

1.3 Coherence/restructuring effects

Even though the semantics of many verb cluster constructions, especially those involving modals, may seem semantically bi- or multi-clausal (cf. the English translation), they actually behave like a monoclausal syntactic unit, i.e. they show so-called restructuring effects = instantiate a so-called coherent construction. The literature contains a plethora of such effects such as a single negation domain, adverbs having scope over the entire sentence etc. (cf. e.g. Haider (2010: 310-321) for a list of such effects in German). I will illustrate the monoclausality by scrambling. It is well-known that German lacks long-distance scrambling, i.e. scrambling across a finite clause boundary:

- (5) * dass [das Buch]_I keiner will, [CP dass ich ____I lese]
that the book no.one wants that I read.1SG
'that no one wants that I read the book' (Standard German)

In contrast, scrambling from the complement of modals and auxiliaries and so-called restructuring verbs (certain lexical verbs which take an infinitival complement with the infinitival particle *zu* 'to') is licit (to remain theory-neutral at this point, I gloss the non-finite complement as XP):

- (6) dass [TP [das Buch]_I keiner [XP ____I lesen] wollte]
that the book no.one read.INF wanted.3SG
'that no one wanted to read the book' (Standard German)

Since the embedded object occurs above the matrix subject, one can be sure that it has left the embedded clause.

1.4 Outline

In this paper, I will argue that there is derivational complex head formation to account for the cluster property. The account differs from previous work in that I assume that the cluster does not arise in syntax but at PF through an operation that involves adjacent verbal terminals. This operation is also responsible for the various orders, which means that the cluster property and the ordering property are directly related. However, I assume that the coherence property does not follow from cluster formation. Rather, I assume that it is due to the fact that the non-finite complement contains less structure.

The paper is organized as follows: In section two, I introduce the empirical domain where the crucial evidence for the PF-sensitivity of cluster formation can be found, viz. Swiss German verb doubling, where a crucial asymmetry with respect to penetrability obtains between ascending and descending orders. The proposal is developed in section three; it contains an analysis of the motion verb construction and two additional arguments in favor of post-syntactic cluster formation, viz. the cluster puzzle and the cross-linguistic predictions with respect to cluster order and penetrability. This section also includes a discussion of previous literature and a few loose ends. Section four concludes.

- (9) a. *Ich ha grad [wele s hole] b. Ich ha s₁ grad [wele __₁ hole]
 I have just wanted it get.INF I have it just wanted get.INF
 ‘I just wanted to get it.’ (Swiss German)

This would be unexpected if *go* were e.g. a complementizer or a preposition because such elements usually block restructuring; cf. e.g. the following example from Dutch where the presence of the complementizer *om* blocks scrambling of the embedded object to the matrix clause (in this case, scrambling takes place from an – apparently extraposed – infinitival clause containing the particle *te* ‘to’ so that the resulting structure instantiates the so-called *third construction*, cf. also fn. 20 and 3.4.1; cf. Broekhuis et al. (1995: 102):

- (10) Jan heeft dat boek geprobeerd (*om) te lezen.
 John has the book tried for to read.INF
 ‘John tried to read that book.’ (Dutch)

Second, as with non-finite modals, the position of *go* is variable within the VP, both under verb second, (11)/(13), and V-final, (12)/(14), thus instantiating Verb Projection Raising (VPR: the a and b examples) and Verb Raising (VR: the c-examples), cf. also Schönenberger and Penner (1995a: 289):⁴

- (11) a. Ich gang [go de Muetter es Buech chauffe].
 b. Ich gang [de Muetter go es Buech chauffe].
 c. Ich gang [de Muetter es Buech go chauffe].
 I go.1s PRT the.DAT mother PRT a book PRT buy.INF
 ‘I’ll go buy a book for the mother.’
- (12) a. dass i gang [go de Muetter es Buech chauffe]
 b. dass i gang [de Muetter go es Buech chauffe]
 c. dass i gang [de Muetter es Buech go chauffe]
 that I go.1s PRT the.DAT mother PRT a book PRT buy.INF
- (13) a. Ich ha [wele de Muetter es Buech chauffe].
 b. Ich ha [de Muetter wele es Buech chauffe].
 c. Ich ha [de Muetter es Buech wele chauffe].
 I have.1s wanted the.DAT mother wanted a book wanted buy.INF
 ‘I wanted to buy the mother a book.’

⁴ The choice between the VR- or the VPR-structure is largely governed by information-structural considerations (non-verbal material in the c-command domain of the particle/the modal is necessarily focal, usually entailing wide focus, while material outside that domain can be presuppositional or focal), cf. Lötscher (1978), Lötscher (1993), Brandner and Salzmann (2009: 94-98).

New arguments for verb cluster formation at PF

- (14) a. dass i ha [wele de Muetter es Buech chauffe]
 b. dass i ha [de Muetter wele es Buech chauffe]
 c. dass i ha [de Muetter es Buech wele chauffe]
 that I have.1s wanted the.DAT mother wanted a book wanted buy.INF]

The possibility of VPR argues against treating *go* as an infinitival particle like *te/zu* ‘to’ which always occurs immediately preverbally in German/Dutch. Evidence that *go* heads its own projection can be seen in the following example where *go* is topicalized with elements preceding it:

- (15) [De Muetter go es Buech chauffe] gang i nöd.
 the.DAT mother PRT a book buy.INF go.1s I not (Swiss German)

I will therefore treat *go* as a non-finite verb heading its own VP-projection.⁵ Furthermore, I make the same assumption for the finite matrix motion verb. Each thus projects its own verb phrase (even though only one of them has lexical content). I thus differ from Schönenberger and Penner (1995b) who treat the finite form as an expletive; see Brandner and Salzmann (2012: 90) for arguments against that position. Note that on my analysis *go* is not the spell-out of a lower copy of the movement chain involving the finite ‘go’ as might have been expected given that doubling phenomena are often analyzed this way (cf. the contributions in Barbiers et al. (2008a)). A spell-out analysis is unattractive for the following reasons: First, *go* occurs after various motion verbs, including *räne* ‘run’, *lauffe* ‘run’, *springe* ‘run’ or *schicke* ‘send’, cf. Weber (1987: 245):

- (16) D Muetter hät mi gschickt go poschte.
 The mother has me sent PRT do.shopping
 ‘My mother sent me to do grocery shopping.’

A spell-out approach would then have to resort to syntactic decomposition of motion verbs (with *go* just spelling out the abstract motion component) for which there is little independent motivation in Swiss German. Second, in many dialects (e.g. Zurich German), the form of *go* does not correspond to any of the regular inflectional forms of the verb ‘go’, it is e.g. not identical to the infinitive (which would be *gaa* in ZG); rather, the form is simply fossilized, suggesting that doubling is no longer a productive process. Third, and most importantly (see also van Riemsdijk (2002: 160, fn. 22)), it is unclear which intermediate positions would be spelled-out. While verb second movement may generate a copy in T, no copies are expected to arise in V-final structures like (12) because on standard assumptions the verb stays in its base-position (or only moves up to v).^{6, 7}

⁵ There is reason to believe that *go* really just projects to VP, not to vP, cf. Brandner and Salzmann (2012).

⁶ Another argument against treating *go* as the spell-out of a trace comes from examples where *go* occurs multiply within a single sentence (both V2 and V-final), cf. also Weber (1987: 246) and Suter (1976: 150):

(i) ... das mer gönd go en guete Platz go sueche
 that we go.1PL PRT a good seat PRT search.INF
 ‘that we go look for a good seat’

At least in V-final structures, there will simply not be two intermediate copies of the verb. A proper analysis of such examples remains difficult though; see Brandner and Salzmann (2012: 90f.) for discussion.

The same facts can also be illustrated with the causative *laa* ‘let’, which takes a double ‘la’ (I will disregard *aafa* because it involves a number of additional complications that are orthogonal to the point to be made):

- (17) Er laat de Hans **la** en Arie singe.
 he lets the.ACC John PRT an aria sing.INF
 ‘He lets John sing an aria.’

As with *go/cho*, restructuring effects are obligatory:

- (18) Er laat **s** de Peter la ___ chauffe b. *Er laat de Peter la **s** chauffe
 he lets it the.ACC Peter PRT buy.INF He lets the.ACC Peter PRT it buy.INF
 ‘He lets Peter buy it.’

Second, we find both VR- and VPR-structures, both in verb-second and V-final structures:

- (19)a. Er laat de Peter [**la** de Unggle bsueche].
 b. Er laat de Peter [de Unggle **la** bsueche].
 He lets the Peter PRT the uncle PRT visit.INF
 ‘He lets Peter visit the uncle.’ (Swiss German)

- (20)a. dass er de Peter laat [**la** de Unggle bsueche]
 b. dass er de Peter laat [de Unggle **la** bsueche]
 that he the Peter lets PRT the uncle PRT visit.INF
 ‘that he lets Peter visit the uncle’ (Swiss German)

⁷ Next to the uses in the text where *go* has the flavor of a purpose marker it also occurs in contexts where one would rather classify it as an infinitive marker, viz. in extraposed subject or object sentences or in complements of nouns, cf. Burgmeier (2008: 49):

(i) die händ dōch chèn Aaspruch, go sāge, es seg es Unrächt ...
 they have.3PL PRT no right PRT say.INF it is a injustice
 ‘They don’t have the right to say that this is unjust ...’ (Swiss TV SF1, „Der Club“, August 29, 2006)

Given that *go* probably goes back to the directional preposition *gen* ‘towards’ cf. Lötscher (1993), Brandner and Salzmann (2012) and fn. 31, the co-existence of these uses is little surprising: The development from a directional preposition via a purpose marker to an infinitival marker represents an unmarked grammaticalization path (cf. Heine & Kuteva 2002: 39f., 247f.). *Go* would thus be parallel to *zu(m)* where all three functions coexist in Alemannic (the directional use of *go* has become rare, but it is still attested in some areas, especially in eastern Switzerland, Bodensee-Alemannic and Swabian). There is a complication, though, in that *go* can co-occur with the other infinitive marker *z*, as e.g. in the following example, cf. Weber (1987: 307):

(ii) Es isch nōd schön, d Lüt eso go z verschrecke
 it is not nice the people that PRT to scare.INF
 ‘It’s not nice to scare people that much.’

