is never forced). This does, in a sense, bring Barbiers' account much closer to the non-movement accounts discussed in section 3.1.

### 3.2.3 Verb projection raising

(76) a. da Valère wilt<sub>1</sub> vele boeken lezen<sub>2</sub> West Flemish  
 that Valère wants<sub>1</sub> many books read<sub>2</sub>  
 ‘that Valère wants to read many books’ [Haegeman 1998b:261]

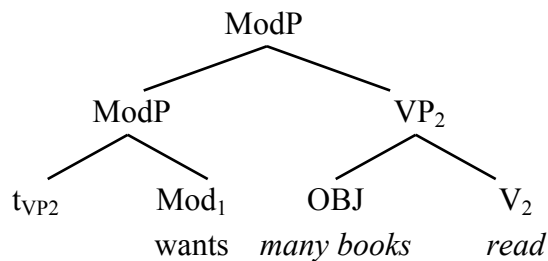
71

- b. da Jan ... {da boek} wilt<sub>1</sub> {da boek} vuor Marie kopen<sub>2</sub>  
 that Jan ... {that book} wants<sub>1</sub> {that book} for Marie buy<sub>2</sub>  
 ‘that Jan wants to buy that book for Marie’ [Haegeman 1992:148]
- c. as hulle daar moet<sub>1</sub> goeie onderwys gee<sub>2</sub> Afrikaans  
 if they there must<sub>1</sub> good education give<sub>2</sub>  
 ‘if they must provide a good education there’ [Robbers 1997:76]
- d. das si am Grendel wöt<sub>1</sub> sini velore chlaue zruggeh<sub>2</sub> Swiss German  
 that she to.the Grendel wants<sub>1</sub> his lost paw back.give<sub>2</sub>  
 ‘that she wants to return the lost paw to Grendel’

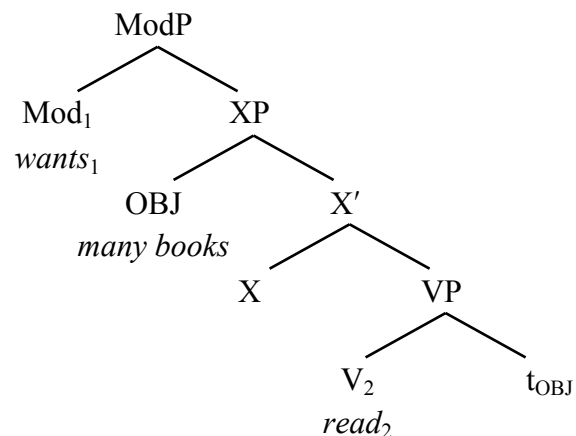
To derive these configurations under the head-final approach, the entire lower VP moves to the right of the modal verb (i.e., the object remains inside the VP as in (77a)). Equivalently, it could be assumed (see for instance Vanden Wyngaerd 1989) that the object moves to a functional projection outside the VP (e.g., Spec,AgrOP) and that this functional projection rather than the VP undergoes movement to the right. To derive the examples in (76) under a head-initial approach, no verb or verb phrase movement takes place; the only movement operation necessary is object movement to a functional projection between the modal and the main verb (as in (77b)).

(77) Verb Projection Raising: 1 - [...2]

- a. Head-final base



- b. Head-initial base



Note that VPR is not an obligatory phenomenon (except for certain elements; see below)—in all of the above examples, the objects can also precede the higher verb. For the head-final approach, this means that the object can move out of the VP prior to VP-movement in (77a) (or under the



AgrOP-movement analysis, one has to allow for either AgrOP movement or VP movement). For the head-initial approach, these facts indicate that the object can move (further) to a functional projection above the higher verb. Thus, either approach has to involve some notion of optionality—in head-final approaches, optionality comes in as a choice between overt and covert movement of the object or a choice between two (or more) categories that can undergo VPR; in head-initial approaches, optionality comes in as an option for the ordering of functional projections (i.e., whether XP is below or above ModP in (77b)).

Looking across the West Germanic languages and dialects, there are some interesting restrictions and cross-linguistic generalizations regarding what elements can, cannot or must intervene between the verbs of a verb cluster. First, VPR is not possible in all West Germanic languages/dialects. Examples such as the ones in (76), for instance, are ungrammatical in Dutch. Although Dutch allows the 1-2 order in these clusters, the object has to precede the higher verb. Second, we find that even among the languages that do allow VPR, the types of elements that can occur inside a cluster are not the same across these languages. Although the distribution shows some degree of variation, there is also a clear generalization that emerges. As can be seen in the summary in Table 8 (illustrations will be provided below), the ‘bigger’ or the more independent an element is, the less likely it is to be licensed between the verbs of a verb cluster. Furthermore, the elements in Table 8 seem to stand in an implicational relation: if a language allows VPR with elements of a particular category in Table 8, it also allows VPR with the elements to its left. In other words, if ‘bigger’ elements can occur in a cluster, all ‘smaller’ elements are also allowed to occur between the verbs of a cluster.

Table 8. Verb projection raising (with modals, auxiliaries)

Language	Separable particles	Low adverbs, idioms, bare Ns	Indefinite objects, PPs	Definite objects
Afrikaans	Obligatory <sup>a</sup>	Possible	%Possible	Impossible
Dutch	Possible	Marginal <sup>b</sup>	Impossible	Impossible
German	Obligatory <sup>a</sup>	Possible	%Possible	Impossible
Swiss	Obligatory <sup>a</sup>	Possible	Possible	Possible
West Flemish	Possible	Possible	Possible	Possible

Notes:

<sup>a</sup> obligatory with ‘idiomatic’ particles (cf. Robbers 1997 citing Le Roux 1989); ‘transparent’ particles are marginal in stranded position in Swiss German and German (Wurmbrand 1999)

<sup>b</sup> only possible if element is morphologically not complex (cf. Neeleman 1994)

The implicational hierarchy in Table 8 is reflected directly in the results of the corpus study provided in Dubenion-Smith (2010). As pointed out under Table 1, the corpus used in this study is based on transcripts of 187 dialect recordings of speakers from West Central German dialects compiled as part of the Zwirner corpus between 1955-1970. In the recorded embedded clauses with 1-2 verb clusters, 12.6% of the utterances involved VPR, and the elements part of the verb cluster distribute as follows: separable particles (55%), adjectives/adverbs (20%), PPs (15%), and indefinite NPs (10%).

Let us begin with the distribution of particles which are considered the smallest category. The claim that these elements are small finds support in the assumption (which most researchers make in one form or another) that particles form a complex head (whether base generated or de-

rived) at some level with the verb they are associated with. Importantly for the present discussion, as shown in (78), these elements can be part of verb clusters in all languages, and they must be part of the cluster and cannot be stranded in Afrikaans, German, and Swiss German.<sup>33</sup>

- (78) a. Die bende sal ons {\*aan} bly {aan} rand Afrikaans  
 The gang will us {\*on} remain {on} assault  
 ‘The gang will go on assaulting us’ [Robbers 1997: 61, fn. 14]
- b. dat Jan zijn moeder {op} wil {op} bellen Dutch  
 that Jan his mother {up} wants {up} call  
 ‘that Jan wants to call his mother’ [Neeleman 1994: 24]
- c. dass er das Buch {\*durch} hätte {durch} sehen sollen German  
 that he the book {\*through} had {through} look shall  
 ‘that he should have looked through the book’
- d. dass er da Buech {\*doere} het soele {doere} laese Swiss German  
 that he the book {\*through} has shall {through} read  
 ‘that he should have read (through) the book’  
 [M. Schönenberger p.c., E. Haeberli p.c.]
- e. dan ze hem vu dienen cursus {in} moeten {in} schrijven West Flemish  
 that they him for that course {in} must {in} write  
 ‘that they must register him for that course’ [Rutten 1991: 60]

There are two basic ways to capture this distributional difference regarding particle placement. First, it has been assumed that the difference lies in the properties of the particles in the two language groups. Second, it has been assumed that the difference lies in the properties of verb cluster formation. Starting with the head-final approach, particle incorporation could be seen as obligatory in the non-stranding languages and optional (or optionally covert) in the stranding languages. Thus, independent of whether verb clusters are formed by head or phrasal movement, in a language with obligatory overt incorporation of particles, particles will always be carried along with the verb or verb phrase and hence the particle will show up next to the verb it is associated with. According to the second option, the difference between Dutch and West Flemish, on the one hand, and Afrikaans, German, and Swiss German, on the other hand, would lie in a difference in the process of verb cluster formation. If—as discussed in section 3.2.1.1—it is assumed that particles are immobile, the difference could be seen as a contrast between the possibility (Dutch, West Flemish) vs. impossibility (Afrikaans, (Swiss) German) of head movement to derive verb cluster reordering. In the second group of languages, only phrasal movement is allowed and hence stranding of the particles would be prohibited. Finally, under a pure phrasal movement approach to verb cluster formation, it could be assumed that Dutch and West Flemish allow leftward movement of particles prior to VP-movement (and hence stranding of particles when the VP undergoes rightward movement), whereas particles cannot undergo independent movement in Afrikaans, German, and Swiss German.

<sup>33</sup> As mentioned in the notes to Table 8, the prohibition against stranding of particles only holds for idiomatic particles; particle constructions with a transparent meaning readily allow particle stranding in Afrikaans (cf. Robbers 1997), and under the right focus conditions, also in German and Swiss German (see Wurmbbrand 1999, Ackema 2004 for discussions of this distinction). In Dutch and West Flemish, however, both types of particles can be stranded. The statements in the text should thus be seen as referring only to idiomatic particles.

While any of these assumptions successfully derives the different word orders, the ultimate goal would be to relate the different particle positions and ordering restrictions to other properties of the languages under consideration. More can be said about the cross-linguistic differences of particle placement in West Germanic verb clusters, however, for the present discussion the crucial observation is that particles are always allowed to interrupt a verb cluster, but that particle stranding (or leftward migration of particles) is subject to language specific constraints.

(79) a. dat hulle hier kom<sub>2</sub> water drink<sub>3</sub> het<sub>1</sub> Afrikaans  
that they here come<sub>2</sub> water drink<sub>3</sub> have<sub>1</sub>  
'that they have come here to drink water' [Robbers 1997: 75]

a'. Jy sal seker moet<sub>2</sub> vinnig reageer<sub>3</sub>  
You will certainly must<sub>2</sub> quickly reply<sub>3</sub>  
'You will certainly have to reply quickly' [Robbers 1997 :83]

a''. Jy sal dit moet<sub>2</sub> in ag neem<sub>3</sub>  
You will it must<sub>2</sub> into consideration take<sub>3</sub>  
'You will have to take this into consideration' [Robbers 1997: 73]

b. dat Jan de deur {groen} wil<sub>1</sub> {groen} verven<sub>2</sub> Dutch  
that Jan the door {green} wants<sub>1</sub> {green} paint<sub>2</sub>  
'that Jan wants to paint the door green' [Neeleman 1994 :24]

b'. dat Jan de meloen {open} zal<sub>1</sub> {open} snijden<sub>2</sub>  
that Jan the melon {open} will<sub>1</sub> {open} cut<sub>2</sub>  
'that Jan will cut open the melon' [Neeleman 1994 :237]

b''. dat Jan de meloen {helemaal open} zal<sub>1</sub> {\*helemaalopen} snijden<sub>2</sub>  
that Jan the melon {all.the.way open} will<sub>1</sub> {\*all.the.way open} cut<sub>2</sub>  
'that Jan will cut the melon all the way open' [Neeleman 1994 :237]

b'''. dat Jan de bloemen {water} heeft<sub>1</sub> willen<sub>2</sub> {?\*water} geven<sub>3</sub>  
that Jan the flowers {water} has<sub>1</sub> want.IPP<sub>2</sub> {?\*water} give<sub>3</sub>  
'that Jan wanted to water the flowers' [F. Veraart p.c.]