Analyzing *go+z* as one complex infinitival marker is implausible because (a) both can also occur on their own and (b) they can be separated by other material, cf. e.g. Staub et al. (1881: 2, 323), as cited in Burgmeier (2006: 47):

(iii) Es ist schlecht, d’ Lüt go a(z’)lüge
 it is bad the people PRT lie.to
 ‘It is bad to lie to people.’

In some of the examples found in the literature, *go* might be licensed by a deleted/silent motion verb (which is a general possibility, cf. van Riemsdijk 2002), but at least in the examples at hand that does not seem to be very plausible. I will leave these issues for further research. A good data source for such questions in Burgmeier (2006).

Given these properties, I will analyze *la* as a verbal element as well.

2.2 The puzzle: descending orders

In all the examples we have looked at so far, the finite verb occurred at the beginning of the verb cluster (even in V-final structures), which constitutes a 123 order if the double *go/cho/la* is classified as V₂. Interestingly, the finite verb can also occur at the end of the verb cluster while *go* still precedes the infinitive, leading to a 231 order:

(21)a. *dass i go₂ poschte₃ gang₁ 231*
 that I PRT do.shopping.INF go.1SG
 ‘that I go to do shopping’

b. *dass i gang₁ go₂ poschte₃ 123*
 that I go.1sg PRT do.shopping.INF
 ‘that I go to do shopping’

(Swiss German)

This is in itself an interesting fact since 231 orders (and the alternation with 123) is otherwise only systematically found with perception verbs, benefactives and inchoatives in Swiss German, according to Löttscher (1978: 3, fn. 2), Schmid (2005: 76), while other cluster types of the language usually only allow the orders 123, 132 or 312 (and perhaps 321), cf. Seiler (2004), Wurmbrand (2004a). Here is an example with a perception verb:⁸

⁸ It needs to be stressed that the 231 order with verb doubling is completely unmarked (i.e. just as unmarked as 123) while 231 orders with inchoatives/perception verbs/benefactives are usually marked for most speakers and never seem to be the preferred order (which is usually either 213 or 123). In a dialect syntax project carried out at the University of Zurich (Glaser (in preparation), 231 orders were accepted by rather few informants even with cluster types that should admit 231 according to Löttscher (1978) and Schmid (2005). The only clearly dissenting voice in this area is Cooper (1995: 154), who considers 231 and even 213 grammatical with Mod-Mod-Inf clusters in Zurich German, but this clashes with the rest of the literature and the empirical results in Glaser (in preparation).

In his survey of cluster orders in Dutch dialects, Barbiers (2005) finds 231 orders in Aux-Mod-Inf-clusters (‘has want read’), accepted by 9,7% of the speakers, and with rather high frequency (40,8%) in clusters he classifies as Aux-Asp-Inf clusters (Barbiers and Bennis (2010: 36, fn. 11) point out that the difference in acceptability may be due to methodological reasons; see also Haegeman and Oosterhof (2012) for further empirical details on 231 in Flemish):

i. *gaan zwemmen is.*
 go.INF swim.INF is
 ‘went swimming’

These may be instances of the motion verb construction where the finite has been elided (in Swiss German equivalents, the perfective motion verb can be present, in Flemish it cannot according to my informants); according to Karen De Clercq (p.c.), in Flemish, where this construction is most frequent, the form of the motion verb in such examples is different from the regular infinitive. It arguably represents a particle similar to *go*, cf. fn. 31. If this is correct, a large part of the 231 orders found in Dutch can then be related to a special element found in this construction, viz. a special form of ‘go’. The 231 cases with Aux-Mod-Inf orders, however, are not amenable to such an explanation, the same holds for 231 in Afrikaans, where it constitutes the neutral order, cf. Biberauer (2013). In other words, then, 231 orders cannot be set aside.

Note incidentally that 231 orders violate the Final-over-Final-Constraint formulated in Biberauer et al. (2008), another indication that such orders are special. See Biberauer and Walkden (2010) and Biberauer (2013) for discussion from the perspective of the FOFC.

- (22)a. ... wil er en ghöore₂ choo₃ hät₁
 because he him hear.INF come.INF has
 ‘... because he heard him come’ Lötscher (1978: 3, fn. 2)

But what is crucial is the following observation: in 231 orders, *go* must immediately precede the infinitive, VPR is not a possibility. There is thus an asymmetry between ascending (i.e. 123) and descending (231) orders, compare the following example with (12) above, repeated here as (24):⁹

- (23) a. *dass i [go₂ de Muetter es Buech chauffe₃] gang₁ (231)
 b. *?dass i [de Muetter go₂ es Buech chauffe₃] gang₁
 c. dass i [de Muetter es Buech go₂ chauffe₃] gang₁
 that I PRT the.DAT mother PRT a book PRT buy.INF go.1s
- (24) a. dass i gang₁ [go₂ de Muetter es Buech chauffe₃] (123)
 b. dass i gang₁ [de Muetter go₂ es Buech chauffe₃]
 c. dass i gang₁ [de Muetter es Buech go₂ chauffe₃]
 that I go.1s PRT the.DAT mother PRT a book PRT buy.INF

We find the same pattern with *la*: In principle, both 123 and 231 orders are possible:

- (25)a. dass er si laat₁ la₂ gaa₃ b. das er si la₂ gaa₃ laat₁
 that he her lets PRT go.INF that he her PRT go.INF lets
 ‘that he lets her go’

But as with *go*, VPR is blocked in 231 orders. Compare the following example with (20) from above, repeated here as (27):

- (26)a. *dass er de Peter [la₂ de Unggle bsueche₃] laat₁ (231)
 b. dass er de Peter [de Unggle la₂ bsueche₃] laat₁
 that he the Peter PRT the uncle PRT visit.INF lets
 ‘that he lets Peter visit the uncle’ (Swiss German)
- (27)a. dass er de Peter laat₁ [la₂ de Unggle bsueche₃] (123)
 b. dass er de Peter laat₁ [de Unggle la₂ bsueche₃]
 that he the Peter lets PRT the uncle PRT visit.INF (Swiss German)

⁹ The crucial empirical data are drawn from a questionnaire study (using a 5-point Likert-scale with 1 = best and 5 = worst) that I carried out at the University of Zurich in spring 2008. The informants were students of German linguistics in their second semester. While the a- and b-sentences received scores around 4 on average (with b receiving slightly higher results), the c-examples were rated as nearly perfect by most speakers (average = 1,5). No such contrast was found in the ascending examples even though speakers had a slight preference for the VR-structure (which is rather unsurprising given that VPR-structures are information-structurally more marked and thus less easy to contextualize in an experiment).

As far as I can tell, this also holds for 231 orders with other verb types in Swiss German; unfortunately, such examples were not part of the questionnaire, but I have not found any conflicting observations in the literature and the native speakers I have consulted all judged VPR with 231 as strongly deviant (and certainly clearly worse than 231 orders without VPR). Schallert (2012: 267, 279, 280, 285) makes the same observation for the Alemannic dialect spoken in Vorarlberg: while 231 orders are marginally possible with benefactives, inchoatives and perception verbs, VPR is ruled out in this order (but available with 123, 132 and 213). This asymmetry motivates my proposal in the next subsection.¹⁰

3 The proposal: post-syntactic complex head formation under adjacency

3.1 Assumptions

3.1.1 Locus and properties of cluster formation

In what follows I will presuppose an architecture of the PF-component as in Embick and Noyer (2001: 566) and Embick (2007) where hierarchical syntactic structure is gradually transformed into a linear string. During the mapping from syntax to a structure interpretable for the phonetic form, movement operations of different kinds take place. They have in common that their properties differ in fundamental ways from movement operations within the syntactic component. The earlier such operations apply, the more syntax-like they will be. For instance, the operation Lowering involves adjunction of a head to the head of its complement and thus refers to hierarchical structure (but since it goes downward, it cannot take place in syntax proper). The operation Local Dislocation, on the other hand, which also affixes/adjoins a head X onto another head Y, applies at a later stage in the PF-component and is sensitive to linear precedence and adjacency. In what follows, I will argue that verb cluster formation + inversion is a late PF-process akin to Local Dislocation.

As in much recent work, such a model assumes that linear order is established post-syntactically. Since linear order plays a crucial role in the syntax of verb clusters, I will briefly spell out my assumptions: Linearization in the present context involves two steps: There is an initial linearization that regulates the ordering of specifiers, heads and complements. In a second (later) step, the verbal elements can be reordered by means of cluster formation. As for the initial linearization, I do not adopt an LCA-based algorithm as proposed in Kayne (1994) where asymmetric c-command automatically implies precedence. Instead, I opt for a more flexible system as in Abels and Neeleman (2012: 66-68) that incorporates ordering statements for all sisterhood relationships. Under such an approach (cf. also Richards (2008)), Merge is seen as essentially symmetric and therefore in need of ordering; ordering statements can refer to both properties of the head and properties of the non-head (see also Cooper (1995), Barbiers (2000), Schmid and Vogel (2004)). The latter is particularly advantageous to express ordering generalizations that depend on the syntactic category of the non-head (rather than its semantic

¹⁰ The same asymmetry between 123 and 231 is found in Afrikaans according to Biberauer (2013) (see also the contrast with *te*-complements noted in Robbers (1997: 93)). The only elements that can intervene between V2 and V3 involve single words most of which can probably be interpreted as being incorporated (cf. also the data in den Dikken (1994: 83), Robbers (1997: 75, ex 80a; 79: ex. 93b)) so that the structures do not instantiate real VPR; the same holds for Swiss German, cf. examples (47)–(49) below. Note finally that the contrast is not as strong for some speakers because VPR is restricted for many even under 123 according to Marc de Vos (p.c.).

function). A relevant example provided by Abels and Neeleman (2012) is the following: In Dutch, as in English, a number of different adverbial and argumental functions can be realized either as a PP or as an adverbial phrase or DP. Cutting across these functions is the generalization that prepositional phrases can appear either pre- or postverbally, whereas adverbs and DPs are rigidly preverbal. Reference to syntactic category of these non-heads allows for a very simple formulation of the generalization. The possibility to refer to the syntactic category of the non-head also has advantages for linearizing arguments of verbs in languages like German and Dutch. While nominal complements always appear to the left of the verb in West-Germanic languages, verbal complements often and finite clausal complements (nearly) always appear to the right. It is certainly correct (as stressed by one of the anonymous reviewers) that LCA-based approaches (or more generally: approaches with a uniform base order) can derive the generalization by making certain movement operations sensitive to syntactic category (e.g. DP-movement to the left under a VO-order as in Zwart (1994) or VP-/CP-extrapolation under an OV-order as in the more traditional approaches). But in at least some of these cases (e.g. the DP-arguments), there is little independent evidence for the movement operation (such as reconstruction or freezing effects etc.); the putative base-position can only be inferred from theory-internal considerations like the LCA. A system with linearization statements that can refer to properties of the non-head is more parsimonious in that it can avoid movement operations in exactly those cases where a non-head always surfaces in a non-LCA-compatible position; cf. Zwart (2011: 247-280) for an overview over the discussion about Dutch.¹¹

I will consequently make the following assumptions about linearization in Swiss German (and West-Germanic more generally): Specifiers are uniformly linearized before the head of the same projection. As for complements: In the verbal domain, VPs and extended projections of V (vP, TP, CP) are linearized to the right of their selecting V-head while NPs/DPs are linearized to the left.¹² The default is thus a right-branching VP-structure, which is possible as a surface form in Dutch and Swiss German but not e.g. in Standard German where additional operations apply, cf. 3.2 below. This part of linearization takes place at an early point of the PF-derivation.

The order of the verbal terminals is determined after the relative order of specifier, head and complement (and modifiers) is established. If nothing happens, the verbal elements appear in a 123 order. (Partially) descending orders are derived via cluster formation at PF. This involves complex head formation (= re-bracketing) + inversion of verbal terminals under adjacency; this operation thus essentially represents an instance of *local dislocation* as described in Embick and Noyer (2001) and Embick (2007).¹³ Before we turn to verb clusters, it is instructive to look at a canonical instance of

¹¹ As the reader will notice in section 3.1.4 below, the right-branching approach adopted here requires quite a bit of evacuation movement to derive VR-structures. At first sight, this seems to be faced with the same criticism that was raised against LCA-based systems above. However, there is one crucial difference: In most varieties, these movement operations are optional, with non-movement resulting in VPR; in other words, there is independent evidence for a base-position to the right of the matrix verb.

¹² Since I do not adopt an LCA-based system, modifiers can in principle also be linearized to the left or to the right. This means that right-adjunction, which may be necessary anyway for extraposition, is a possibility.

¹³ My analysis bears certain similarities to that by Cooper (1995: 192-197) who assumes that cluster formation takes place “somewhere between S-structure and PF if not at PF” and postulates adjacency as a precondition. Unfortunately, the precise properties of the mechanism do not become clear to me, i.e. whether it is an instance of head-movement or

local dislocation, viz. the Latin clitic conjunction *-que*. Even though it conjoins two conjuncts, it does not appear between the two conjuncts but rather after the first (morphosyntactic) word of the second conjunct. Schematically, what happens is the following:

- (28)a. Input: (Conjunct1 X Y) *-que* (Conjunct2 W Z)
 b. Surface: (Conjunct1 X Y) __ (Conjunct2 W-*que* Z)

In the following example *-que* conjoins two DPs and ends up attached to the adjective of the second conjunct; the placement of *-que* is thus sensitive to linear order:

- (29) a. [[boni pueri] *-que* [bonae puellae] →
 good.NOM.PL.MSC boys and good.NOM.PL.FEM girls
 ‘good boys and good girls’
 b. boni pueri bonae=*que* puellae
 good.NOM.PL.MSC boys good.NOM.PL.FEM=*and* girls

In the case of verb clusters, a verb is adjoined to another and inverted. Importantly, what is inverted here are not syntactic sisters (i.e. V and VP, as e.g. in Haegeman and van Riemsdijk (1986) or Wurmbrand (2004b)), but linearly adjacent verbs. Without any further qualifications, this mechanism can derive the following orders (apart from 123): 132 (V₃ inverts with V₂: 1[32]), 213 (V₂ inverts with V₁: [21]3), and 321 (V₃ inverts with V₂ and [V₃V₂] invert with V₁: [[32]1]). In the last case, the second cluster forming operation thus involves a complex element, viz. V₃+V₂ that inverts with V₁. This is in accordance with the conditions described in Embick and Noyer (2001: 577) according to which segments of complex heads (so-called subwords) can only be involved in Local Dislocation if the target is also a subword. Since V₁ is an independent element, it does not count as a subword (but rather as a so-called morphosyntactic word). Therefore, Local Dislocation cannot involve only V₃ but has to involve the entire complex [V₃+V₂].

So far, the mechanism has the same effects as leftward head-movement from a right-branching base. We will see in section 3.2 how 231 orders come about in this system. In section 3.4.1 I will show how the entire range of attested orders in West-Germanic can be derived with this mechanism. It is well-known that West-Germanic varieties differ from each other with respect to the possible cluster orders. In what follows, I will capture these differences by means of linearization constraints (similar to the inversion parameters in Haegeman and van Riemsdijk (1986), and the lexical entries in Bader and Schmid (2009a)). A simple example would be a constraint requiring participles to precede auxiliaries

reanalysis and whether what is inverted are linearly adjacent verbal terminals or structural sisters. Her quote on p. 205 is rather opaque: “verb cluster formation [...] results in a constituent that is not compatible with the standard X’-theoretical assumptions about structure. It remains to be investigated whether clustering and inversion are general syntactic phenomena.”

they depend on.¹⁴ An alternative that locates the variation outside the grammar will be discussed in 3.4.4.

3.1.2 *Coherence/restructuring effects*

What is arguably special about this approach is that the clustering + reordering-operation is dissociated from syntax. As a consequence, unlike in other approaches (e.g. Haider (2003)), cluster formation cannot be made responsible for what are usually called restructuring effects or effects of monosententiality that arise with verb clusters, cf. e.g. the obligatory fronting of weak pronouns in (8)–(9). Rather, I adopt the widespread assumption that these effects follow from the fact that such complements are generated with less structure (as e.g. in den Dikken (1995), Wurmbrand (2007)). For instance, a modal only takes a VP/vP as its complement, which does not have a proper middle-field that could host weak pronouns. As a consequence, the entire structure behaves like a monoclausal unit even though one is not dealing with a complex head/verb (in syntax). The small size will also account for other monoclausal properties like the lack of a separate negation domain etc. Non-finite complements without restructuring effects, on the other hand, (so-called non-coherent infinitives) are taken to be larger in size, i.e. CPs, so that they can host their own negation, adverbials etc. and arguments cannot scramble out of them. As a consequence, they behave more like independent clauses.^{15, 16}

One might object here that under the present analysis, it is completely accidental that (in varieties like Standard German) verbs that take small VP-complements also have to undergo cluster formation + inversion while in approaches based on complex heads (e.g. Haider (2003)) monosententiality is a direct consequence of cluster formation. This may seem like a drawback, but there is evidence that restructuring effects can obtain independently of complex head formation. In all these cases, the restructuring verb/modal and the dependent non-finite verb are not adjacent. Wurmbrand (2007: 263f.) presents examples of long-distance passive with part of the non-finite clause being in clause-final/extrapolated position. Other examples include scrambling and pronoun fronting in Verb Projection Raising. Here is an example from West Flemish with long pronoun fronting:

¹⁴ Since at the level where cluster orders are generated the hierarchical syntactic structure is partially lost and embedding relations among verbs are no longer directly visible, the linearization constraints have to be formulated as correspondence constraints that relate the linear structure to the hierarchical one; a constraint like *Inf >> Mod* then implies that an infinitive has to precede at PF the modal verb it was c-commanded by in the hierarchical structure that formed the input for PF-computations. This seems essentially the type of constraint used in Schmid and Vogel (2004).

¹⁵ The exact size of restructuring/coherent non-finite complements is a matter of debate. What everyone seems to agree on is that they cannot be CPs, but just about any size between VP and TP has been proposed in recent work, cf. e.g. Wurmbrand (2004c) for VP and vP and den Dikken (1996) and Salzmann (2011) for TP. Furthermore, there is reason to believe that not all transparent complements are of equal size, cf. ter Beek (2008) for detailed discussion. Ter Beek additionally argues that it is not the size of a complement as such that determines transparency but rather its *phase-structure* (with transparency effects only obtaining in the absence of a phase-head in the non-finite complement). Since these issues are orthogonal to the questions pursued in this paper, I will label the projections as VPs for simplicity's sake.

¹⁶ Distinguishing between coherent and non-coherent infinitives is far from trivial as there is a certain gradience involved: Even so-called restructuring verbs differ to what extent they allow for the monoclausal construction. For empirical details, see e.g. Grosse (2005), Schmid et al. (2005).

- (30) dat et₁ Jan ee proberen an Marie ___₁ te verkopen
 that it John has try.INF at Mary to sell.INF
 ‘that John tried to sell it to Mary’ Haegeman (1992: 110)

Similar transparency effects can be found in the third construction, cf. e.g. Bayer and Kornfilt (1994: 26). Restructuring effects in other languages also frequently obtain without (overt) complex head formation, cf. e.g. Roberts (1997: 424) on Italian. Given these facts, it is clearly preferable to separate the syntax that leads to monoclausal behavior and the operation that builds complex heads.