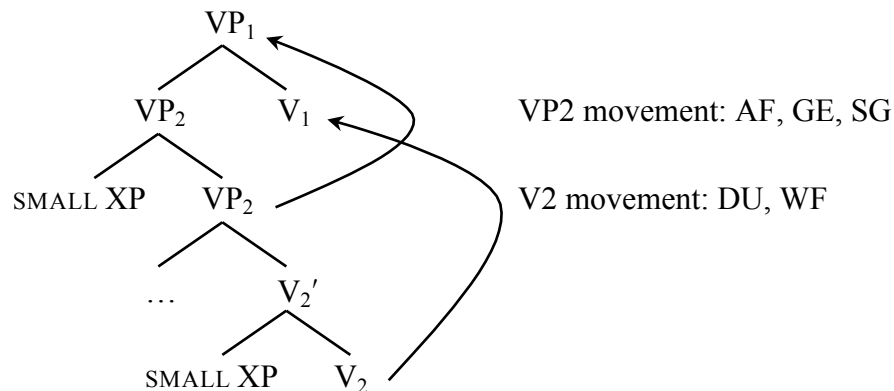
c. dass er das Buch {genau} hätte<sub>1</sub> {genau} durchsehen<sub>3</sub> sollen<sub>2</sub> German  
that he the book {carefully} had<sub>1</sub> {carefully} through.look<sub>3</sub> shall<sub>2</sub>  
'that he should have looked through the book carefully'

- c'. daß er vor der Abreise noch {Blumen} hätte<sub>1</sub> {Blumen} gießen<sub>3</sub> sollen<sub>2</sub>  
 that he before the departure still {flowers} had<sub>1</sub> {flowers} water<sub>3</sub> shall<sub>2</sub>  
 'that he should have watered the flowers before his departure'
- c''. daß er dieses Faktum {in Betracht} hätte<sub>1</sub> {in Betracht} ziehen<sub>3</sub> müssen<sub>2</sub>  
 that he this fact {in consideration} had<sub>1</sub> {in consideration} take<sub>3</sub> must<sub>2</sub>  
 'that he should have taken this fact into consideration'
- d. dass er vor de Abreis no {Blueme} het<sub>1</sub> soele<sub>2</sub> {Blueme} guesse<sub>3</sub> SG  
 that he before the departure still {flowers} had<sub>1</sub> shall<sub>2</sub> {flowers} water<sub>3</sub>  
 'that he should have watered the flowers before his departure'  
 [M. Schönenberger p.c., E. Haeberli p.c.]
- d'. dass er da Buech {gnau} het<sub>1</sub> soele<sub>2</sub> {gnau} doere laese<sub>3</sub>  
 that he the book {carefully} had<sub>1</sub> shall<sub>2</sub> {carefully} through read<sub>3</sub>  
 'that he should have read through the book carefully'  
 [M. Schönenberger p.c., E. Haeberli p.c.]
- d''. dass er das {i Betracht} het<sub>1</sub> soele<sub>2</sub> {i Betracht} zieh<sub>3</sub>  
 that he that {in consideration} had<sub>1</sub> shall<sub>2</sub> {in consideration} take<sub>3</sub>  
 'that he should have taken this into consideration'  
 [M. Schönenberger p.c., E. Haeberli p.c.]
- e. da Jan dat boek absoluut wilt<sub>1</sub> thuis lezen<sub>1</sub> West Flemish  
 that Jan that book absolutely wants<sub>1</sub> at.home read<sub>1</sub>  
 'that Jan absolutely wants to read that book at home' [Haegeman 1992: 190]
- e'. da Valère oa<sub>1</sub> willen<sub>2</sub> morgen no Gent goan<sub>3</sub>  
 that Valère had<sub>1</sub> want.IPP<sub>2</sub> tomorrow to Gent go<sub>3</sub>  
 'that Valère had wanted to go to Gent tomorrow' [Haegeman 1995 :72]

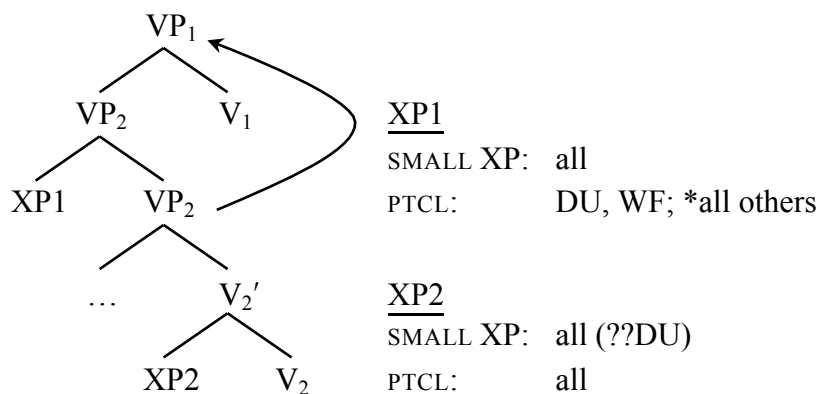
The fact that the elements in (79) behave like particles in that they are allowed to split up a verb cluster (across West Germanic) could be seen as support for the claim that resultatives, bare nouns, idiom chunks, and perhaps certain adverbs share a common property with particles (e.g., they form a complex head with the verb as suggested by Neeleman 1994, or they target the same functional projection—PredP in Zwart's 1996 terms). On the other hand, the fact that these elements do not have to be part of the verb cluster in any of the languages discussed, shows that they also contrast with particles in some crucial respect. For instance, depending on one's analysis of particle constructions, the elements in (79) do not have to form a complex head with the verb, even in languages where particles are required to do so; or in the head-initial setting, the landing site of the elements in (79) can be further away from the verb than the one where particles must move to. The (mechanical) derivation of the word orders attested in verb clusters involving 'small' elements such as bare nouns, adverbs etc. is again fairly straightforward. The diagrams in (80) illustrate possible head-final and head-initial derivations and summarize the assumptions necessary to capture the distribution of particles and other small elements in verb clusters. As in the case of particles, the interesting (and still open) question regarding the distribution of these elements is whether the assumptions in (80) can be related to other properties of the particular languages or whether they simply constitute irreducible language-specific facts that do not follow from other conditions or constraints but that must be learned for each language.

(80) Verb Projection Raising with particles and other ‘small’ elements

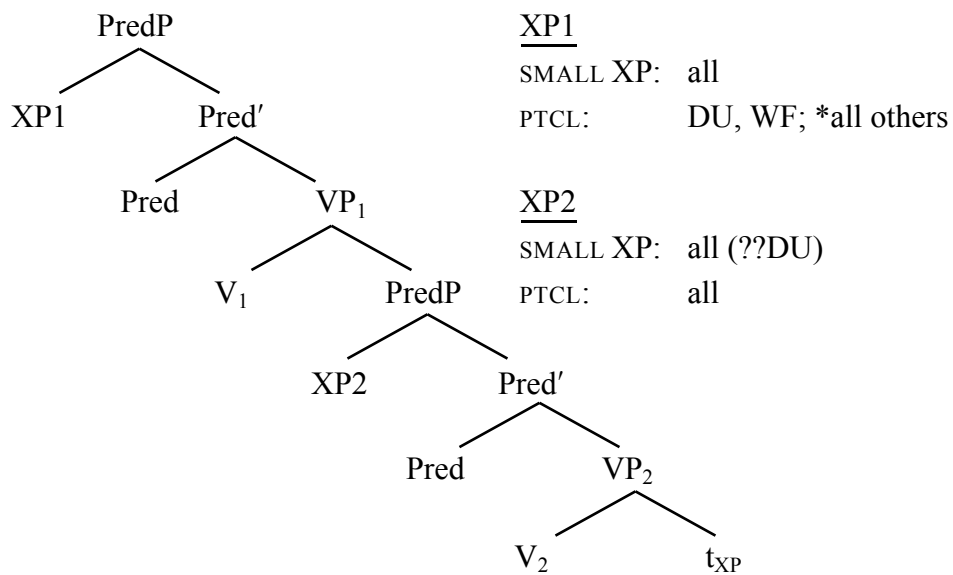
a. Head-final base: Head and phrasal movement



b. Head-final base: Phrasal movement only



c. Head-initial base



As shown in Table 8, the final two categories that have to be distinguished in VPR constructions are, on the one hand, indefinite (weak) noun phrases and prepositional phrase, and definite (strong) noun phrases on the other hand. While the phrases of the former type are allowed to occur within a verb cluster in Swiss German, West Flemish, and (at least for some speakers) in Afrikaans and German, the latter are only allowed to separate the verbs of a cluster in Swiss German and West Flemish. The examples in (81) illustrate these differences (further examples illustrating VPR in German can be found in den Besten and Broekhuis 1992, Broekhuis 1992: 190).

- (81) a. Dink jy ek sal kan<sub>1</sub> in Pretoria bly<sub>2</sub> Afrikaans  
 Think you I will can<sub>1</sub> in Pretoria stay<sub>2</sub>  
 ‘Do you think I will be able to get a place to stay in Pretoria?’ [Robbers 1997: 82]
- a’. %Ek sal nou kan<sub>1</sub> twee boeke betaal<sub>2</sub>  
 I will now can<sub>1</sub> two books pay<sub>2</sub>  
 ‘I will now be able to pay for two books’ [Robbers 1997: 77]
- a’’. \*Sy sal graag wil<sub>1</sub> die boek lees<sub>2</sub>  
 She will gladly want<sub>1</sub> the book read<sub>2</sub>  
 ‘She would be happy to read the book’ [Robbers 1997: 78; my paraphrase]
- b. dat Jan {naar Wenen} heeft<sub>1</sub> willen<sub>2</sub> {\*naar Wenen} gaan<sub>3</sub> Dutch  
 that Jan {to Vienna} has<sub>1</sub> want.IPP<sub>2</sub> {\*to Vienna} g<sub>3</sub>  
 ‘that Jan wanted to go to Vienna’ [F. Veraart p.c.]
- b’. dat Jan {een huis} wil<sub>1</sub> {\*een huis} kopen<sub>2</sub>  
 that Jan {a house} wants<sub>1</sub> {\*a house} buy<sub>2</sub>  
 ‘that Jan wants to buy a house’ [Haegeman and van Riemsdijk 1986: 419]
- b’’. dat Jan {het huis} wil<sub>1</sub> {\*het huis} kopen<sub>2</sub>  
 that Jan {the house} wants<sub>1</sub> {\*the house} buy<sub>2</sub>  
 ‘that Jan wants to buy the house’ [F. Veraart p.c.]
- c. %dass er nicht hätte<sub>1</sub> nach Turkestan fahren<sub>3</sub> sollen<sub>2</sub> German  
 that he not had<sub>1</sub> to Turkestan go<sub>3</sub> shall<sub>2</sub>  
 ‘that he should not have gone to Turkestan’
- c’. %dass er zumindest hätte<sub>1</sub> einen Trostpreis gewinnen<sub>3</sub> sollen<sub>2</sub>  
 that he at least had<sub>1</sub> a consolation prize win<sub>3</sub> shall<sub>2</sub>  
 ‘that he should at least have won a consolation prize’
- c’’. \*dass er vor der Abreise hätte<sub>1</sub> die Kakteen gießen<sub>3</sub> sollen<sub>2</sub>  
 that he before the departure had<sub>1</sub> the cacti water<sub>3</sub> shall<sub>2</sub>  
 ‘that he should have watered the cacti before his departure’
- d. dass er noed het<sub>1</sub> soele<sub>2</sub> noch Turkeschtan fahre<sub>3</sub> Swiss German  
 that he not had<sub>1</sub> shall<sub>2</sub> to Turkestan go<sub>3</sub>  
 ‘that he should not have gone to Turkestan’ [M. Schönenberger p.c., E. Haeberli p.c.]
- d’. dass er vor de Abreis het<sub>1</sub> sole<sub>2</sub> pBlueme/t Rose guesse<sub>3</sub>  
 that he before the departure had<sub>1</sub> shall<sub>2</sub> the.flowers/the roses water<sub>3</sub>  
 ‘that he should have watered the flowers/roses before his departure’

[M. Schönenberger p.c., E. Haeberli p.c.]

d''. dass er het<sub>1</sub> soele<sub>2</sub> da Buech gnau doerelaese<sub>3</sub>  
 that he had<sub>1</sub> shall<sub>2</sub> the book carefully through.read<sub>3</sub>  
 'that he should have read through the book carefully'

[M. Schönenberger p.c., E. Haeberli p.c.]