3.1.3 *Illustration: a simple example*

The workings of the approach can be illustrated with a simple example involving a two-verb cluster. In the syntax, such a structure will look as follows (recall that linear order does not play a role at this point):¹⁷

- (31) dass er [vp₁ wett [vp₂ es Buech läse]]
 that he wants a book read.INF (Swiss German)

In some varieties, e.g. in Swiss German, this structure can be directly linearized, leading to a VPR-structure, essentially like (31). In others like Standard German, (31) is not a possible surface form. It requires a VR-structure with the infinitive preceding the modal. VR-structures (where the verbal elements are adjacent) are derived by (syntactically) moving the XPs out of the embedded VP into the higher VP, cf. (32a). At PF, this structure is first linearized as a right-branching structure, basically as in (32a). Depending on the variety, the verbal elements may now undergo cluster formation + inversion since they are adjacent, cf. (32b):¹⁸

- (32) a. dass er [vp₁ [es Buech]₁ wett [vp₂ ___₁ läse]] movement in syntax
 that he a book wants read.INF
 b. dass er es Buech [läse+wett] inversion at PF
 that he a book read.INF+wants VR (21) (Swiss German)

In Swiss German, inversion is optional here since 2-V-clusters with V₁ = modal allow both 12 and 21 orders. In Standard German, however, only the descending variant in (32b) is grammatical;

¹⁷ For the purpose of this paper, I will label all verbal elements involved in verb clusters as lexical verbs, i.e. including modals, auxiliaries and others even though there may be reasons to classify some of them as functional elements, cf. Wurmbrand (2004c); as far as I can tell, nothing of what follows hinges on this (what is important, though, is that the types of elements can be differentiated in some way because linearization constraints can be sensitive to this difference).

¹⁸ That verb cluster formation is sensitive to adjacency has been proposed in earlier work, e.g. in van Riemsdijk (1998: 639-645) where cluster-formation involves syntactic head-movement. Given that head-movement is normally not subject to such a constraint, an adjacency condition seems stipulative. Things may be different if verb cluster formation is conceived of as reanalysis as in Haegeman and van Riemsdijk (1986) where adjacency is an intrinsic property of the mechanism. The present proposal, which locates an adjacency-sensitive operation in the post-syntactic component, is more in line with current conceptions about the architecture of grammar. Furthermore, the reanalysis approach makes a number of incorrect predictions with respect to penetrability (see section 3.4.3) and fails to account for the cluster puzzle in section 3.3.

consequently, evacuation of the lower VP is obligatory. If this fails to apply, the verbs will not be adjacent so that they cannot invert. As a consequence, the derivation crashes because it violates a linearization constraints requiring infinitives to precede modals. Note that movement from the lexical VP to the projection of the modal is in principle optional in both varieties. For instance, it does not apply in the following VP-topicalization example:

- (33) [Ein Buch lesen] will Peter nicht.
 a book read.INF wants Peter not
 ‘Peter does not want to read a book.’ (Standard German)

The varieties differ with respect to the consequences in V-final structures. In Swiss German, both movement and non-movement are a possibility (since infinitives can, but do not have to precede modals), leading to either Verb Raising (32) or Verb Projection Raising (31). In Standard German, however, only a derivation with movement leads to a converging result because otherwise the verbs cannot invert and the infinitive fails to precede the modal. In other words, the difference between languages with VPR and those without is largely due to differences in linearization requirements on verbs, i.e. VPR is blocked when descending orders are required while it is expected to occur whenever ascending orders are licensed.¹⁹ Note that the cluster formation operations as such are also optional in a technical sense. They are not triggered by the linearization constraints; rather, if they fail to apply, the resulting structure may violate a linearization constraint and will therefore crash.^{20, 21}

¹⁹ Standard Dutch, which has ascending orders but no VPR, raises interesting questions in this context. See section 3.4.4.

²⁰ Since the complements of non-restructuring predicates like e.g. *auffordern* ‘urge’ are CPs, they do not show any transparency effects (arguments cannot scramble out). Accordingly, these infinitivals usually appear in clause-final (= “extraposed”) position. Alternatively, like other CPs, they can be topicalized or scrambled (in which case they are separated from the matrix verb by a prosodic break, cf. Wurmbrand (2007: 256f.)). Predicates that are optionally restructuring like e.g. *versuchen* ‘try’, can select either VPs or CPs. When they select CPs, they behave like non-restructuring predicates. When they select VPs, the arguments of the infinitive can be scrambled out so that the verbs become adjacent and can be inverted:

- i. dass Hans versuchte [_{VP} das Buch zu lesen] → scrambling
 that John tried the book to read.INF
- ii. dass Hans [das Buch]_I versuchte [_{VP} ____I zu lesen] → inversion
 that John the book tried to read.INF
- iii. dass Hans das Buch zu=lesen+versuchte

The 3rd construction could then obtain if only part of the VP is scrambled out or if the verbs are not inverted (i.e. appear as in [ii]). However, this construction involves a number of additional complexities that space constraints prevent me from doing full justice to, see also section 3.4.1 below.

Note that on my analysis there are no true intraposed infinitival clauses (in the sense of being leftward sisters of V). Rather, they are either in scrambled position or involve scrambling + inversion as in (iii). For psycholinguistic evidence that intraposed infinitives are preferentially parsed as coherent/mono-clausal constructions (and thus not as clausal left-hand sisters of V), cf. Bayer et al. (2005: 87ff.) and Bader and Schmid (2009b). See also fn. 38 below.

²¹ The fact that PF-operations play a prominent role in this approach should not lead to the impression that these operations are also responsible for LF-phenomena such as the scope facts discussed in Haegeman and van Riemsdijk (1986). They observe that scopal elements outside the VPR-cluster can have ambiguous scope with respect to a modal that is part of the cluster; once the scopal element is inside the cluster and follows the modal, however, it can only have narrow scope. In my approach these facts obtain as follows (essentially as in den Dikken (1996)): All that counts are c-command relationships between the (copies of the) scopal elements, there is no Quantifier Raising: A QP that surfaces within the cluster consequently cannot gain scope over a c-commanding modal. However, once the QP moves to a position c-

3.1.4 *On evacuation movement*

As in other right-branching approaches to West-Germanic (starting with Zwart 1994), questions arise with respect to the movement operation that is needed to empty the VP in VR-structures (or partial VPR-structures). As discussed in Salzmann (2011), it cannot be equated with traditional scrambling: first, this movement affects not only arguments but also all other elements that might occur within a VP, i.e. at least low adverbs and predicative adjectives modifying the direct object and thus elements which cannot be reordered in the middle field. Second, if it affects arguments, none of the typical scrambling effects (freezing, prevention of focus projection) can be observed, at least as long as the arguments occur in canonical order. For the purposes of this paper, I will assume that all movement in and to the German middle field, i.e. “evacuation movement” and regular (reordering) scrambling, is driven by optional unspecific edge-features on little *v* as in Assmann and Heck (2013). The major reason for edge-feature driven movement is that no contentful trigger for scrambling could be found. It is thus straightforward to extend this to evacuation movement where a contentful trigger is even less obvious.²² The different syntactic and information-structural effects of the two reorderings (viz. evacuation movement vs. regular scrambling) can be attributed to independent factors, e.g. that only topical DPs lead to freezing effects or that only non-canonical orders disrupt focus projection. The fact that some elements, e.g. predicative adjectives, apparently can undergo only evacuation movement but not regular scrambling is not a fact about syntax but due to independent surface generalizations about the order in the middle field (e.g. as in Müller (1999)). In other words, from a syntactic perspective, word order in the German middle field is completely free. Importantly, evacuation movement is not just a stipulation needed under right-branching accounts of West-Germanic. Rather, movements of this type also need to be posited under left-branching accounts. For instance, they are found with partial VPR and the third construction when an argument occurs in a higher VP-projection as in the following examples:

- (34) a. dass er hät [em Hans]_i wele ____i es Buech schänke
 that he has the.DAT John wanted a book give.INF
 ‘that he wanted to give John a book’ (VPR, Swiss German)
- b. dass er [dem Hans]_i versuchte ____i ein Buch zu schenken
 that he the.DAT John tried a book to give.INF
 ‘that he tried to give John a book’ (3rd construction, Colloquial German)

As shown in Salzmann (2011: 460-462), Bayer and Kornfilt (1994: 45), Wöllstein-Leisten (2001), Geilfuss-Wolfgang (1991), the moved DP does not show the properties of regular scrambling (like

commanding the modal, ambiguity results as it can be interpreted both above and below the modal. A non-movement alternative deriving these facts is proposed in Salzmann (2011).

²² Hinterhölzl (2006) employs so-called licensing movement where for each type of constituent in the lower VP there is a designated landing site in the projection of the higher verb. While this derives the correct result, the solution is rather mechanical.

freezing effects).²³ Even more interesting are cases of remnant movement like the following from Fanselow (2002: 102):

- (35) Angemalt hat er sein Haus rot.
 painted has he his house red
 ‘He painted his house red.’

Here one has to assume that the predicate adjective *rot* ‘red’ moves out of VP prior to VP-topicalization even though the adjective is normally immobile in the middle field. Another interesting case is discussed in Bouma (2003) who provides verb cluster data with adverbs in the matrix VP that are semantically related to the subordinate VP. To summarize, the need for edge feature-driven movement in the middle field exists independently of the branching direction that one assumes for the VP.^{24, 25}

3.2 Applying the model to the motion verb construction

I will start with the derivation of the ascending orders in (12), repeated for convenience:

- (36) a. dass i gang [**go** de Muetter es Buech chauffe]
 b. dass i gang [de Muetter **go** es Buech chauffe]
 c. dass i gang [de Muetter es Buech **go** chauffe]
 that I go.1s PRT the.DAT mother PRT a book PRT buy.INF

The syntactic structure will look as follows (note that this is just a hierarchical structure at this point):

- (37) dass i [_{VP1} gang [_{VP2} **go** [_{VP3} de Muetter es Buech chauffe]]]
 that I go.1s PRT PRT the.DAT mother a book buy.INF

The order (36a) is basically the direct linearization of the structure in (37): neither XP-movement nor cluster-formation is necessary to derive the VPR-structure. (36b/c) are derived by moving XPs from the lowest VP to the projection of *go* (alternatively, the XPs are directly generated in the higher projection, cf. the previous subsection); then, the structures are linearized as right-branching structures.