WF e. da Jan ... {da boek} wilt<sub>1</sub> {da boek} vuor Marie kopen<sub>2</sub> West Flemish  
 that Jan ... {that book} wants<sub>1</sub> {that book} for Marie buy<sub>2</sub>  
 'that Jan wants to buy that book for Marie' [Haegeman 1992: 148]

e'. da Valère wilt<sub>1</sub> vele boeken lezen<sub>2</sub> / een hus kopen<sub>2</sub>  
 that Valère wants<sub>1</sub> many books read<sub>2</sub> / a house buy<sub>2</sub>  
 'that Valère wants to read many books/buy a house'

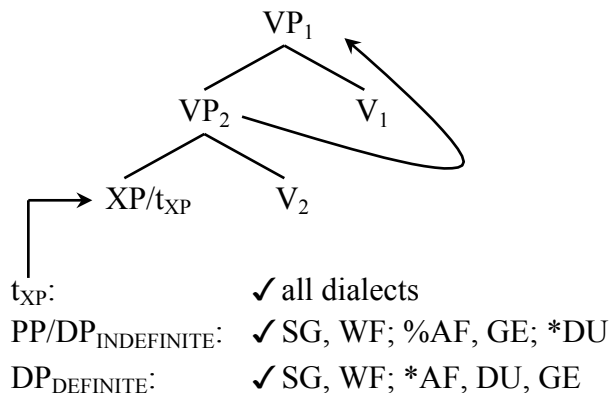
[Haegeman 1998b: 261/Haegeman and van Riemsdijk 1986: 419]

e''. da Valère willen<sub>1</sub> [ Marie dienen boek geven<sub>2</sub> ] eet<sub>3</sub>  
 that Valère want.IPP<sub>1</sub> [ Marie that book give<sub>2</sub> ] has<sub>3</sub>  
 'that Valère has wanted to give Marie that book' [Haegeman 1998b: 260]

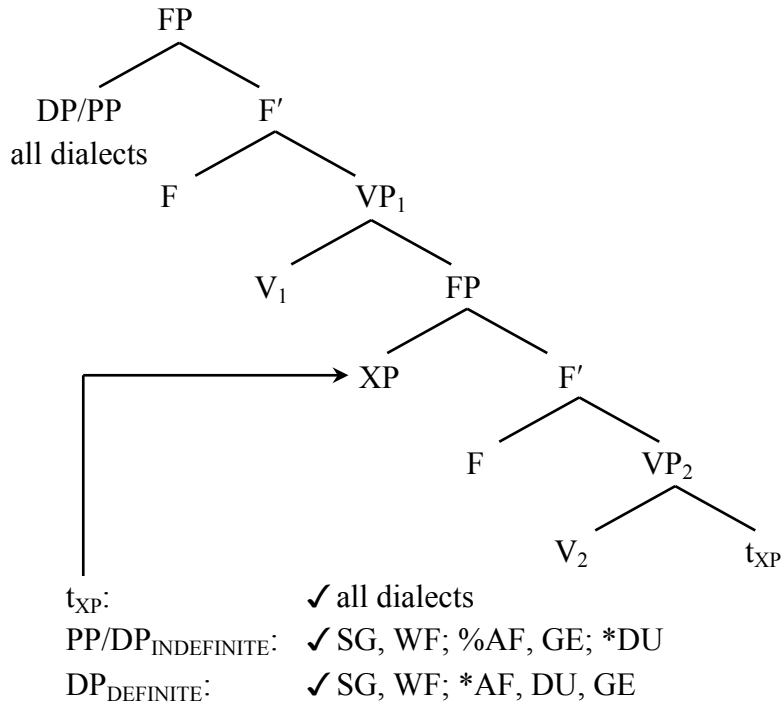
Under a head-final approach (cf. (82a)), definite objects must leave the VP prior to verb projection raising in Afrikaans, Dutch, and German, whereas in West Flemish and Swiss German, objects and prepositional phrases can be part of the projection that undergoes rightward movement. For some Afrikaans and German speakers, indefinite objects and prepositional phrases can also remain inside the VP. Under a head-initial approach (cf. (82b)), all XPs must leave the VP; indefinite objects and prepositional phrases can target a position lower than the higher verb in Swiss German and West Flemish, and in some Afrikaans and German dialects. Definite objects can target the intermediate position in Swiss German and West Flemish, but they must move to a position higher than the highest verb in all other dialects.

## (82) Verb Projection Raising with objects and prepositional phrases

### a. Head-final base: XP-movement account



b. Head-initial base



Parts of the above distribution of objects and prepositional phrases in verb clusters can be related to inherent differences between these elements. In particular, a frequent assumption is that definite and indefinite objects are subject to different licensing conditions (or more accurately, the conditions differ depending on whether an object represents old or new information). It is commonly assumed that definite objects in contrast to indefinite objects must leave the VP for semantic reasons (see for instance Diesing 1990, 1996, 1997, Bobaljik 1995 among many others). Thus, if definite objects are required to move to a VP-external position (head-final approach) or they must target a higher specifier than indefinite objects and prepositional phrases (head-initial approach), the difference encountered in Afrikaans and German between definite and indefinite objects can be related to the semantic difference of the elements involved. The distribution of objects in verb clusters thus provides further evidence for this view. However, unfortunately, this semantic correlation only captures parts of the distribution. Further language-specific assumptions are again necessary to account e.g., for the fact that definite objects can target the lower FP (in a structure like (82b)) in Swiss German and West Flemish or the fact that the lower FP is unavailable for all objects in Dutch (and for some Afrikaans and German speakers).

In conclusion, the VPR phenomenon is not an absolute property of a language (i.e., languages do not simply allow or disallow VPR structures), but the felicity of these structures depends crucially on the type of material that occurs between the verbs of a cluster. Comparing the composition of verb clusters across West Germanic, some interesting generalizations regarding the material that can occur interspersed in a verb cluster can be found. However, the distribution of verb projection raising material is also subject to language-specific properties that do not necessarily follow from other properties of grammar.



### 3.3 Motivation for verb clustering

At the beginning of section 3, two questions were raised: i) how are verb clusters derived, and ii) how does the variation across languages and constructions arise. So far, we have mostly been concerned with the first question, although the discussion of Zwart (1996) has also touched on the second question. The question of how to account for the variation found in the distribution of verb clusters is tightly connected to the question of how whatever mechanism one employs to achieve the desired word orders is motivated. A common theoretical mandate, in particular in Minimalist approaches of syntax, is that everything happens for a reason. Thus, if certain verb cluster orders are derived via syntactic movement, the immediate question is why the movements we find appear to be optional: if X moves in language L1 and movement is for a particular reason, why does that reason not apply in language L2 where such movement does not happen? In a perfect world, syntactic computations should not differ from one language to the next, unless it can be shown that the different computations involve different syntactic items or are related to a general parametric difference between the languages under consideration.

One of the biggest challenges verb clusters have posed for syntactic theories is the fact that the variation seems to be to some (a large?) extent arbitrary, and, most importantly, that, in two languages, constructions with identical meaning and the same type of elements involved can still show a difference in distribution in these languages. To deal with this issue, three basic approaches have been put forward. The first approach, which was discussed in section 3.1, assumes that verb cluster orders are determined post-syntactically. This relieves the syntactic computation of the verb cluster problem and puts the burden into PF, an area for which variation is easier to deal with, at least in certain approaches. As was mentioned in the course of the directionality debate, there are differences in word order across languages which appear to be non-derivable. If head-final vs. head-initial configurations are the result of different linearization rules which apply when syntax hands things over to PF, verb clusters are just an extreme version of such language-specific spell-out rules.

The second approach aims at building the motivation for reordering into the syntactic structure, usually via features that trigger movement or the needs of various elements to fulfill certain licensing conditions. In this approach, the variation is attributed to different feature compositions (e.g., feature strength, the presence/absence of a movement-triggering feature, or different feature values). If these features or conditions can be independently motivated, an account deriving the distribution of verb clusters by means of these features or licensing conditions would offer an insightful explanation for the phenomenon and hence be clearly preferable to the descriptive accounts outlined in section 2.5. In the next section, it will be shown that there are several ways features can be used to mechanically derive the possible word orders, but that at a deep level, the original problem remains—while the movement operations may seem to have a nominal motivation, the features themselves become an issue and the new question is why constructions that are seemingly identical are nevertheless represented by such different features.

This issue has also been explicitly raised in Barbiers (2005), which will bring us to the third and last approach to be summarized in section 3.3.2. Barbiers proposes a syntactic account that derives two crucial properties: the general impossibility of the 2-1-3 order and the fact that the 2-3-1 order, cross-linguistically, is only found in auxiliary-modal-V constructions. In contrast to other syntactic accounts, Barbiers' syntax does not distinguish (in any obvious way) between possible and impossible orders in specific languages or dialects, but only regulates when movement is possible in principle in each construction (in any language). Thus, the syntax proposed derives a larger set of word orders than are attested for each language. Similar to the PF-view, a

large part of the cross-linguistic variation is kept out of syntax proper, in that the featural makeup of the elements involved as well as the syntactic licensing conditions are identical for constructions of same type across languages, independently of whether movement applies or doesn't apply in a given language. In Barbiers' model, all verb cluster movement operations are in principle optional, and to restrict the possible word orders in different languages/dialects, language-specific optionality assumptions need to be added. This proposal, in a very interesting fresh way, singles out what can be seen as the universal parts of verb clusters and separates those parts from the parts that are language-specific.

### 3.3.1 Featural approaches

In this section, I summarize three types of approaches that aim at deriving the variation found in verb clusters from the featural make-up of the constructions involved: i) over/covert approach; ii) overt positional licensing approach; iii) combination of i) and ii).

To illustrate the problem of motivating verb cluster movements and accounting for the variation, it will suffice to look at the distribution of clusters with two verbal elements. Since most accounts to summarize follow a head-initial base structure, I limit the discussion to head-initial settings (see section 2.5 for head-final generalizations and derivations). Table 9 summarizes the distribution of two-verb clusters in the languages discussed and also lists whether movement is necessary, possible, or impossible in the two constructions. The two rows labeled *Unattested* should be read as: so far, no language has been attested that disallows movement in auxiliary-participle constructions; and no language has been found that has optional movement in auxiliary-participle constructions but obligatory movement in modal-infinitive constructions.

Table 9. Five language types (two verbs): Head-initial derivations

Language	AUX-PART	Movement	MOD-INF	Derivation
German, Frisian	2-1	Obligatory	2-1	Obligatory
Dutch (1=NON-FIN)	2-1/1-2	Optional	*2-1/1-2	Impossible
Swiss-1	2-1	Obligatory	2-1/1-2	Optional
Swiss-2, Dutch (1=FIN)	2-1/1-2	Optional	2-1/1-2	Optional
West Flemish, Afrikaans	2-1	Obligatory	*2-1/1-2	Impossible
Unattested	*2-1/1-2	Impossible		
Unattested	2-1/1-2	Optional	2-1/*1-2	Obligatory

The descriptive account offered in section 2.5, Table 3 derives this distribution by rules demanding, allowing, or prohibiting the reordering of the elements involved. Such inversion rules, however, are considered to be unmotivated since, although they cover the variation found in the distribution of verb clusters, they suffer in terms of motivation, predictability, and restrictiveness. More specifically, nothing seems to prevent a rule that disallows inversion between auxiliaries and participles (there appears to be exactly such a rule for modal-infinitive constructions in West Flemish and Afrikaans), or a rule that requires inversion for modals and auxiliaries (as we find for auxiliary and participles in several of the languages).

The first attempt at deriving the word order variation is by invoking a (universal) requirement that verbal elements must undergo movement, however, languages can differ regarding whether this movement is overt or covert (similar to the claim in Zwart 1996, illustrated in section 3.2.1.5, that infinitives must move covertly in Dutch, but overtly in German). For West

Germanic two-verb clusters, one would then have to assume the following overt/covert settings (note that it does not matter for the discussion here if covert movement is seen as movement at LF or overt movement with pronunciation of the lower copy).

(83) Overt/covert settings for two-verb clusters (head-initial base)

German, Frisian:	Infinitives and participles move overtly
Dutch (I=NON-FIN):	Infinitives move covertly, participles overtly or covertly
Dutch, Swiss-2:	Infinitives and participles move overtly or covertly
Swiss-1:	Infinitives move overtly or covertly, participles overtly
West Flemish, Afrikaans:	Infinitives move covertly, participles overtly

While the settings in (83) derive the possible word orders in the languages listed, the question arises whether this approach goes beyond the descriptive accounts presented in section 2.5. Like it is the case for the statement of inversion rules, the criteria of predictability and restrictiveness are not met. There is again no reason why there should not be a language with the settings *participles move (only) covertly* or *infinitives move overtly and participles move overtly or covertly*. Furthermore, the overt/covert assumptions in (83) do not correlate with any other property found in the language(s) discussed but only serve the purpose of deriving the word orders in verb clusters.

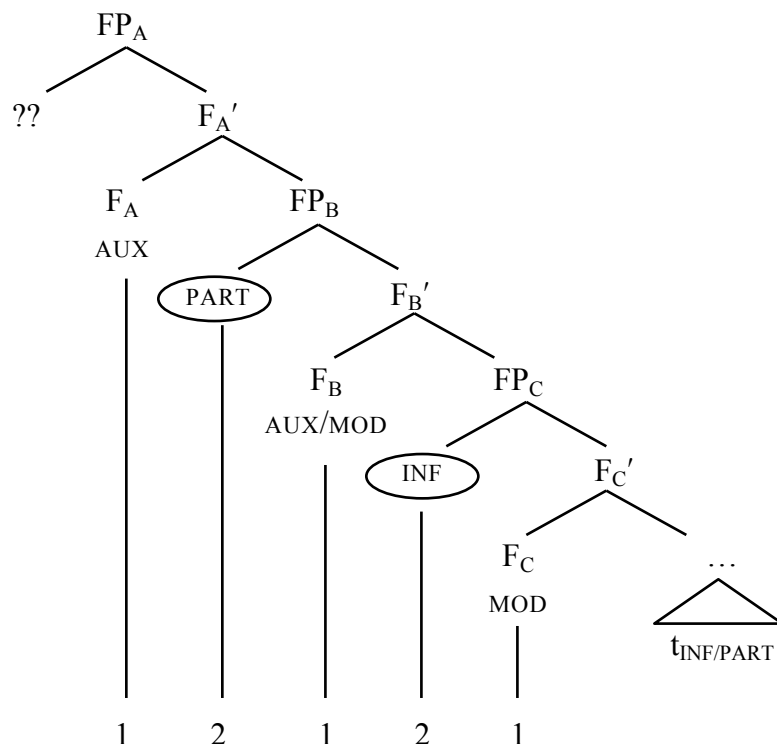
A different and at first sight more promising approach to word order variation in verb clusters is what I will refer to as *positional* approach here. The core idea of positional analyses is that word order differences are not the result of a difference in the timing of movement, but rather arise from different positions an element can/must occupy in different languages. This approach is a direct consequence of expansions of phrase structure (as for instance suggested by Rizzi 1997 or Cinque 1999) and the Minimalist idea that movement is triggered by the presence of features. In particular, an increase of the number of functional projections has offered new landing sites for movement, and hence new ways to encode differences between languages without reference to an overt/covert distinction. The step from the traditional verb (projection) raising approach to an analysis of verb cluster movement as movement to functional projections also (partially) solves a constant problem arising for V-to-V and VP-to-VP raising analyses—namely the question of how movement can be triggered by a lexical head (since, as pointed out by den Dikken 2000, lexical heads typically do not serve as attractors).

The idea that word order variation is the result of different positions rather than an overt/covert distinction is reflected (fully or partly) in many analyses of verb clusters (see Haegeman 1995, 1998b, c, den Dikken and Hoekstra 1997, Robbers 1997, Koopman and Szabolcsi 2000 among many others). The challenge of positional analyses is to avoid over-generation and to limit the options offered by the in principle unrestricted number of functional projections (i.e., although the core idea of positional analyses is that certain elements can occur in different positions in different languages, it is not the case that they can occur just anywhere). The general approach to the over-generalization problem is to assume that one of the elements in a two-verb cluster is fixed to one (derived) position, whereas the other element is arranged around it. Haegeman (1995, 1998b, c) for instance assumes that the positions of the participle and the infinitive are fixed, whereas the positions of the auxiliary and the modal are variable. Robbers (1997) and Zwart (1996), on the other hand, assume that the position of the auxiliary is fixed whereas the position of the participle is variable (Zwart's account also makes use of an overt/covert distinction).

A positional account for two-verb clusters is outlined in (84). As can be seen, three functional projections are necessary to account for the word order differences in the languages discussed. Spec,FP<sub>B</sub> has been chosen as the designated position for participles, whereas Spec,FP<sub>C</sub> is the position targeted by infinitives. Under such a positional approach, participles (cross-linguistically) move to Spec,FP<sub>B</sub> and infinitives move to Spec,FP<sub>C</sub> in overt syntax. Modals and auxiliaries, on the other hand, are variable in that they can occur in either F<sub>A</sub>, F<sub>B</sub> or F<sub>C</sub>. Thus, if the auxiliary occupies F<sub>A</sub>, the 1-2 order is derived; if it occupies F<sub>B</sub> or F<sub>C</sub>, the 2-1 order is derived. Similarly, if the modal is in F<sub>A</sub> or F<sub>B</sub>, the 1-2 order is derived, if it is in F<sub>C</sub>, the 2-1 order is derived (to limit the positions, it would be sufficient to assume that auxiliaries either occupy F<sub>A</sub> or F<sub>B</sub>, whereas modals occupy either F<sub>B</sub> or F<sub>C</sub> in (84a)). The variation found across West Germanic can now be derived from the settings in (84b).

(84) Positional analysis of two-verb clusters

a. Structure



b. German, Frisian:	AUX: F <sub>B/C</sub>	MOD: F <sub>C</sub>
Dutch (1=NON-FIN):	AUX: F <sub>A,B/C</sub>	MOD: F <sub>A/B</sub>
Dutch, Swiss-2:	AUX: F <sub>A,B/C</sub>	MOD: F <sub>A/B,C</sub>
Swiss-1:	AUX: F <sub>B/C</sub>	MOD: F <sub>A/B,C</sub>
West Flemish, Afrikaans:	AUX: F <sub>B/C</sub>	MOD: F <sub>A/B</sub>

This approach is appealing since it reduces the variation found across languages to simple claims about the position of certain elements. Despite this point of attractiveness, the question to be asked again is how this account fares in terms of predictability and restrictiveness. Does this approach make any predictions about impossible combinations and are the assumptions in (84) independently motivated? Unfortunately, it seems that the answers are again negative. First, it is

not clear under this approach why certain combinations are unattested. In particular, there are no languages or dialects that require auxiliaries to occupy  $F_A$  only; and there are no languages that allow auxiliaries in any position but require modals to occur in  $F_C$ . Regarding the former case, one might suggest that languages that only allow auxiliaries in  $F_A$  and not in  $F_B$  do not exist since overt specifier-head configurations between elements that are in some checking relation (such as auxiliaries and participles) have to be at least possible in every language. However, this seems to be falsified by West Flemish and Afrikaans which only allow modals in  $F_{A/B}$  and not in  $F_C$  (i.e., an overt specifier-head configuration between two elements in a checking relation is *not* allowed in these languages). Thus, while the positional approach derives the possible orders it does not provide any obvious way other than by stipulation to exclude the unattested configurations.

An answer to the question of whether the movements/positions are motivated by properties independent of word order depends on whether it is possible to assign a function or meaning to the functional projections in (84). As for  $F_B$  and  $F_C$ , it could be assumed that these heads are the designated positions for auxiliaries ( $F_B$  encodes e.g., aspectual information such as perfect(ive)) and modal verbs ( $F_C$  encodes modality). Furthermore, the idea that these heads must enter into a checking relation, in particular a specifier-head relation, with a participle in Spec,FP<sub>B</sub> or an infinitive in Spec,FP<sub>C</sub> is a fairly standard assumption in theories that define licensing as checking in specifier-head configurations. A question, however, arises regarding the highest projection  $F_A$ —what type of information or features are expressed by this projection? It is not obvious what the answer to this question is. A similar concern arises for  $F_B$ . Assuming that  $F_C$  represents the base position for modals, one has to ask why modals can or must move further to  $F_B$  in languages such as West Flemish or Afrikaans.

This question becomes an even more important issue if we look closer at the way positional analyses (in particular pure surface positional analyses that dispense with covert movement) handle optionality. Although positional analyses are built on the claim that certain elements can occur in different positions (both within one language or across languages), upon closer inspection, this optionality seems in fact to be inherently inconsistent with a pure positional approach. To illustrate this problem, consider Dutch and Swiss-2. As shown in (84), modals and auxiliaries can occur in any of the functional heads postulated (i.e., they can either precede or follow the corresponding participle or infinitive). There are two ways to derive this optionality. First, it could be assumed that modals and auxiliaries do not have universally designated positions but are inserted in different positions in different languages and/or constructions. Second, one could imagine that modals and auxiliaries are base generated (universally) in designated positions (i.e., modals in  $F_C$  and auxiliaries in  $F_B$ ) and undergo further movement in some languages. Regarding the first option, it should be noted that this option does not provide a way to relate the assumption about where elements are inserted in a particular language to any other property (even the idea that the functional projections in (84) are related to notions such as aspect or modality would be given up in this approach). Thus, this option is entirely descriptive in that it simply states the order of elements in a clause and hence does not provide any insight into the distribution of verbal elements. According to the second option, modals and auxiliaries are inserted universally in the lower positions; in Dutch and Swiss-2, however, these elements have a choice of moving on or remaining in their base positions. Since in a pure positional account no covert movement is available, it must be the case that all licensing/checking requirements are met in surface positions. For Dutch and Swiss-2 modals and auxiliaries, this means that since these elements can stay in the lower positions  $F_C$  and  $F_B$  in the 2-1 orders, they must also be licensed in the lower positions (otherwise, the 2-1 orders should be ungrammatical, given that covert move-

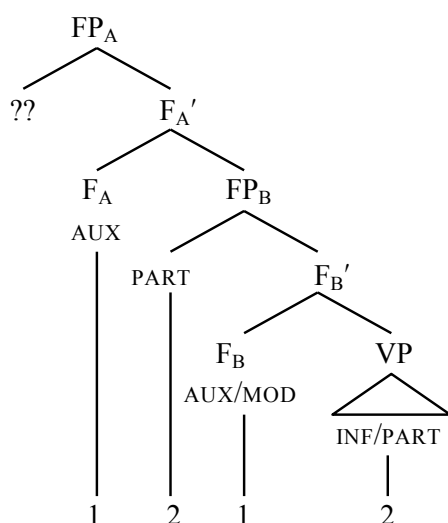
ment is unavailable). However, if this is the case a serious question arises for the orders that involve movement. What triggers this movement? The conclusion one must draw is that this movement must be untriggered and cannot be caused by any licensing/checking requirement. If there was such a requirement, it would not be met in the case where movement does not occur (again since this approach rejects covert movement). Thus, under these premises, movement as in Dutch or Swiss would have to be movement that is not motivated by any syntactic licensing requirement. The only way licensing could be built into such an approach is if it is assumed that the 1-2 order and the 2-1 order involve crucially different features or properties that must be licensed in different positions. However, the works on verb clusters seem to agree that apart from word order there is no difference between examples showing the 1-2 order and those showing the 2-1 order (furthermore, this claim would lead one to conclude that the constructions corresponding to *John must leave* are entirely different constructions in West Flemish and German, despite the fact that they mean the same).

To conclude, a positional approach that is based on the claims that i) there is no covert movement, and ii) movement is motivated by feature checking leads to a paradox that can only be solved by giving up one of the two assumptions. If the second assumption is given up, i.e., if movement is not considered to be motivated, the assumptions in (84) reduce to stipulations about whether certain elements can or cannot precede other elements. Hence, an account along these lines would not go beyond simple descriptive statements about the word order in verb clusters, and therefore cannot be seen as superior to the descriptive accounts outlined in section 2.5. The second option—i.e., a positional account that allows covert movement—seems more promising and will be discussed in the remainder of this section.

In a combined positional and overt/covert account, the minimal structure for two-verb clusters would be as in (85). In (85), auxiliaries occur in  $F_B$ , infinitives and participles move overtly or covertly to  $Spec,FP_B$ , and modals can occupy either  $F_A$  or  $F_B$ . Thus, if VP-movement is covert, the 1-2 orders are derived; if VP-movement is overt, the 2-1 order is derived in the auxiliary-participle construction, and either the 1-2 or the 2-1 order in the modal-infinitive construction depending on the position of the modal. The settings necessary to account for the distribution of two-verb clusters in the languages discussed are listed in (85b).

(85) Positional analysis of two-verb clusters & covert movement

a. Structure



- |    |                          |  |
|----|--------------------------|--|
| b. | German, Frisian:         | MOD in F <sub>B</sub> ; overt movement                             |
|    | Dutch (1=NON-FIN):       | non-fin MOD in F <sub>A</sub> ; overt/covert                       |
|    | Dutch, Swiss-2:          | MOD in F <sub>A</sub> or F <sub>B</sub> ; overt or covert movement |
|    | Swiss-1:                 | MOD in F <sub>A</sub> or F <sub>B</sub> ; overt movement           |
|    | West Flemish, Afrikaans: | MOD in F <sub>A</sub> ; overt movement                             |

As these settings show, the advantage of a combined positional and overt/covert system over an account that makes use of only the overt/covert distinction is that a single overt/covert setting is sufficient in (85) for all constructions (and languages), and no construction specific stipulations are necessary to derive differences between auxiliary-participle constructions and modal-infinitive constructions. Nevertheless, the system in (85) involves two types of parameters that have to be set for each language—an option between overt and covert movement and an option between two positions for modals. The two positions for modals are essential to account, for instance, for the difference between German and West Flemish. In both languages, overt VP-movement takes place, which guarantees that auxiliary-participle constructions appear in the 2-1 order. However, to derive the 1-2 order in modal-infinitive constructions in West Flemish and Afrikaans but not in German (by keeping the claim that VP-movement is overt in these languages; if it were overt for participles but covert for infinitives in West Flemish and Afrikaans, we would end up again with a system such as the one in (83)), two positions for modals are necessary.

Does this account finally meet the criteria of predictability and restrictiveness? Let us first consider if a positional account involving overt and covert movement can exclude the unattested cases. In the scenario in (85), these would be languages with obligatory covert movement, and languages with overt or covert movement for participles but obligatorily covert movement for infinitives. Thus, to exclude these scenarios, all one has to do is to add a ban on obligatorily covert movement and a requirement that the overt/covert parameter is set uniformly for all constructions in one language. Thus, while these assumptions would have to be evaluated with respect to other constructions (in particular, they must be tested in three-verb clusters), it is important to note that this account appears to be the most successful one in providing a way to approach this over-generalization problem.