- (38) a. dass i [_{VP1} gang [_{VP2} [de Muetter]₁ **go** [_{VP3} ___₁ es Buech chauffe]]]
 that I go.1s the.DAT mother PRT a book buy.INF

²³ As pointed out by an anonymous reviewer, movement can be avoided in the VPR case if it is derived by means of reanalysis as in Haegeman and van Riemsdijk (1986) where an element can end up in a higher VP without movement (i.e. when it is not affected by VP-reanalysis). This does not extend to the third construction, though, at least not under the predominant (partial) remnant extraposition analysis and to the remnant movement case in (35).

²⁴ A base-generation alternative with the same empirical coverage is proposed in Salzmann (2011).

²⁵ Adopting a right-branching structure thus leads to a relatively complex analysis of varieties like Standard German which have almost exclusively descending clusters. On the other hand, the analysis of varieties like Swiss German or West Flemish with preponderantly ascending orders and systematic VPR becomes very simple. This is the usual trade-off if both languages are to be derived from the same base. Since West-Germanic varieties are usually neither purely ascending or purely descending, I consider this preferable to an approach where different base orders are assumed for different varieties.

- b. *dass* i [_{VP1} *gang* [_{VP2} [*de* Muetter]₁ [*es* Buech]₂ **go** [_{VP3} ___₁ ___₂ *chauffe*]]]
 that I go.1s the.DAT mother a book PRT buy.INF

But what about the descending orders in (23), repeated from above?

- (39) a. **dass* i [**go** *de* Muetter *es* Buech *chauffe*] *gang*
 b. *?*dass* i [*de* Muetter **go** *es* Buech *chauffe*] *gang*
 c. *dass* i [*de* Muetter *es* Buech **go** *chauffe*] *gang*
 that I PRT the.DAT mother PRT a book PRT buy.INF go.1s

The starting point is again (37). In the derivation of (39c) the arguments of the lexical verb move to (or are base-generated in) the projection of the matrix motion verb:²⁶

- (40) *dass* i [_{VP1} [*de* Muetter]₁ [*es* Buech]₂ *gang* [_{VP2} **go** [_{VP3} ___₁ ___₂ *chauffe*]]]
 that I the mother a book go.1s PRT buy.INF

To derive the 231 order [*go chauffe gang*], one could be tempted to invert the matrix motion verb *gang* ‘go.1s’ with its syntactic sister VP2 [*go chauffe*] ‘PRT buy.INF’ in (40). But once it is possible to invert *gang* with its sister, one could, given the structures in (37) or (38a), also derive the ungrammatical VPR-orders in (39a/b) where *go* is not adjacent to the infinitive. Instead, as proposed above, inversion applies to adjacent verbal terminals (morphosyntactic words in the terminology of Embick and Noyer (2001)).

What we need to achieve is that the motion verb inverts with [*go chauffe*]. Given that inversion only affects terminals, *go* and *chauffe* first need to rebracket for this to become possible. Crucially, rebracketing follows from an independent property of *go* (Lötscher (1993)): it is a (category-insensitive) proclitic. Evidence for the clitic nature of *go* comes from two facts: First, it is phonetically reduced (the vowel is often reduced to schwa); second, it can be shown to be a pro-clitic because it cannot be stranded (as opposed to other members of the verb cluster like non-finite modals):²⁷

- (41) a. **[De Muetter hälffe]₁ gang* I nöd *go* ___₁ .
 the mother help.INF go.1sg I not PRT

²⁶ Importantly, (40) is also a possible surface form. Movement to VP₁ is thus not restricted to descending orders.

²⁷ As pointed out by a reviewer of a NELS-abstract, this predicts that examples like (41a) should be rescuable by inserting a DP onto which *go* can cliticize (basically as in the ascending orders in (12a/b), (36a/b)). This can be tested by means of topicalization of transitive VPs where one of the arguments is left behind. Interestingly, the example with *go* does not improve. Interestingly, but the problem seems to be more general in that the corresponding example with a modal is also quite unacceptable:

- (i) * [*es* Buech *schänke*]₁ *gang* i [*go* *de* Muetter ___₁]
 a book give.INF go.1s I PRT the.DAT mother
 (ii)?? [*es* Buech *schänke*]₁ *han* i [*wele* *de* Muetter ___₁]
 a book give.INF have.1s I wanted the.DAT mother (Swiss German)

Perhaps such examples are ruled out by information structure: VP-topicalization involves contrast, which is arguably not compatible with the VPR-structure in the middle-field, which usually entails a wide-focus interpretation, cf. Lötscher (1978).

Additionally, the Standard German version of (i), which does not include *go*, is also degraded, suggesting that there is indeed an independent factor leading to the degradedness of (i).

New arguments for verb cluster formation at PF

- b. ?[De Muetter hällffe]₁ han i nöd wele ___₁.
 the mother help.INF have.1s I not wanted

The same facts can be observed for the doubles *cho* and *la* (and *afa*). Returning to the descending orders in the *go*-construction, due to its clitic nature, *go* thus cliticizes onto the infinitive in (40), i.e. forms a unit with it, a case of string-vacuous local dislocation (= rebracketing). Then, the entire [*go+inf*]-complex can be inverted with the matrix verb:

- (42) a. dass i de Muetter es Buech gang [go=chauffe] go cliticizes onto Inf
 that I the mother a book go.1s PRT=buy.INF
 b. dass i de Muetter es Buech [[go=chauffe]+gang] inversion
 that I the mother a book PRT=buy.INF go.1s (Swiss German)

The properties of *go* are visible at this stage of the derivation because in the framework of Embick and Noyer (2001), operations like Local Dislocation occur after vocabulary insertion and are thus sensitive to the properties of individual lexical items such as being a clitic.

The derivation also shows that cyclicity is crucial: cliticization of *go* precedes inversion; the derivation can be compared with certain cases involving Latin *-que* ‘and’ as discussed in Embick and Noyer (2001: 575f.): Unlike in example (29) above where *-que* follows the first morpho-syntactic word of the second conjunct, there are also examples where it seems to attach to an entire PP, cf. (43):

- (43) in rebus=**que**
 in things.PL.ABL=and
 ‘and in things’

This would be surprising given that the placement of *-que* was argued to follow non-syntactic principles. Upon closer inspection, a different interpretation of (43) is possible: Such cases always involve light monosyllabic prepositions. They can be argued to be proclitic on the noun, i.e. they undergo string-vacuous local dislocation and form a unit with the noun. Then, *-que* is inverted with the newly formed complex (“+” marks rebracketed material, “*” indicates precedence):

- (44) -que * [in * re+bus] → -que * [in+re+bus] → [in+re+bus+que]

Independent evidence for this reanalysis comes from the fact that once an adjective is added, *-que* attaches to the adjective (and not to the noun); again, it targets the first morphosyntactic word:

- (45) in bonis=**que** rebus
 in good.PL.ABL things.PL.ABL
 ‘and in good thing’

The fact that a prosodic property of *go* is involved not only explains why we obtain an order that is marked otherwise (normally, V2 is not clitic); it also provides clear evidence for the PF-nature of

cluster-formation since the derivation of the correct cluster order presupposes a kind of cliticization that is highly unlikely to take place in syntax: *go* simply adjoins to the following constituent, it does not target a particular syntactic position. Independent evidence for this can be seen in the ascending orders (36a/b) where *go* precedes a DP. In the following example, it cliticizes onto a determiner/noun (the indefinite article is pro-clitic as well, a linking *n* is inserted for phonological reasons):²⁸

- (46) a. dass i [_{VP1} gang [_{VP2} **go** [_{VP3} es Buech chauffe]]] → rebracketing
 that I go.1s PRT a book buy.INF
- b. dass i gang [**gon+es+Buech**] chauffe
 that I go.1s PRT+a+book buy.INF (Swiss German)

One may ask at this point what rules out inversion between *gang* and [*go+es+Buech*]. There is adjacency and [*go+es+Buech*] is just one complex element, i.e. a morphosyntactic word in the relevant sense. Verb cluster-formation/Inversion is blocked because – by assumption – it only applies to verbal elements. [*go+es+Buech*], however, is not verbal: according to Embick (2007: 320f.), elements that are adjoined at PF do not become the head of the resulting complex head (basically as in syntax). Since the clitic is adjoined to the host and not vice versa, the clitic *go* does not affect the syntactic category of the resulting complex head. Given a cyclic derivation, the indefinite article will first cliticize onto the noun, the result [*es+Buech*] being a complex head of category N. Thereafter, *go* is introduced and cliticizes onto the complex head, which remains of type N: [*go+[es+Buech]*]. Therefore, inversion is blocked. [*go+chauffe*] in (42), however, certainly is verbal because the infinitive is verbal, and thus allows for inversion.

This approach predicts that whenever an element can become part of the infinitive, it can intervene between *go* and the infinitive. This prediction is borne out; the following examples illustrate this for separable verbal particles, non-referential nouns, and resultative adjectives:

- (47)a. dass i gang [go =[mit + spile]] → inversion:
 that I go.1s PRT with play.INF
 ‘that I go to play along’
- b. dass I [[go = [mit+spile] + gang]
 that I PRT with+play.INF go.1SG
- (48)a. dass i gang [go = [radio + lose]] → inversion:
 that I go.1s PRT= radio + listen.to.INF
 ‘that I go to listen to the radio.’

²⁸ One might object that the unit of *go* + V3 could also be formed by means of right-adjunction of V3 to *go* (=V2). While this is true, this implies that there is actually no procliticization of *go* (rather, movement of V3 is altruistic); however, the fact that *go* clearly does procliticize in ascending structures like (46) suggests that a generalization is missed under a head-movement analysis. See also fn. 45.

New arguments for verb cluster formation at PF

- b. dass i [[go = [Radio + lose]] + gang]
 that I PRT radio listen.to.INF go.1s

- (49)a. dass I s Schnitzel gang [go=[flach+chlopfe]] → inversion:
 that I the Schnitzel go.1SG PRT flat+hammer.INF
 ‘that I go to hammer the Schnitzel flat’

- b. dass I s Schnitzel [[go=[flach+chlopfe] + gang]
 that I the Schnitzel PRT flat+hammer.INF go.1s

In these cases, the particle, the noun and the adjective form a unit with the infinitive. *Go* then cliticizes onto this complex (i.e. undergoes string-vacuous local dislocation) and since the complex is verbal, inversion is possible.²⁹

What still needs to be addressed is what happens if the arguments do not move to the projection of the matrix motion verb (i.e. if the basis is (37) or (38a)). Inversion is blocked in both cases since we do not have the required input structure with adjacent verbs as in (40). As a consequence, only an ascending order is possible, viz. (36a/b).