Without engaging in an endless discussion of larger verb cluster constructions, there are, however, two concerns that cast some doubt on the straightforwardness of the account outlined in (85). First, the system in (85) crucially relies on the assumption that cross-linguistically, auxiliaries occur in a fixed position, whereas modals and main verbs have a choice. It is not easy to see why there should be this distinction. In particular, the difference between modals and auxiliaries does not seem to follow from any obvious property of these elements and hence one could easily imagine the opposite situation (it is also important to note in this respect that this difference between modals and auxiliaries is essential to exclude the impossible orders as suggested above). The second question the account in (85) is faced with is whether the settings can be independently motivated or whether they only serve the purpose of deriving (two-)verb clusters. One way to motivate the two functional projections in (85) is to assume that FP<sub>B</sub> is the projection responsible for the licensing of auxiliaries in F<sub>B</sub> and main verbs in Spec,FP<sub>B</sub>, whereas FP<sub>A</sub> is the projection responsible for the licensing of modals.<sup>34</sup> Under this assumption, modals which start

---

<sup>34</sup> The claim that movement to functional projections is motivated by the need to check features is itself based on an assumption which appears to be controversial. What kind of features do verbs, modals, auxiliaries etc. check? Features commonly assumed are so-called ‘V’ features. Although it seems that these rather vague features do no more

out in  $F_B$  or a lower VP, move (overtly or covertly) to  $F_A$  to check whatever “modal” features have to be checked. Thus, in West Flemish and Afrikaans, for instance, modals move overtly, whereas in German and Frisian, they move covertly. However, assuming that much, a re-evaluation of the point that has been mentioned as an advantage of this system becomes necessary. Although no construction-specific settings are necessary under this approach for auxiliaries, infinitives, and participles, and a general overt/covert setting is sufficient for VP-movement, the distribution of modals does require additional assumptions. In particular, German and Frisian require different timing settings for VP-movement (which must be overt) and movement of modals (which must be covert). Thus, the claim that there is a uniform overt/covert setting should be qualified, in particular, different timing settings must be allowed and assumed for modals, on the one hand, and main verbs, on the other hand. To conclude, although the system in (85) appears attractive in its restrictiveness, it nevertheless—like the other approaches sketched—involves crucial stipulations that do little more than state the orders of verbal elements in a cluster.

In sum, the discussion in this section has shown that to account for the distribution of West Germanic two-verb clusters, all three approaches outlined above must make crucial reference to category-specific assumptions (either in terms of the timing of movement or the position a particular category targets, or both). In particular, what seems to be unavoidable is the assumption of two differences that so far have not been correlated with any other (universal or language-specific) property: i) an overt vs. covert difference or different positions for different categories within one language (i.e., an element such as modal or auxiliary occurs in  $FP_A$  vs.  $FP_B$  based on the category of that element), and ii) an overt vs. covert difference or different positions for elements of the same category in different languages (i.e., an element of a particular type occurs in  $FP_A$  in one language but in  $FP_B$  in another language). While it is undoubtedly the case that descriptive accounts such as the ones presented in Table 3 are unattractive for their lack of explanatory value, it is not clear at the current stage of the research on verb clusters that it is trivial (or possible at all) to come up with an alternative that does not suffer from these very same inadequacies.

### 3.3.2 *Optionality in syntax*

The question of what motivates movement and how the possible and in particular impossible word orders can be accounted for is central in Barbiers (2005). Since the approach presented there provides a new direction for the verb cluster research and has consequences for most of the questions raised in this overview, as well as for syntactic theory in general, it will be discussed in more detail here.

Barbiers analysis is based on the distribution of verb clusters in the Dutch dialects, as determined by the SAND project. To remind the reader, the main facts from Table 2 are replicated in Table 10. As shown, Dutch dialects allow four, possibly five of the six word orders (the 2-1-3 order is not found) in IPP constructions. In double modal constructions and modal-auxiliary-V constructions, four orders are found—the 2-1-3 and 2-3-1 orders are not attested robustly in these constructions. As we have seen, the sum of the options found in the Dutch dialects thus represents the sum of possibilities of verb cluster orders in general in West Germanic.

---

than account for movement in the absence of any other good reason (see also Koopman 1984 for a similar observation regarding features such as  $[\pm\text{AFFIX}]$ ), it will be granted for the discussion here that movement that is triggered by these kind of features is motivated movement.



Table 10 Verb clusters with three verbal elements in the Dutch dialects

Language	MOD-MOD-V	AUX-MOD-V	MOD-AUX-V	AUX-AUX-V
	FIN-INF-INF	FIN-IPP-INF	FIN-INF-PART	FIN-PART-PART
Dutch Dialects	1-2-3	1-2-3	3-1-2	no information
	3-1-2	3-2-1	1-3-2	
	3-2-1	2-3-1	1-2-3	
	1-3-2	1-3-2	3-2-1	
		[3-1-2]		

One goal of Barbiers' analysis is to account for i) the cross-linguistic lack of the 2-1-3 order and ii) the limitation of the 2-3-1 order to the IPP construction. Since these generalizations hold for verb clusters throughout West Germanic (and not only for Dutch; see again Table 2), Barbiers' approach extends readily to verb clusters in general.

Barbiers account is a syntactic account involving feature licensing. Specifically, it is assumed that modals and auxiliaries have uninterpretable features (*u*Fs) that need to be licensed by Agree with a corresponding interpretable feature. The distribution of features assumed is given in (86). As shown, modals (including IPP modals) are equipped with a *u*Event feature, which requires that modals establish a dependency with a main verb, as main verbs are the only elements that are assumed to have an interpretable counterpart of that feature. The perfect auxiliary *have* is equipped with a *u*Event feature, which makes this element dependent on either a main verb or a modal in the IPP.

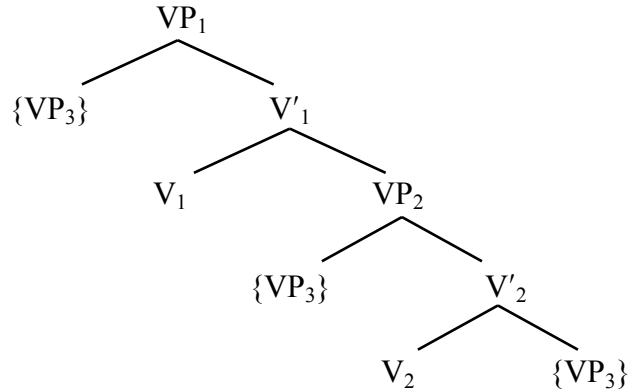
- (86) a. modals: *u*Event  
b. perfect auxiliary (*have*): *u*Perf, *u*Event<sup>35</sup>  
c. main verb infinitive: *i*Event  
d. main verb perfect participles: *i*Perf, *i*Event  
e. IPP modal: *i*Perf, *u*Event

Agree applies cyclically, and after Agree has been established between a *u*F and a corresponding *i*F, the uninterpretable feature is deleted. Crucially, movement is never triggered by features, but it is assumed to happen freely (subject only to language and construction-specific assumptions), whenever there is an Agree relation. In verb clusters, movement is phrasal—the Agreeing XP moves to the specifier of the Agreeing head. This is illustrated in (88). In (88a), V2 Agrees with V3, which allows movement of VP3 to the specifier of V2. If V1 then also Agrees with V3, further movement of VP3 to the specifier of V1 can occur. Optionality is given by { }, hence three word orders are derived in (88a): 3-1-2, 1-3-2, 1-2-3. If, on the other hand, V1 Agrees with V2, movement of VP2 can occur as in (87b), yielding the 3-2-1 and 2-3-1 orders. A final assumption is that if a head Agrees with the specifier of a lower XP (as in (88a)), in addition to movement of that specifier (VP3 in (88a)), movement can also pied-pipe the entire XP. This is illustrated in (88c), which again then derives the 3-2-1 order. Note that the only way for VP3 to pied-pipe VP2

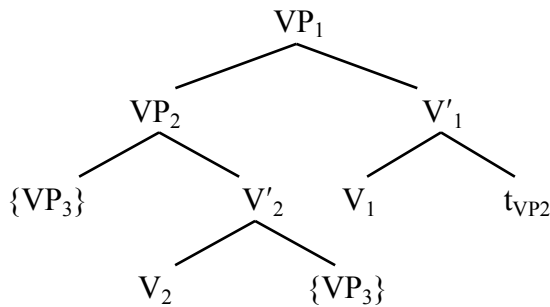
<sup>35</sup> In Barbiers (2005: 252f), it is assumed that the perfect auxiliary does not have a *u*Event feature in the AUX-MOD-V construction, but it does have such a feature in the MOD-AUX-V construction. This assumption is not justified beyond the goal to exclude the 3-1-2 order in IPP constructions. Since the status of this order is still unclear, I present Barbiers' analysis without this stipulation (hence generating the 3-1-2 order in IPP constructions as well).

is if VP3 has moved in the first step, i.e., if it is in the specifier of V2. If this first step of movement has not occurred, VP3 would have to move by itself, yielding only the 3-1-2 order.

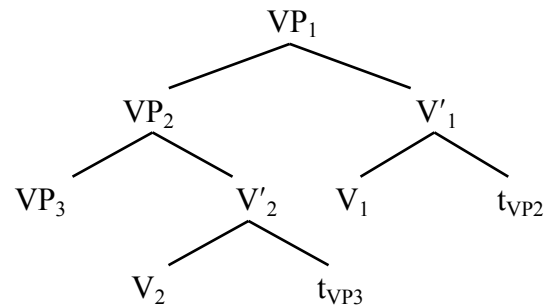
(88) a. V1, V2 Agree with V3



b. V1 Agrees with V2



c. V1 Agrees with V3; pied piping



This system has two consequences. First, the 2-1-3 order cannot be derived. This is independent of the specific Agree relations, but is built into the assumptions of locality of movement: verb cluster movement is phrasal movement (X'-categories do not move), and movement can only target specifiers. This derives essentially that there is no remnant movement. Like a head-final approach that excludes the 2-1-3 order via locality, this approach hence scores in terms of restrictiveness. Second, there is only one way to derive the 2-3-1 order: if V1 Agrees with V2. We will see that this claim, together with the feature system in (86) derives the generalization that the 2-3-1 order can only arise in the IPP construction. This is a major achievement, as no other account so far, has offered a principled way of predicting this generalization.

Armed with these assumptions, let us now see how the (im)possible word orders and the two generalizations noted are derived. Since main verbs always have at least one *i*F, the verbal element merging with a main verb can (and will) always Agree with the main verb. This has the result that in verb clusters with three verbs, movement of VP3 to the specifier of V2 is always an option, yielding the ordering 3-2 in all constructions (if no movement occurs, the order remains the head-initial 2-3 order). As for the Agree relation established by V1 the specific feature combinations become crucial.

The derivations for double modal constructions are given in (89). V2 Agrees with V3, erasing the uninterpretable feature on V2. As indicated, movement creates the 3-2 order (with VP3 occurring in the specifier of V2), and no movement the 2-3 order. The finite modal is also

merged with an *uEvent* feature, which triggers Agree between V1 and V3 (note that, by assumption, only main verbs have an *iEvent* feature, thus V3 is the only verb that V1 can Agree with in this configuration). Since the *uEvent* feature of V2 has been deleted, no issue of locality arises by V1 Agreeing with V3 across V2. As for movement, there are now four options to consider—movement and no movement from both outputs in (89a), given as (i) and (ii). Movement of VP3 in structure (i) yields the 3-1-2 order, and if VP2 is pied piped also the 3-2-1 order. Movement of VP3 in structure (ii), also yields the 3-1-2 order. If Agree between V1 and VP3 is not followed by movement, structure (i) yields the 1-3-2 order, and structure (ii) yields the 1-2-3 order. Thus, the four possible orders are derived in this system, and crucially, the impossible 2-3-1 order cannot be derived. To derive the 2-3-1 order, V1 would need to Agree with V2, which, given the assumed feature distribution is not possible (see below for further comments on this point).

(89) a.	Merge:	MOD.INF <sub>2</sub>	V.INF <sub>3</sub>	Move	No move
		<i>uEvent</i>	<i>iEvent</i>		
	Agree:	<del><i>uEvent</i></del>	<i>iEvent</i>	3-2 (i)	2-3 (ii)
b.	Merge:	MOD.FIN <sub>1</sub>	MOD.INF <sub>2</sub>	V.INF <sub>3</sub>	
		<i>uEvent</i>	—	<i>iEvent</i>	
	Agree:	<del><i>uEvent</i></del>	—	<i>iEvent</i>	3-1-2, 3-2-1 (i;pp) 1-3-2 (i), 1-2-3 (ii)

The derivations for the modal-auxiliary-V construction are given in (90). The derivations are basically identical to the double modal construction, with the only difference that the perfect auxiliary also has a *uPerf* and the main verb participle the corresponding *iPerf*. Thus in the first step of Agree, two features are licensed and erased. Since the Agree relations are established between the same elements, the same movement options arise, and the output orders are therefore the same as in the double modal construction.