Motion verbs can thus occur in either 123 or 231 order. All other orders are ungrammatical:

- (50) a.* dass i de Muetter es Buech [go+gang] chauffe 213
 that I the mother a book PRT+go.1s buy.INF
- b.* dass i de Muetter es Buech [[chauffe+go] gang] 321
 that I the mother a book buy.INF PRT go.1s
- c.* dass i de Muetter es Buech gang [chauffe +go] 132
 that I the mother a book go.1s buy.INF PRT
- d.* dass i de Muetter es Buech [chauffe [gang + go]] 312
 that I the mother a book buy.INF go.1s PRT

²⁹ I leave open at this point how the nouns, adjectives and particles in ex. (47)–(49) become part of the verb, i.e. whether they are directly generated as part of the verb (more precisely: as a head adjoined to the V°-head) or whether the complex arises derivationally. Given the thrust of the argument in this paper (and the puzzle to be introduced in the next subsection), the most obvious solution would be one in terms of PF-cluster formation. I believe that such a solution is indeed feasible (though by no means trivial), especially given that complex predicate formation with particles, nouns and adjectives is often subject to prosodic restrictions (cf. e.g. Gallmann (1999: 275)). Such effects can also be found in the examples (47)–(49). For instance, in (49), if *flach* ‘flat’ is modified by *ganz* ‘fully’, it can no longer appear together with *flach* between *go* and the infinitive, rather, both have to occur outside the *go*-complex:

i. dass I s Schnitzel {ganz flach} gang {ganz flach} go={*ganz flach} chlopfe
 that I the Schnitzel totally flat go.1SG totally flat PRT={totally flat} hammer.INF

Given space constraints, however, I will leave detailed discussion of this issue for another occasion. Note finally, that the restrictions in the Swiss German *go*-constructions in descending orders seem to be parallel to those found in West Flemish verb doubling throughout, cf. fn. 31. Furthermore, the prosodic nature of the restrictions seems parallel to what has been found for Dutch verb raising in e.g. Neeleman (1994: 230ff.); cf. Hoeksema (1994) for a slightly more careful statement on the restrictions on Dutch verb raising.

All these orders could theoretically be derived by means of cluster-formation + inversion as indicated by the brackets if *go* were a regular (i.e. non-clitic) verb (312 orders will be discussed in 3.4.1). 213 and 312 can be ruled out by cyclicity: cliticization of *go* onto V3 precedes any operations that might involve V1 so that cliticization bleeds inversion/rebracketing with V1. In the case of 132 and 321, which is based on it, inversion can probably not be prevented derivationally (since the rebracketing operation that precedes inversion respects the properties of *go*), but the resulting structure is incompatible with the needs of *go*: there is no host anymore as in these clusters nuclear stress falls on the lexical verb, cf. Wagner (2004). Potential extraposed material will not help because extraposed material is contained in a prosodic phrase separate from that of the verb cluster.

4 out of six orders involving doubling verbs are thus independently ruled out. The alternation between 123 and 231 can be subsumed under a more general property of (finite) verbs selecting bare infinitives in Swiss German: Like modals, perception verbs, inchoatives, benefactives and verbs like *leere* ‘learn’, the motion verb *gaa* ‘go’ can both precede or follow the infinitive: Since *go* forms a unit with the infinitive, it is as if we are dealing with a 2-V-cluster, cf. (32).³⁰ The possibility of VPR in 123 orders and its impossibility in 231 orders need not be stipulated, it simply follows from independent properties of *go* and the cluster-forming/inversion mechanism. This concludes my analysis of the orders in the *go*-construction.³¹

³⁰ A related idea is expressed in Barbiers and Bennis (2010: 36f.) in the context of Dutch clusters of the type *gaan zwemmen is* and in Biberauer (2013) for 231 in Afrikaans more generally.

³¹ Outside Alemannic (including Alsatian), Verb doubling is also found in Swabian, South Tyrolean and (West-)Flemish, cf. Burgmeier (2006) and Brandner and Salzmann (2009: 82ff.) for overviews. The construction seems to have the same semantics in all varieties, but there are some interesting syntactic differences:

In West-Flemish, the particle *gon*, which seems to be a reduced variant of the infinitive *goan* ‘to go’, only occurs in VR-constructions, VPR is ruled out, even in ascending orders, cf. Haegeman (1990: 82f.); (according to Karen De Clercq, who provided the following example, the construction is also marginally possible in descending, viz. 231 order)

i. dan ze atent [_{VP1} goan {√*en* boek} [_{VP2} gon [_{VP3} {**en* boek} lezen]]]
 that they always go.3PL a book PRT a book read.INF
 ‘that they always go read a book’

The only exception are zero-level elements like bare-nouns and separable particles (the restrictions thus seem to be the same as in the descending Swiss German construction, cf. (47)–(49)). Haegeman treats *gon* as a prefix that selects a non-finite VP and because of the Stray Affix filter triggers incorporation of the infinitive. Given that *gon* occurs in a more peripheral position than the *ge*-prefix (which directly attaches to the stem), it is arguably more adequate to treat it as a category-sensitive clitic. It would thus differ from Swiss German *go* only in its category-sensitivity.

In some Alemannic varieties (Vorarlberg German, Liechtenstein German, Bodensee-Alemannic, Swabian), the particle appears as *gi* (see Brandner and Salzmann (2009), Brandner and Salzmann (2012)). It differs rather strongly from *go*: It has a fixed position at the beginning of the cluster, can be followed by non-verbal material and largely blocks restructuring. Crucially, it can also be followed by non-verbal material if the motion verb occurs at the end of the cluster, thus apparently instantiating a 231 order with VPR, cf. Dobler and Rothmayr (2001: 8), Schallert (2010: 56f.):

ii. dass er [gi d Katz füttra] kunnt
 that he PRT the cat feed.INF comes
 ‘that he comes to feed the cat’

(Vorarlberg German)

These differences in behavior can be related to *gi*’s origin and synchronic categorization, cf. Brandner and Salzmann (2012): First, the particle goes back to the directional preposition *gen* ‘towards’. While the Swiss German form has developed into a verbal element, *gi* has not acquired any verbal properties but rather has developed into a prepositional complementizer (West-Flemish *gon* must have a different history, cf. Brandner and Salzmann (2012: 94) for discussion). This accounts for the absence of restructuring properties and the possibility of VPR with a cluster-final motion verb: Since we are not dealing with a verb cluster, the *gi*-phrase can be linearized to the left of the motion verb, just like other PPs, it does not depend on cluster-formation/inversion to precede a clause-final verb (note that this is expected under the

3.3 Further evidence: Post-syntactic cluster formation solves the cluster puzzle

In this section I will show that the post-syntactic cluster formation approach provides a novel and very straightforward solution to what I will call the cluster puzzle.³² As pointed out in 1.1, the verbal elements clustering at the end of the clause form an impenetrable unit when they occur in descending order. Extraposition targeting any of the non-final VPs leads to ungrammaticality. I repeat a more complex version of example (2) illustrating the same fact (the structures indicated presuppose a left-branching structure for ease of exposition only, for a right-branching structure, cf. below; for discussion on German, cf. Haider (2003: 92ff.), Bayer et al. (2005: 91ff.); on Dutch, cf. Hoeksema (1988: 152) and van Riemsdijk (1998: 640f.)):

- (51) *dass man* [_{VP1} [_{VP1} [_{VP2} [_{VP2} [_{VP3} [_{VP3} ___₁ *reden*] ***darüber**₁] können] ***darüber**₁] sollte] √**darüber**₁]
 that one talk.INF about.it
 can.INF about.it should about.it
 ‘that one should be able to talk about it’ (Standard German)

This property follows straightforwardly if the verbal elements form a complex syntactic head, either derivationally (as e.g. in Evers (1975)) or by means of base-generation as e.g. in Haider (2003), Bader and Schmid (2009a)) as extraposition cannot target segments of V. However, as pointed out in Wurmbrand (2007: 257ff.), syntactic cluster formation involving the formation/base-generation of a complex syntactic head leads to a serious problem once VP-topicalization is taken into account: Extraposition to a non-final VP is suddenly possible if it undergoes topicalization to Spec, CP:

- (52) [_{VP3} [_{VP3} ___₁ *reden*] **darüber**₁]₃ sollte₂ man schon ___₃ können ___₂.
 talk.INF about.it should one indeed can.INF

This is utterly surprising given the standard assumption that verb-second structures are derived from verb-final structures. It thus seems that (52) is derived from an ungrammatical structure. One might want to derive (52) by allowing for excorporation or by taking cluster formation to be optional, but both solutions are too simplistic: As for the first solution, while one might consider excorporation of V1 in (52) to be motivated since the verb moves to C, there remains a problem with V2 as it would have to excorporate as well to allow for topicalization of a VP that involves a verb that was part of a complex head.³³ However, there is no evidence that a verb in clause-final position would ever undergo movement; consequently, excorporation of the V₂ *können* cannot be motivated so that (52) cannot be

theory of Biberauer et al. (2008) as the FOFC does not apply if the dependent head belongs to a different syntactic category):

iii. *dass i* [_{PP} *uf Berlin*] *gang*
 that I to Berlin go.1S (Swiss German)

The cross-linguistic variation we find in the domain of verb doubling can thus be reduced to differences in the properties of the relevant lexical items.

³² For further evidence for post-syntactic cluster formation see also Wurmbrand (2004b).