(90) a.	Merge:	AUX.INF <sub>2</sub>	V.PART <sub>3</sub>	Move	No move
		<i>uEvent</i>	<i>iEvent</i>		
		<i>uPerf</i>	<i>iPerf</i>		
	Agree:	<del><i>uEvent</i></del>	<i>iEvent</i>		
		<del><i>uPerf</i></del>	<i>iPerf</i>	3-2 (i)	2-3 (ii)
b.	Merge:	MOD.FIN <sub>1</sub>	AUX.INF <sub>2</sub>	V.PART <sub>3</sub>	
		<i>uEvent</i>	—	<i>iEvent</i>	
				<i>iPerf</i>	
	Agree:	<del><i>uEvent</i></del>	—	<i>iEvent</i>	
				<i>iPerf</i>	3-1-2, 3-2-1 (i;pp) 1-3-2 (i), 1-2-3 (ii)

Lastly the derivations for IPP constructions are given in (91). As a special assumption, IPP modals are assumed to have an *iPerf*. The first step of the derivation is as before—V2 Agrees with V3, yielding the 3-2 or 2-3 order in the lower part of the structure. The second step, however, is different from the previous derivations. Since V1 is merged with a *uPerf*, and only V2 has a corresponding *iPerf*, V1 establishes an Agree relation with V2. This has the important consequence that VP2 is now licensed to move, which correctly derives the possibility of the 2-3-1 order.<sup>36</sup>

<sup>36</sup> The derivation in (91), may in principle also derive the 3-1-2 order if the Agree relation between V1 and V3 trig-

(91) a.	Merge:	MOD.IPP <sub>2</sub> <i>uEvent</i> <i>iPerf</i>	V.INF <sub>3</sub> <i>iEvent</i>	Move	No move
	Agree:	<del><i>uEvent</i></del> <i>iPerf</i>	<i>iEvent</i>	3-2 (i)	2-3 (ii)
b.	Merge:	AUX.FIN <sub>1</sub> <i>uEvent</i> <i>uPerf</i>	MOD.IPP <sub>2</sub> <i>iPerf</i>	V.INF <sub>3</sub> <i>iEvent</i>	
	Agree:	<del><i>uEvent</i></del> <del><i>uPerf</i></del>	— <i>iPerf</i>	<i>iEvent</i>	3-2-1 (i), 2-3-1 (ii) 1-3-2 (i), 1-2-3 (ii)

As mentioned in fn. 35, Barbiers excludes the 3-1-2 order as a possible order of IPP constructions, and to derive this property, it is assumed that AUX does not come with a *uEvent* feature in this case. To avoid this stipulation and still derive the lack of the 3-1-2 order one could appeal to the order of Agree relations established by V1 and restrict movement to the goal of the first Agree relation. If V1 Agrees with the closer V2 first in (91b), movement of 2P would take preference over movement of 3P.

While this system, as mentioned, is very attractive in deriving not only the possible but also excluding the impossible word orders and in accounting for the generalizations regarding the distribution of the 2-1-3 and 2-3-1 orders, some points about the predictability of the system should be noted nevertheless. As we have seen, the major work is done by the features assumed, and one has to ask whether the distribution in (86) is motivated. Even if we eliminate the stipulation regarding the missing *uEvent* feature on the auxiliary in IPP constructions (see fnn. 35 and 36), one crucial stipulation remains: the claim that modals do not have an *iEvent* feature (but they do have an *iPerf* when they are in the IPP construction). The lack of an *iEvent* feature is crucial to exclude the 2-3-1 order in double modal constructions, that is, one of the main generalizations. If modals had such a feature, it could not be explained why only the IPP construction allows the 2-3-1 order. The reason why one should be worried about this stipulation is that modals behave very much like main verbs regarding their event properties. For instance, they can be modified separately from the main verb, in particular by an event modifier such as *again*. With the right intonation and appropriate context, the sentence below allows three different interpretations, as indicated.

- (92)    dass    er    sie    wieder    besuchen    dürfen    will  
           that    he    her    again    visit        may        want  
           ‘that he again wants to visit her’  
           ‘since he wants to again be allowed to visit her’  
           ‘since he wants to be allowed to visit her again’

---

gered by the *uEvent* feature can be followed by movement of just VP3. As mentioned in fn. 35, Barbiers excludes the 3-1-2 order as a possible order of IPP constructions, and to derive this property, assumes that AUX does not come with a *uEvent* feature in this case. To avoid this stipulation and still derive the lack of the 3-1-2 order one could appeal to the order of Agree relations established by V1 and restrict movement to the goal of the first Agree relation. If V1 Agrees with the closer V2 first in (91b), movement of VP2 could be assumed to take preference over movement of VP3.

Barbiers states that “The exact label of the feature is immaterial. Any feature that distinguishes between modal/aspectual verbs and main verbs will do.” [Barbiers 2005: fn. 19]. If, however, the feature is simply some diacritic that only main verbs have, the question arising is why the two features, *iPerf* and *iEvent*, differ the way they differ. To motivate *iPerf* on modals in the IPP construction, semantic reasons are given. But the same semantic reasons would then favor an *iEvent* feature on modals. If the features bear no semantic relation to the elements involved, the restrictiveness and predictability of the approach is once again threatened. Why do modals have an *iPerf* but not an *iEvent* feature?

To close the discussion of the motivation of verb cluster movement, the account provided in Barbiers (2005) provides a very interesting new perspective. Barbiers explicitly acknowledges that optionality exists and that building this property into optional EPP features or the like, doesn’t solve the problem, but simply masks it, and he concludes: “Thus, Minimalist accounts of optional movement so far have failed to eliminate optionality from the grammar. [...] Rather than trying to eliminate optionality from the grammar, I propose that optionality of movement is an inherent property of the grammatical system. Movement does not have a trigger. It is freely available provided that it satisfies general conditions such as the ones discussed above. In a sense, then, this view on movement is a return to the Principle and Parameters (P&P) approach of movement (Chomsky 1981).” [Barbiers 2005: 254f] In light of the difficulties and challenges verb clusters have posed for almost any syntactic approach, it seems that this conclusion is something that researchers will have to consider in the future works on the topic.

#### 4. CONCLUSION—WHY DO VERBS CLUSTER?

This overview has shown that in the West-Germanic languages, constructions with multiple verbs such as auxiliaries and modals show variable word orders in different languages and constructions—verbs ‘cluster’ in ways that has been made precise in this overview. Verb cluster phenomena have been discussed in the generative literature since the 1970s and many interesting and important observations and generalizations have been put forward. Although the empirical distribution of possible and impossible word orders in VCs across West-Germanic is more complex than early works have assumed, our current knowledge of the micro- and macro-variation is fairly clear thanks to many detailed empirical studies (most noteworthy the SAND project on Dutch dialects). The theoretical tools available for deriving verb clusters are diverse, both within and across different frameworks, and the accounts on the market are generally successful in generating the different verb cluster configurations attested. Despite the significant empirical and theoretical advancements the research on verb clusters has made, there is, however, one issue which has not received a fully satisfactory treatment yet. As we have seen in the course of this overview, many interesting accounts have been suggested addressing the question of *how* verb clusters are derived. But what still appears to be an open question is the question of *why* the elements of a verb cluster are inverted in certain languages and constructions.

The question of why verbs cluster can be seen as a narrow or a broad question. The narrow question couches the answer as part of the syntactic mechanism(s) that derive the different word orders. While certain syntactic feature checking/valuation/licensing accounts appear to have an answer for why verbs cluster—there are features (i.e., lexical properties of certain elements) that make (certain) verbs do it—we have also seen that these accounts do not really explain why verbs cluster, at least not at a deep level, but they only shift the question to a question about why the features are distributed the way they are. It is obviously possible to derive any word order by

allowing features or linearization rules to be assigned as needed. But it is also clear that these feature assignments and rules do no more than restate the empirical facts, often on a case-by-case basis.

One may ask at this point whether there is really something for narrow grammar to explain here or whether the diverse empirical distribution is either simply the result of underivable lexical properties of certain elements or a consequence of a variety of grammatical (and not necessarily syntactic) and/or, even more broadly, extra-grammatical factors. The answer to this question may, as so often, lie somewhere in-between. The reason why one should perhaps not discard the syntactic relevance in regulating verb clusters is that there are generalizations and tendencies that can be found in the distribution of verb clusters which may indicate that there is a deeper ‘plan’ in the grammar of verb clusters. In particular, there are two generalizations that broadly characterize the distribution of verb clusters. Haider (2003, et seq.) notes that verb clustering is only found in OV languages, and he relates this to an unwelcome computational complexity created by head-final structures, which can be resolved by reordering. While there is a strong tendency of verb clusters being restricted to OV languages, it is not clear whether there is an implicational relation between the two properties. Hungarian and Yiddish may be languages that allow verb clusters, yet they are VO languages. And, of course, an OV order cannot be seen as a sufficient condition for verb clustering, given that verb clusters are basically only found in West-Germania. A further tendency concerns the possible orders in three-verb clusters: of the six logically possible combinations, five are attested robustly; the 2-1-3 order, however, is rare if existent at all as an unmarked order in a verb cluster. Barbiers (2005) and Abels (2011) provide accounts that can derive this distribution. If these tendencies can be maintained by further research, the grammar of clustering needs to be restricted in ways that limit the generation of (certain orders in) verb clusters to the attested orders and languages. While these two generalizations have consequences for how verb clusters are derived, they do not provide insights into the question what orders are possible in which language, which construction, and why. The following generalizations, however, may do so.

A popular approach in the 1980s and 1990s was to relate verb cluster reordering to the IPP effect, specifically to identify the trigger of different word orders with the IPP property. As we have seen in this overview, it is certainly true that across West-Germanic, most word order variation occurs in IPP constructions. However, as shown in Zwart (2007) (also already in Wurmbrand 2006), a causal relation between the IPP effect and reordering cannot be maintained. Neither is it the case that the IPP necessarily triggers reordering, nor is it the case that reordering triggers the IPP. Zwart (2007) does, however, note two important generalizations: i) IPP occurs without exception in 1-2-3 clusters; and ii) IPP is only found in languages with circumfixal participles (*ge*-V-*en*). This is, there is a one-way implication: if IPP then ‘*ge*’. Furthermore, we have seen in this overview that all verb cluster languages allow participles to precede auxiliaries. Most strikingly, this is the case even in languages such as Dutch or West Flemish that otherwise only allow English-type orders. Put differently, there is no language that allows the 2-1 order in modal-infinitive constructions, but not in auxiliary-participle constructions. Building on these generalizations, Wurmbrand (2012a) further observes that all ‘*ge*’-languages allow (many require) the participle to precede the auxiliary. Setting aside optionality as in Dutch, this fact can be taken to indicate that ‘*ge*’ participles and auxiliaries enter a syntactic dependency, triggered (directly or indirectly) by the properties of ‘*ge*’, that must be met in a configuration where the participle c-commands the auxiliary. As far as participles are concerned, a movement approach to reordering then has the potential of providing a configurational motivation for why auxiliary-participle con-

structions, but not modal-infinitive constructions, can (often have to) occur in the descending 2-1 order. If such a syntactic account can be provided for auxiliary-participle constructions, however, the question of why there is verb cluster reordering (and also VPR) in other contexts, and why only in certain constructions and languages, becomes a real question again, which, according to the current state-of-knowledge does not receive a satisfactory answer in narrow syntactic approaches.

Due to the highly varied and, at least from a syntactic point of view, seemingly arbitrary distribution of word orders, several researchers have therefore pursued multi-causal approaches. Lötscher (1978) provided one of the first multivariate approaches including grammatical, functional, and processing factors as triggers for different verb cluster and VPR orders. Recent multivariate approaches include Sapp (2006, 2011), Coussée (2008), Sutter (2009), and Dubenion-Smith (2010), and many important advances have been made in these approaches. For instance, based on extensive corpus and experimental data, Sapp (2006, 2011) has shown that among the factors showing statistical effects on the word order of two-verb clusters are: morphological and syntactic properties, the phonological shape of the elements involved, pragmatic properties, time period and dialects, as well as socio-linguistic factors such as occupation. Another factor singled out in Sapp's work (which is also stressed in Lötscher 1978, Schmid and Vogel 2004, and Dubenion-Smith 2010) is focus, which may point to a combination of prosodic, pragmatic, and possibly functional factors influencing the distribution of verb clusters. Furthermore, processing and parsing factors have already been brought into the picture in Lötscher's (1978) account, and been revived and extended as part of the explanation of verb clusters in Haider's (2003) directionality approach, Kaufmann's (2007) account of verb clusters in Mennonite Low German, as well as Dubenion-Smith's (2010) account of the statistical findings of the word order distribution in the Zwirner corpus of the West Central German dialects.

While there does not seem to be a simple answer to the question of what the deep motivation of verb cluster formation is and why this phenomenon only exists in certain languages, the decades-long research on this topic has brought to light many interesting data and generalizations, as well as numerous insightful theoretical observations about the (narrow and broad) grammatical properties of verb clusters as well as their relevance for several important issues of the general architecture of grammar.

## 5. REFERENCES

- Abels, Klaus. 2011. Hierarchy-order relations in the Germanic verb cluster and in the noun phrase. In *Groninger Arbeiten zur germanistischen Linguistik*, 1-28. Groningen: University of Groningen.
- Abels, Klaus. 2013. On 2-3-1. In *UCL Working Papers in Linguistics* 25, 67-101.
- Abels, Klaus, and Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15.1:25-74.
- Ackema, Peter. 2004. Do preverbs climb? In *Verb clusters: A study of Hungarian, German, and Dutch*, ed. by Katalin É. Kiss and Henk van Riemsdijk, 359-393. Amsterdam/Philadelphia: John Benjamins.
- Aissen, Judith, and David Perlmutter. 1976. Clause reduction in Spanish. In *Proceedings of the Second Annual Meeting of the Berkeley Linguistic Society*, ed. by Henry Thompson et al., 1-30. Berkeley: University of California, Berkeley Linguistic Society.
- Bader, Markus, and Tanja Schmid. 2009. Verb clusters in colloquial German. *Journal of Comparative Germanic Linguistics* 12.3:175-228.

- Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: The University of Chicago Press.
- Barbiers, Sjef. 2005. Word order variation in three-verb clusters and the division of labour between generative linguistics and sociolinguistics. In *Syntax and variation: Reconciling the biological and the social*, ed. by Leonie Cornips and Karen Corrigan, 233-264. Amsterdam: John Benjamins.
- 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 / Syntactic Atlas of the Dutch Dialects Volume II*. Amsterdam: Amsterdam University Press.
- Bech, Gunnar. 1955. *Studien zum deutschen Verbum infinitum*. Tübingen: Max Niemeyer Verlag.
- Bennis, Hans, and Teun Hoekstra. 1989. Why Kaatje was not heard sing a song. In *Sentential complementation and the lexicon: Studies in honour of Wim de Geest*, ed. by Dany Jaspers, Wim Klooster, Yvan Putseys and Pieter Seuren, 21-40. Dordrecht: Foris Publications.
- Besten, Hans den, and Hans Broekhuis. 1992. Verb projection raising in het Nederlands. *Spektator* 21:21-34.
- Besten, Hans den, and Jerold A. Edmondson. 1983. The verbal complex in continental West Germanic. In *On the formal syntax of the Westgermania*, ed. by Werner Abraham, 155-216. Amsterdam: John Benjamins.
- Besten, Hans den, and Corretje Moed-van Walraven. 1986. The syntax of verbs in Yiddish. In *Verb second phenomena in Germanic languages*, ed. by Hubert Haider and Martin Prinzhorn, 115-135. Dordrecht: Foris Publications.
- Besten, Hans den, and Jean Rutten. 1989. On verb raising, extraposition and free word order in Dutch. In *Sentential complementation and the lexicon: Studies in honour of Wim de Geest*, ed. by Dany Jaspers, Wim Klooster, Yvan Putseys and Pieter Seuren, 41-56. Dordrecht: Foris Publications.
- Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45.2:169-225.
- Bloemhoff, Henk. 1979. Heranalyse van een Stellingwerper oppervlaktestructuur. *Us Wurk: Tydskrift foar Frisistyk* 28:31-38.
- Blom, Elma, and Eric Hoekstra. 1996. IPP en werkwoordsvolgorde in het Achterhoeks. *Taal en Tongval* 48:72-83.
- Bobaljik, Jonathan D. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation, MIT, Cambridge, MA.
- Bobaljik, Jonathan D. 1998. Floating quantifiers: Handle with care. *Glott International* 3.6:3-10.
- Bobaljik, Jonathan D. 2004. Clustering theories. In *Verb clusters: A study of Hungarian, German, and Dutch*, ed. by Katalin É. Kiss and Henk van Riemsdijk, 121-145. Amsterdam/Philadelphia: John Benjamins.
- Bok-Bennema, Reineke, and Brigitte Kampers-Manhe. 1994. Transparency effects in the Romance languages. In *Issues and theory in Romance linguistics: Selected papers from the Linguistic Symposium on Romance Languages XXIII*, ed. by Michael Mazzola, 199-217. Washington, D.C.: Georgetown University Press.
- Broekhuis, Hans. 1992. Chain-government: Issues in Dutch syntax. Doctoral dissertation, University of Amsterdam.
- Broekhuis, Hans. 1993. Verb projection raising. *Spektator* 22:28-47.



- Broekhuis, Hans, Hans den Besten, Kees Hoekstra, and Jean Rutten. 1995. Infinitival complementation in Dutch: On remnant extraposition. *The Linguistic Review* 12:93-122.
- Büring, Daniel, and Katharina Hartmann. 1997a. Doing the right thing. *The Linguistic Review* 14:1-42.
- Büring, Daniel, and Katharina Hartmann. 1997b. The Kayne mutiny. In *Rightward movement*, ed. by Dorothee Beerman, David LeBlanc and Henk van Riemsdijk, 59-79. Amsterdam/Philadelphia: John Benjamins.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36.3:315-332.
- Coppen, P-A., and M. Klein. 1992. Het einde van verb raising. In *Nieuwe Eskapades in de Neerlandistiek: Opstellen van vrienden voor M.C. van der Toorn bij zijn afscheid als hoogleraar Nederlandse Taalkunde aan der Katholieke Universteit Nijmegen*, ed. by M. Klein. Groningen: Wolters-Noordhoff.
- Corrigan, Kate P. 2000. "What bees to be maun be": aspects of deontic and epistemic modality in a northern dialect of Irish English. *English World-Wide* 21.1:25-62.
- Coussée, Evie. 2008. Motivaties voor volgordevariatie: Een diachrone studie van werkwoordsvolgorde in het Nederlands. Doctoral dissertation, Ghent University, Ghent, Belgium.
- Diesing, Molly. 1990. Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory* 8:41-79.
- Diesing, Molly. 1996. Semantic variables and object shift. In *Studies in comparative Germanic syntax II*, ed. by Höskuldur Thráinsson, Samuel David Epstein and Steve Peter, 66-84. Dordrecht: Kluwer Academic Publishers.
- Diesing, Molly. 1997. Yiddish VP order and the typology of object movement in Germanic. *Natural Language and Linguistic Theory* 15.2:369-427.
- Dikken, Marcel den. 1989. Verb projection raising en de analyse van het IPP-effect. *Tabu* 19:59-75.
- Dikken, Marcel den. 1990. Verb incorporation in French causative constructions. Ms., University of Leiden. Leiden.
- Dikken, Marcel den. 1992. Particles. Doctoral dissertation, Holland Institute of Generative Linguistics, Leiden.
- Dikken, Marcel den. 1994. Minimalist verb (projection) raising. In *Minimalism and Kayne's asymmetry hypothesis*, ed. by Jan-Wouter C. Zwart, 71-88. Groninger Arbeiten zur germanistischen Linguistik 37. Groningen: University of Groningen.
- Dikken, Marcel den. 1995a. *Particles: On the syntax of verb-particle, triadic, and causative constructions*. New York/Oxford: Oxford University Press.
- Dikken, Marcel den. 1995b. Verb (projection) raising, scope, and uniform phrase structure. In *Proceedings of the North Eastern Linguistics Society Annual Meeting 25*, ed. by Jill N. Beckman, 95-110. Amherst: University of Massachusetts, GLSA.
- Dikken, Marcel den. 1996. The minimal links of verb (projection) raising. In *Minimal ideas: Syntactic studies in the Minimalist framework*, ed. by Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson and Jan-Wouter C. Zwart, 67-96. Amsterdam/Philadelphia: John Benjamins.

- Dikken, Marcel den. 2000. Verb clusters in Germanic and Hungarian: An overview. Plenary talk at the general meeting of the *Dutch-Hungarian Study Center on Language Contact*, The Netherlands Institute for Advanced Study, Wassenaar.
- Dikken, Marcel den, and Eric Hoekstra. 1997. Parasitic participles. *Linguistics* 35:1057-1089.
- Dubenion-Smith, Shannon A. 2010. Verbal complex phenomena in West Central German: Empirical domain and multi-causal account. *Journal of Germanic Linguistics* 22.02:99-191.
- É. Kiss, Katalin, and Henk van Riemsdijk eds. 2004. *Verb clusters: A study of Hungarian, German, and Dutch*. vol. 69. *Linguistik Aktuell/Linguistics Today*. Amsterdam/Philadelphia: John Benjamins.
- Erb, Marie Christine. 2001. Finite auxiliaries in German. Doctoral dissertation, University of Tilburg.
- Evers, Arnold. 1975a. The Guillotine principle. In *Linguistics in the Netherlands 1972-1973*, ed. by A. Kraak, 147. Amsterdam: Von Gorcu.
- Evers, Arnold. 1975b. The transformational cycle of Dutch and German. Doctoral dissertation, University of Utrecht.
- Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and Scope: a case for overt QR. In *Proceedings of the 18th West Coast Conference on Formal Linguistics*, ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen and Peter Norquest, 132-144. Somerville, MA: Cascadilla Press.
- Gonçalves, Anabela. 1998. On restructuring constructions in European Portuguese. In *Proceedings of the sixth Conference of the Student Organisation of Linguistics in Europe*, ed. by Tina Cambier-Langeveld, Anikó Lipták and Michael Redford, 75-88. Leiden: SOLE.
- Grewendorf, Günther, and Joachim Sabel. 1994. Long scrambling and incorporation. *Linguistic Inquiry* 25.2:263-308.
- Guasti, Maria Teresa. 1992. Incorporation, excorporation and lexical properties of causative heads. *The Linguistic Review* 8:209-232.
- Guasti, Maria Teresa. 1993. *Causatives and perception verbs: a comparative approach*. Turin: Rosenberg and Sellier.
- Guasti, Maria Teresa. 1996. Semantic restrictions on Romance causatives and the incorporation approach. *Linguistic Inquiry* 27.2:294-313.
- Guasti, Maria Teresa. 1997. Romance causatives. In *The new comparative syntax*, ed. by Liliane Haegeman, 124-144. London: Longmans Linguistics Library.
- Guéron, Jacqueline, and Teun Hoekstra. 1988. T-Chains and the constituent structure of auxiliaries. In *Constituent structure*, ed. by Anna Cardinaletti, Giuliana Giusti and Guglielmo Cinque, 35-100. Dordrecht: Foris Publications.
- Haan, Germen J. de. 1992. The verbal complex in Frisian. *Us Wurk: Tydskrift foar Frisistyk* 41.3-4:59-92.
- Haan, Germen J. de. 1993. The third construction in Frisian. In *Dialektsyntax*, ed. by Werner Abraham and Josef Bayer, 117-130. Opladen: Westdeutscher Verlag.
- Haan, Germen J. de. 1996. Recent changes in the verbal complex of Frisian. In *A Frisian and Germanic Miscellany. Published in Honour of Nils Århammar on his Sixty-Fifth Birthday, 7 August 1996*, ed. by Adeline Petersen and Hans F. Nielsen, 171-184. Nordfriisk Instituut, Bredsted: Odense University Press. Published as NOWELE 28-29.

- Haegeman, Liliane. 1990. The syntax of motional *goan* in WF. In *Linguistics in the Netherlands 1990*, ed. by Reineke Bok-Bennema and Peter Coopmans, 81-90. Dordrecht: Foris Publications Publications.
- Haegeman, Liliane. 1992. *Theory and description in generative syntax: A case study in West Flemish*. Cambridge: Cambridge University Press.
- Haegeman, Liliane. 1994. Verb raising as verb projection raising: Some empirical problems. *Linguistic Inquiry* 25:509-522.
- Haegeman, Liliane. 1995. IPP constructions and V-movement in West Flemish. In *Geneva Generative Papers*, ed. by Michal Starke, Eric Haeberli and Christopher Laenzlinger, 50-76. Geneva: Université de Genève.
- Haegeman, Liliane. 1998a. V-positions and the middle field in West Flemish. *Syntax* 1:259-299.
- Haegeman, Liliane. 1998b. Verb movement in embedded clauses in West Flemish. *Linguistic Inquiry* 29:631-656.
- Haegeman, Liliane, and Henk van Riemsdijk. 1986. Verb projection raising, scope, and the typology of rules affecting verbs. *Linguistic Inquiry* 17:417-466.
- Haider, Hubert. 2003. V-clustering and clause union: Causes and effects. In *Verb constructions in German and Dutch*, ed. by Pieter Seuren and Gerard Kempen, 91-126. Amsterdam/Philadelphia: John Benjamins.
- Haider, Hubert. 2011. Grammatische Illusionen: Lokal wohlgeformt – global deviant. *Zeitschrift für Sprachwissenschaft* 30.3:223-257.
- Hinterhölzl, Roland. 1997. A VP-based approach to verb raising. In *Proceedings of the North Eastern Linguistics Society Annual Meeting 27*, ed. by Kiyomi Kusumoto, 187-201. Amherst: University of Massachusetts, GLSA.
- Hinterhölzl, Roland. 1998. The syntax of IPP constructions and the structure of the lower middlefield in Westgermanic. In *Proceedings of the 22nd Annual Penn Linguistics Colloquium*, ed. by Alexis Dimitriadis, Hikyoung Lee, Christine Moisset and Alexander Williams, 59-70. Philadelphia: University of Pennsylvania.
- Hinterhölzl, Roland. 1999. Restructuring infinitives and the theory of complementation. Doctoral dissertation, University of Southern California, Los Angeles.
- Hoekstra, Eric, and Willem Taanman. 1996. Een West-Friese gradatie van het infinitivus-pro participio effect. *Nederlandse Taalkunde* 1:13-25.
- Hsiao, Franny P. 1999. A minimalist look at the German clause structure and parametric variation. M.A. Thesis. Toronto, Ontario: University of Toronto.
- Hsiao, Franny P. 2000. Relating syntactic variation in West Germanic to checking parameters. In *Proceedings of the eighth Conference of the Student Organisation of Linguistics in Europe*, ed. by Christine Czinglar, Katharina Köhler, Erica Thrift, Erik Jan van der Torre and Malte Zimmermann, 155-169. Leiden: SOLE.
- Ijbema, Aniek. 1997. Der IPP-Effekt im Deutschen und Niederländischen. In *Groninger Arbeiten zur germanistischen Linguistik* 40, ed. by Kleanthes K. Grohmann, 137-163. Groningen: University of Groningen.
- Kaan, Edith. 1992a. A minimalist approach to extraposition. M.A. Thesis: University of Groningen.
- Kaan, Edith. 1992b. A minimalist approach to extraposition of CP and verb (projection) raising. In *Yearbook of the research group for linguistic theory and knowledge representation*, ed. by D. Gilbers and S. Looyenge, 169-179. Language and Cognition 2. Groningen: University of Groningen.