³³ Given the structure in (51), this would arguably involve topicalization of VP₁ and not VP₃.

derived (needless to say, there may also be conceptual reasons to reject an excorporation approach). As for optionality, once cluster formation is simply optional, the ungrammatical versions of (51) remain unaccounted for. A more sophisticated approach could try to resort to co-occurrence restrictions: Suppose that cluster-formation results from feature-driven head-movement in syntax. The attracting features must thus be made dependent on other movement operations. Recall that cluster formation must be blocked when a) the finite verb undergoes V2 and b) there is topicalization of the lexical VP. In the case of V-to-C-movement, one could think of a co-occurrence restriction that blocks attracting features on verbs that themselves move to C. In the case of VP-topicalization the selecting verb would have to lack an attracting feature if its VP-complement is marked for undergoing topicalization. Positing the required co-occurrence restriction would, however, amount to a reformulation of the descriptive observation so that I conclude that this type of optionality also does not work. Note that this result is independent of the branching direction of the VP. In conclusion, then, it remains completely unclear how both (51) and (52) can be derived under an approach that assumes a complex head in syntax.

Interestingly, apart from Wurmbrand (2007), the issue has not been addressed in much detail. Bader and Schmid (2009a: 202, fn. 11) mention that for cases like (52), it may be necessary that V1 actually selects a VP and not a V° as they assume elsewhere. But once this possibility is granted, the cluster property in (51) cannot be derived anymore. Haegeman and van Riemsdijk (1986: 451) simply mention that VR- and VPR-clusters permit extraction and thus are not “lexical” in the sense of being impenetrable. Consequently, the cluster puzzle remains unaccounted for under their approach as well. Haider (1990: 104), Haider (2010: 307) proposes a more radical solution: He assumes that the topicalization-structure and the V-final structure are not derivationally, but only representationally related. This has drastic consequences if the VP undergoes long-distance movement or occurs with a 4-verb-cluster, as there will be at least part of a complex head clause-finally:

- (53) [VP4 [VP4 ___1 *reden*] **darüber**₁] sollte₂ man schon [können wollen ___2].
 talk.INF about.it should one indeed can.INF want.INF
 ‘One should want to be able to talk about it.’

The interpretation of the topicalized constituent as a complement of only a part of that complex head (i.e. V3 *können* ‘can’) is non-trivial (like other interpretive aspects such as adverbial modification, see Wurmbrand (2007) for critical discussion). Haider tentatively proposes that the (base-generated) topicalized XP binds an X°-trace within the cluster; the resulting mismatch is assumed to be licensed by underspecification of the bar-level of V. Unfortunately, the potentially wide-ranging consequences of this proposal are left unexplored.³⁴

³⁴ Interestingly, a very similar solution can be found in HPSG-approaches (De Kuthy and Meurers (2001: 178f.), Müller (2002)) where the filler (the topicalized [partial] VP) and the gap differ with respect to the feature [lex]: Coherent verbs in the middle field are obligatorily [+lex] and thus directly combine with the dependent verb. This leads to a cluster/complex head with the arguments of the dependent verb being inherited by the coherent verb. The verb in the topicalized constituent, however, can be [-lex] and can thus combine with a non-verbal complement. The mismatch with

An alternative base-generation approach can be found in Zwart (1993: 258-264) who proposes that topicalization should be reanalyzed as left-dislocation plus deletion of the fronted proform (which anaphorically refers to the base-generated topic). However, this reanalysis cannot be generally correct because there are systematic asymmetries between topicalization and left-dislocation, cf. Haider (1990: 100f.), Haider (2003: 95f.), and Hinterhölzl (2006: 196-198). What is relevant in the case at hand is that left-dislocation is blocked if the fronted verb (phrase) contains lexical specifications (i.e. specifications for oblique case or PP-, wh- and CP-complements) while topicalization is not (while the above-mentioned authors star such examples, it seems to me that they are not completely ungrammatical; but the contrast is clear enough to show that the two constructions need to be distinguished):

- (54) Gewartet (??das) hat er auf sie nicht.
 waited that has he not her not
 ‘He didn’t wait for her.’

One can now construct examples with extraposition in the fronted VP which show that the cluster puzzle cannot be solved by means of the left-dislocation reanalysis.³⁵

- (55) a. [[__₁ Warten] [mit dem Essen]₁]₃ (??das) hätte₂ er auf sie nicht __₃ müssen __₂.
 wait.INF with the dinner that had.SUBJ he on her not must.INF
 ‘He would not have had to wait for her with dinner.’
- b. [[__₁ Zugeben] [gegenüber dem Chef]₁]₃ (??das) hätte₂ er schon __₃ müssen __₂,
 admit.INF towards the boss that had.SUBJ he indeed must.INF

respect to [lex] is allowed because by assumption the feature [lex] is not part of the structure that is obligatorily shared by filler and gap. In other words, the structure preservation principle does not hold.

³⁵ Admittedly, there are also data that argue in favor of base-generation: Zwart (1993: 262-263) presents mismatches between the morphological shape of (elements of) the topic and the subcategorization requirements of the selecting verbs in the middle field:

- i. Marie (**te**) kussen (dat) zou ik nooit durven proberen.
 Mary to kiss.INF that would I never dare.INF try.INF
 ‘I would never dare to try to kiss Mary.’
- ii. Ik zou nooit durven proberen Marie ***(te)** kussen.
 I would never dare.INF try.INF Mary to kiss.INF
 ‘I would never dare try to kiss Mary.’
- iii. Ik zou **dat** nooit durven proberen
 I would that never dare.INF try.INF
 ‘I would never dare to try that.’

While the infinitive particle *te* ‘to’ is optional in topicalized position, it is obligatory in the middle-field if *proberen* takes a verbal complement. The variant of (i) without *te* can be related to (iii) where *proberen* takes a nominal complement, the proform *dat*. A similar case is the bleeding of the IPP effect as in *Lezen heeft hij het boek niet *willen/gewild* ‘He didn’t want to read the book’, cf. Hoeksema (1988: 159), which follows if there is an underlying ‘that’ that triggers the non-IPP form of *willen*. A derivational account of such mismatches is certainly not trivial, but a late-insertion approach to morphology that treats these effects as PF-phenomena is certainly a possibility (for instance, if IPP is sensitive to adjacency in Dutch). See also the derivational account of Hinterhölzl (2006: 198-208). Note further that German shows the inverse pattern from Dutch: IPP is preferred under topicalization but impossible under left-dislocation. For a recent analysis showing that left-dislocation should receive a treatment different from topicalization, cf. Ott (to appear). In conclusion then, while there remain problematic data for a derivational account, there is sufficient evidence against a wholesale unification of left-dislocation and topicalization. Most importantly, data like (55) that instantiate the cluster puzzle *cannot* be reanalyzed.

dass er nicht zufrieden ist.
 that he not satisfied is
 ‘He should have admitted to the boss that he is not satisfied.’

Under the PF-perspective proposed here, the crucial property that solves the puzzle is the timing: In (52), V-to-C-movement and topicalization occur in syntax and thus before PF-cluster formation can apply; i.e. the syntactic operations destroy the context for cluster-formation (there is only one verbal element in the prefield and one clause-finally); we are dealing with a bleeding relationship.³⁶ The derivation of (51) and (52) proceeds as follows: the starting point will be the following VPR-structure (again, this is just a hierarchical structure at this point):

(56) dass man [VP₁ sollte [VP₂ können [VP₃ darüber reden]]] (Standard
 that one should can.INF about.it talk.INF German)

Extrapolation can in principle target either VP₁, VP₂ or VP₃:

(57) a. dass man [VP₁ sollte [VP₂ können [VP₃[VP₃ ___₁ reden] darüber₁]]]
 that one should can.INF talk.INF about.it
 b. dass man [VP₁ sollte [VP₂ [VP₂ können [VP₃ ___₁ reden]] darüber₁]]
 that one should can.INF talk.INF about.it
 c. dass man [VP₁ [VP₁ sollte [VP₂ können [VP₃ ___₁ reden]]] darüber₁]
 that one should can.INF talk.INF about.it

All options are possible in the derivation of (51) since the verbal elements will always be adjacent so that cluster formation can apply (since inversion is necessary in Standard German, extraposition or, alternatively, movement of the PP to VP₁/base-generation in VP₁ is obligatory). (52) then simply involves V-to-C-movement of V₁ and topicalization of VP₃ in (57a) followed by cluster formation at PF.^{37, 38}

Approaches without complex head formation are hard pressed to account for the contrast between (51) and (52) since no complex head is formed so that extraposition should be possible in both cases (I will

³⁶ I assume that linearization statements (for verbal elements but also for phrasal constituents) only hold for a certain linearization domain, viz. the vP, cf. Richards (2008) (this will include all the VPs that are part of the cluster clause-finally). The fact that the finite modal precedes the non-finite modal at surface structure in (52) (which is ruled out in the verb-final structure in Standard German), therefore does not lead to a clash. The two elements belong to different linearization domains.

³⁷ Note that examples with stranded CP-complements like (55b) require CP-extraposition even under a right-branching basis. That this is independently necessary is shown in Salzmann (2013a) with respect to the placement of the infinitival particle *zu* ‘to’.

³⁸ Bayer et al. (2005: 92f.) point out that the adjacency requirement observed in (51) also holds with intraposed non-restructuring predicates. As the authors speculate themselves, given the general parsing preference for monoclausal structures (recall fn. 20), this fact arguably indicates that the preferred parse overrides the lexical specification of the non-restructuring predicate and forces a monoclausal parse. In other words, there are no intraposed non-restructuring infinitives, at least not in the verb’s complement position. Given my assumptions, this entails that the structure must involve inversion and thus adjacency at PF so that extraposition is blocked.

that the adjacency requirement on verbal elements in clause-final position results from cluster formation of verbal elements and not from prosodic restrictions on extraposition.³⁹

The present proposal with cluster formation at PF has an additional advantage: It derives the otherwise puzzling fact that verb-second movement never involves a complex head; rather, what is moved is always just the finite verb:

(60)a. *Man [reden₃ können₂ sollte₁] darüber schon.
one talk.INF can.INF should about.it indeed
'One should be able to talk about it.'

b. Man sollte₁ darüber schon reden₃ können₂.
one should about.it indeed talk.INF can.INF

Since cluster formation applies after syntax, there is no complex head yet at the point where the verb undergoes V-to-C movement. Again, movement in syntax bleeds cluster formation (involving V1) at PF. Approaches with a complex head in syntax (derived or base-generated) have to make extra assumptions to rule out movement of the complex head (such as e.g. the complexity constraint in Neeleman and Weerman (1993: 460ff.)).