- Kathol, Andreas. 1996. Order variability in German and Dutch verb clusters. In *Computational Linguistics in the Netherlands 1995*, ed. by Gert Durieux, Walter Daelemans and Steven Gillis, 147-164. Antwerp: UIA, University of Antwerp, Center for Dutch Language and Speech.
- Kathol, Andreas. 1998a. Constituency and linearization of verbal complexes. In *Complex predicates in nonderivational syntax*, ed. by Erhard Hinrichs, Andreas Kathol and Tsuneko Nakazawa, 221-270. San Diego, Calif: Academic Press.
- Kathol, Andreas. 1998b. Linearization of verb clusters in West Germanic. In *Proceedings of the 9th Western Conference on Linguistics*, ed. by Vida Samiian, 149-161. Fresno: California State University.
- Kayne, Richard. 1989. Null subjects and clitic climbing. In *The null subject parameter*, ed. by Osvaldo Jaeggli and Kenneth Safir, 239-261. Dordrecht: Kluwer Academic Publishers.
- Kayne, Richard. 1990. Romance clitics and PRO. In *Proceedings of the North Eastern Linguistics Society Annual Meeting 20*, ed. by Juli A. Carter, Rose-Marie Dechaine, William Philip and Tim D. Sherer, 255-302. Amherst: University of Massachusetts, GLSA.
- Kayne, Richard. 1991. Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22.4:647-686.
- Kayne, Richard. 1993. Toward a modular theory of auxiliary selection. *Studia Linguistica* 47:3-31.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Koopman, Hilda. 1984. *The syntax of verbs: From verb movement rules in Kru languages to Universal Grammar*. Dordrecht: Foris Publications.
- Koopman, Hilda. 1995. On verbs that fail to undergo V-second. *Linguistic Inquiry* 26:139-163.
- Koopman, Hilda. 1999. Derivations and representations: remnant movement and complexity filters. Talk given at the *Workshop on Remnant Movement, Feature Movement, and Their Implications on the T-Model*, Potsdam.
- Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
- Koster, Jan. 1994. Predicate incorporation and the word order of Dutch. In *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, ed. by Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi and Raffaella Zanuttini, 255-277. Washington, D.C.: Georgetown University Press.
- Lötscher, Andreas. 1978. Zur Verbstellung im Zürichdeutschen und in anderen Varianten des Deutschen. *Zeitschrift für Dialektologie und Linguistik* 45:1-29.
- Meer, G. van der. 1990. De Infinitivus Pro Participio: een nieuwe verklaring. *Tabu*:29-48.
- Merkes, Peter Wilhelm. 1895. Der neuhochdeutsche Infinitiv als Teil einer umschriebenen Zeitform: Historisch-grammatische Betrachtungen. Doctoral dissertation, University of Göttingen.
- Meurers, Walt Detmar. 2000. Lexical generalizations in the syntax of German non-finite constructions. Doctoral dissertation, University of Tübingen.
- Neeleman, Ad. 1994. Complex predicates. Doctoral dissertation, Utrecht University.
- Neeleman, Ad, and Fred Weerman. 1993. The balance between syntax and morphology: Dutch particles and resultatives. *Natural Language and Linguistic Theory* 11:433-475.
- Patocka, Franz. 1997. *Satzgliedstellung in den bairischen Dialekten Österreichs*. Frankfurt/Berlin/New York/Paris/Wien: Peter Lang.

- Reis, Marga. 1979. Ansätze zu einer realistischen Grammatik. In *Befund und Bedeutung. Zum Verhältnis von Empirie und Interpretation in Sprach- und Literaturwissenschaft.*, ed. by Klaus Grubmüller, Ernst Hellgardt, Heinrich Jellissen and Marga Reis, 1-21. Tübingen: Max Niemeyer Verlag.
- Richards, Norvin. 1997. What moves where when in which language? Doctoral dissertation, MIT, Cambridge, MA.
- Riemsdijk, Henk van. 1998. Head movement and adjacency. *Natural Language and Linguistic Theory* 16.3:633-678.
- Rizzi, Luigi. 1976. Ristrutturazione. In *Rivista di grammatica generativa*.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris Publications.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar: Handbook of generative syntax*, ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer Academic Publishers.
- Robbers, Karin. 1997. Non-finite verbal complements in Afrikaans: A comparative approach. Doctoral dissertation, University of Amsterdam, Den Haag.
- Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22:209-218.
- Roberts, Ian. 1993. Restructuring, pronoun movement and head-movement in Old French. Ms., University of Wales. Bangor.
- Roberts, Ian. 1997. Restructuring, head movement, and locality. *Linguistic Inquiry* 28:423-460.
- Rooryck, Johan. 1994. Against optional movement for clitic climbing. In *Issues and theory in Romance linguistics: Selected papers from the Linguistic Symposium on Romance Languages XXIII*, ed. by Michael Mazzola, 417-443. Washington, D.C.: Georgetown University Press.
- Rutten, Jean. 1991. Infinitival complements and auxiliaries. Doctoral dissertation, University of Amsterdam.
- Sabel, Joachim. 1996. *Restrukturierung und Lokalität: Universelle Beschränkungen für Wortstellungsvariationen*. Berlin: Akademie Verlag.
- Salzmann, Martin. 2013a. New arguments for verb cluster formation at PF and a right-branching VP: Evidence from verb doubling and cluster penetrability. *Linguistic Variation* 13:81-132.
- Salzmann, Martin. 2013b. New evidence for postsyntactic verb cluster formation and a right-branching base order. In *Proceedings of the North Eastern Linguistics Society Annual Meeting 42*, ed. by Stefan Keine and Shayne Sloggett, 135-148. Amherst: University of Massachusetts, GLSA.
- Salzmann, Martin. 2014. Displaced morphology in German and its implications for the theory of verb clusters. Paper presented at the *CRISSP Seminar, HUBrussel/KU Leuven*, Brussels.
- Sapp, Christopher D. 2006. Verb order in subordinate clauses: From Early New High German to Modern German. Doctoral dissertation, Indiana University.
- Sapp, Christopher D. 2011. *The verbal complex in subordinate clauses from Medieval to Modern German*. Amsterdam/Philadelphia: John Benjamins.
- Schallert, Oliver. 2014. *Zur Syntax der Ersatzinfinitivkonstruktion: Typologie und Variation*. Tübingen: Stauffenberg.
- Schmid, Tanja, and Ralf Vogel. 2004. Dialectal variation in German 3-verb clusters: A surface oriented OT account. *Journal of Comparative Germanic Linguistics* 7.3:235-274.
- Schönenberger, Manuela. 1995. Constituent order in the VP: Verb raising and verb projection raising. In *Topics in Swiss German syntax*, ed. by Zvi Penner, 347-411. Bern: Peter Lang.

- Sportiche, Dominique. 1996. A-reconstruction and constituent structure. Talk given at Université de Montréal, Montréal, Québec.
- Stechow, Arnim von, and Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens: Ein Lehrbuch der modernen generativen Grammatik*. Opladen: Westdeutscher Verlag.
- Sternefeld, Wolfgang. 1990. Scrambling and minimality. In *Scrambling and barriers*, ed. by Günther Grewendorf and Wolfgang Sternefeld, 239-257. Amsterdam/Philadelphia: John Benjamins.
- Sutter, Gert de. 2009. Towards a multivariate model of grammar: The case of word order variation in Dutch clause final verb clusters. In *Describing and modeling variation in grammar*, ed. by Andreas Dufter, Jürg Fleischer and Guido Seiler, 225-254: Walter de Gruyter.
- Svenonius, Peter. 2007. 1...3-2. In *The Oxford handbook of linguistic interfaces*, ed. by Gillian Ramchand and Charles Reiss, 239-288. Oxford/New York: Oxford University Press.
- Terzi, Arhonto. 1996. Clitic climbing from finite clauses and tense raising. *Probus* 8:273-295.
- Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral dissertation, MIT, Cambridge, MA.
- Vogel, Ralf. 2009. Skandal im Verbkomplex: Betrachtungen zur scheinbar inkorrekten Morphologie in infiniten Verbkomplexen des Deutschen. *Zeitschrift für Sprachwissenschaft* 28.2:307-346.
- Wiklund, Anna-Lena. 2005. The syntax of tenselessness: On copying constructions in Swedish. Doctoral dissertation, Umeå University.
- Wiklund, Anna-Lena. 2007. *The syntax of tenselessness: Tense/mood/aspect-agreeing infinitivals*. Berlin: Mouton de Gruyter.
- Williams, Edwin. 2004. The structure of clusters. In *Verb clusters: A study of Hungarian, German, and Dutch*, ed. by Katalin É. Kiss and Henk van Riemsdijk, 173-201. Amsterdam/Philadelphia: John Benjamins.
- Wurmbrand, Susi. 1999. The structure(s) of particle verbs. Ms., McGill University. Montréal, Québec.
- Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin/New York: Mouton de Gruyter.
- Wurmbrand, Susi. 2004a. Syntactic vs. post-syntactic movement. In *Proceedings of the 2003 Annual Conference of the Canadian Linguistic Association*, ed. by Sophie Burelle and Stanca Somesfalean, 284-295: Université du Québec à Montréal.
- Wurmbrand, Susi. 2004b. West Germanic verb clusters: The empirical domain. In *Verb clusters: A study of Hungarian, German, and Dutch*, ed. by Katalin É. Kiss and Henk van Riemsdijk, 43-85. Amsterdam/Philadelphia: John Benjamins.
- Wurmbrand, Susi. 2006. Verb clusters, verb raising, and restructuring. In *The Blackwell Companion to Syntax*, ed. by Martin Everaert and Henk van Riemsdijk, Vol. V, 227-341. Oxford: Blackwell.
- Wurmbrand, Susi. 2007. How complex are complex predicates? *Syntax* 10:243-288.
- Wurmbrand, Susi. 2010. Parasitic morphology in Germanic: Consequences for the theory of feature checking. Storrs: University of Connecticut.
- Wurmbrand, Susi. 2012a. Parasitic participles in Germanic: Evidence for the theory of verb clusters. *Taal en Tongval* 64.1:129-156.

- Wurmbrand, Susi. 2012b. The syntax of valuation in auxiliary-participle constructions. In *Coyote Working Papers: Proceedings of the 29th West Coast Conference on Formal Linguistics (WCCFL 29)*, ed. by Jaehoon Choi et al. University of Arizona: Tucson.
- Wurmbrand, Susi. 2013. Complex predicate formation via voice incorporation. Ms., University of Connecticut. Storrs.
- Wurmbrand, Susi. 2014. Tense and aspect in English infinitives. *Linguistic Inquiry* 45.3:403-447.
- Wurmbrand, Susi. To appear. Restructuring across the world. In *Language Use and Linguistic Structure: Proceedings of the Olomouc Linguistics Colloquium 2014*, ed. by Ludmila Veselovská and Markéta Janebová. Olomouc: Palacký University.
- Wurmbrand, Susi, and Jonathan D. Bobaljik. 2005. Adjacency, PF, and extraposition. In *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, ed. by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz and Jan Koster, 679-688. Berlin/New York: Mouton de Gruyter.
- Wyngaerd, Guido Vanden. 1989. Verb projection raising and the status of infinitival complements. In *Sentential complementation and the lexicon: Studies in honour of Wim de Geest*, ed. by Dany Jaspers, Wim Klooster, Yvan Putseys and Pieter Seuren, 423-438. Dordrecht: Foris Publications.
- Zwart, C. Jan-Wouter. 1995. A note on verb clusters in the Stellingwerf dialect. In *Linguistics in the Netherlands 1995*, ed. by Marcel den Dikken and Kees Hengeveld, 215-226. Amsterdam: John Benjamins.
- Zwart, C. Jan-Wouter. 1996. Verb clusters in continental West Germanic dialects. In *Microparametric syntax and dialect variation*, ed. by James Black and Virginia Motapanyane, 229-258. Amsterdam/Philadelphia: John Benjamins.
- Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement*. Dordrecht: Kluwer Academic Publishers.
- Zwart, C. Jan-Wouter. 2007. Some notes on the origin and distribution of the IPP-effect. In *Groninger Arbeiten zur germanistischen Linguistik* 45, 77-99. Groningen: University of Groningen.

Susi Wurmbrand  
 susi@alum.mit.edu  
 wurmbrand.uconn.edu