3.4 Further evidence: predictions with regard to order and penetrability

In this section, I will show that the post-syntactic approach not only generates all crosslinguistically attested orders but also makes important predictions with respect to the penetrability of these orders.

3.4.1 *Post-syntactic cluster formation generates all cross-linguistically attested orders*

The cluster forming mechanism proposed in this paper is very powerful in that it can generate all six logically possible orders (for simplicity's sake, I will focus on three-verb-clusters in what follows, but the observations extend to larger clusters as well). Let me illustrate in somewhat more detail, how the various orders come about: As pointed out in 3.1.1, the default linearization leads to a 123 order. All other orders involve cluster formation: 132 is derived by means of cluster formation between V3 and V2: 1[32], 321 (as in (51)) by first inverting V3 and V2 and then V1 with [V3+V2]: [[32]1]. 213 can be derived by cluster formation between V2 and V1: [21]3. 231 and 312 are special in that they involve string vacuous cluster formation in addition to inverting cluster formation: In 231 as in the *go*-construction, V2 rebrackets with V3 and then the entire complex undergoes cluster formation with V1: [[23]1]. In 312 V1 and V2 first rebracket and then undergo cluster formation with V3: [3[12]].⁴⁰

³⁹ For reasons unclear to me, data similar to (58b) are judged somewhat better in Sternefeld (2006: 415). Such examples are certainly better than the ungrammatical variants in (51), arguably because they are at least comprehensible. But that does not make them grammatical. Haider (2003: 93), on the contrary, judges an example with a light extraposee as in (58b) ungrammatical. Note also that one easily gets google hits with strings like (59b) but not with strings like (58b).

⁴⁰ Note that rebracketing without inversion will also be needed for more complex clusters, e.g. to derive 4123 and 4312 orders (which have the structures [4[1[23]]] and [[43][12]], respectively).

As pointed out in 1.1 above, the existence of 123, 132, 321 and 312 is uncontroversial within West-Germanic. We have already seen evidence for 231 in Swiss German; the order is also found in West-Flemish (Haegeman (1998)) and in Afrikaans (Robbers (1997)) (recall also fn. 8), cf. Biberauer and Walkden (2010) and Biberauer (2013) for an overview over 231 in West-Germanic. The most controversial order certainly is 213, which is often claimed not to exist (cf. e.g. Wurmbrand (2004a), Barbiers (2005), Abels (2011: 7f.)). It is indeed rare, but one cannot exclude it on principled grounds as it is attested in certain cluster types:

According to Cooper (1995: 154) it is grammatical in Swiss German Mod-Mod-Inf clusters ('must can Verb'), and Schmid and Vogel (2004) give it as grammatical for Aux-Mod-Inf clusters ('has want Verb') in Rheidlerländer Platt (Low German), the dialect of St. Gall and for the dialect of Meran (South Tyrolean). While this clashes with the rest of the literature on Swiss German (Seiler (2004), Glaser (in preparation)) and the overview in Wurmbrand (2004a), 213 is absolutely unmarked in 3 other cluster types, namely with benefactives, perception verbs and inchoatives (Lötscher (1978: 3, 9); see also Schallert (2012) for similar data from Vorarlberg German):

- (61)a. ... wil er en ghöört₂ hät₁ chooo₃
because he him heard has come.INF
'because he heard him come'
- b. Wo s aagfänge₂ hät₁ rägne₃, simer i d Beiz
when it started has rain.INF are.1P in the pub
'When it started to rain, we went into the pub.'
- c. das är mer ghulffe₂ hät₁ abwäsche₃
that he me.DAT helped has do.the.dishes.INF
'that he helped me do the dishes'

Furthermore, in Glaser (in preparation), 213 orders were accepted by 75% of the informants for clusters based on 'has learned X-inf'. As for other varieties, 213 orders have also been noted for clusters with perception verbs in the dialect of Stellingwerf, cf. Zwart (1995). Zwart (2007: 80) reports 213 orders from Luxemburgish with 'has learned X-inf', and den Dikken (1994: 82f.) reports 213 in Middle English with a Fut-Mod-Inf cluster.

It is true that this order tends to be limited to certain cluster types (more or less contentful verbs taking a bare infinitival complement), but that also holds to a large extent for 231, at least in certain varieties (recall that in Swiss German next to the doubling construction with *go* and *la*, 231 is limited to benefactives, inchoatives and perception verbs). One may be tempted to reanalyze 213-structures as instances of the so-called 3rd construction, cf. (10) above, as both involve a V2 that is lexical and both allow non-verbal material between V1 and V3 (recall (34b)):

- (62) wo s mer ghulffe₂ händ₁ s Gschier abwäsche₃
 when they me helped have.1P the dishes wash.INF
 ‘when they helped me wash the dishes’ Lötscher (1978: 2)

The traditional account of the 3rd construction involves (remnant) extraposition of VP3 (given a left-branching basis, cf. e.g. Broekhuis et al. (1995)), and by applying this to (62), one could eliminate 213 orders from the class of attested cluster orders. However, it is far from clear in the first place whether the remnant extraposition analysis is the correct one for the 3rd construction; cf. e.g. ter Beek (2008) for an analysis based on a right-branching basis where VP3 is in its base-position while V1 and V2 invert. This may imply that the class of 213 orders is even larger. The literature contains some indications that VPR and the 3rd construction should be differentiated, but most of the differences are very subtle (cf. Salzmann (2013b) for an overview). In Salzmann (2013a) I have presented a new diagnostic that robustly distinguishes between VPR and the 3rd construction, viz. the placement of the infinitival particle *zu* ‘to’: While *zu* appears misplaced in VPR, i.e. does not occur on the verb governed by a *zu*-selecting constituent but rather on the last verb of the verb cluster, there is no such displacement in the 3rd construction (the preposition/complementizer *ohne* ‘without’ selects a *zu*-infinitive; in VPR, *z(u)* does not occur on V1 *welle* as one would expect but on V2 *stelle*; in the third construction, *zu* regularly occurs on V1 and since V1 also selects a *zu*-infinitive, *zu* also occurs on V2):

- (63)a. ohne mich [_{VP1} *(zu) versuchen₁ [_{VP2} t_{mich} zu mögen₂]] 3rd construction
 without me to try.INF to like.INF (Colloquial German)
 ‘without trying to like me’
- b. ohni mi [_{VP1} (*z) welle₁ [_{VP2} t_{mi} uf d bullesite z stelle₂]] VPR
 without me to want.INF on the cops.side to put.INF ...
 ‘without wanting to side with the cops ...’ (Swiss German)
<http://www.fcbforum.ch/forum/showthread.php?4328-usschritige-nachem-spiel-!/page4;>
 found on 11.3.2013

I conclude in Salzmann (2013a) that the extraposition analysis of the 3rd construction is correct after all (even under a right-branching base). While this limits the class of 213 orders, what is crucial is that 213 orders like (62) behave differently with respect to this diagnostic: The VPR-pattern with *zu*-placement obtains, not that of the 3rd construction (the data are hard to judge, but the preference is quite clear; an alternative with 2 *zus* is also accepted by some speakers, suggesting that both options may be available in the Swiss German grammar).

- (64). ohni en ghört₂ (*z) ha₁ *(z) singe₃
 without him heard to have.INF to sing.INF
 ‘without having heard him sing’

This implies that 213 orders do exist and thus supports cluster theories like the one proposed here that generate all 6 logically possible orders.⁴¹

3.4.2 *Predictions with respect to penetrability*

The fact that a single mechanism is sufficient to generate all relevant orders certainly is a strong argument in favor of it. I will now show that it also makes important predictions with respect to the penetrability of the various orders. Even though I will use (Swiss) German in what follows to illustrate these properties, they are not peculiar to (Swiss) German but hold more generally of West-Germanic. Exceptions like Standard Dutch and West-Flemish will be addressed in 3.4.4

VPR with 123 is uncontroversial. Non-verbal material can appear both between V1 and V2 and between V2 and V3 (recall ex. (13)–(14)). I will henceforth indicate non-verbal material by means of X. 123 clusters can thus be represented by 1X2X3. This is entirely expected under the present approach since no cluster formation is involved so that nothing precludes leaving arguments of V3 in situ or scrambling them to/base-generating them in a projection of V2. 132 allows non-verbal material between V1 and V3, the order can thus be represented as 1X32. Examples can be found in standard German (they are relatively rare, probably for prescriptive reasons), cf. Bader and Schmid (2009a: 224):

- (65) Es ist schön, dass Max neulich hat₁ zu dem Nationalpark fliegen₃ dürfen₂.
It is nice that Max lately has to the national.park fly.INF may.INF
'It is nice that Max has been allowed to fly to the national park lately.' (Standard German)

V3 and V2, however, form an impenetrable unit. Extraposition targeting a position between V3 and V2 leads to ungrammaticality; instead, as in the examples in section 3.4.1, extraposition is only felicitous at the right edge of the cluster:

- (66) a. *dass er hat₁ reden₃ darüber wollen₂ b. dass er hat₁ reden₃ wollen₂ darüber.
that he has talk.INF it.about want.INF that he has talk.INF want.INF it.about
'that he wanted to talk about it' (Standard German)

The properties again follow under the post-syntactic approach: There is no cluster formation between V1 and V3 so that nothing prevents non-verbal material there, e.g. adverbials or arguments of V3. The impenetrability of V3 and V2 follows, however, since cluster formation is involved. Under a right-branching basis, extraposition will always target a right-peripheral position, whether it involves adjunction to VP2 or VP3 (as 3.3 above).

321 orders are always completely impenetrable as shown in section 3.3. This is an automatic result of the present approach since all verbal elements are involved in cluster formation.

⁴¹ Barbiers and Bennis (2010) try to reanalyze some of the orders; they argue that there are essentially only two cluster types, viz. 123 and 321. The other orders obtain if V3 is non-verbal, i.e. adjectival or nominal. 132 thus constitutes a two-verb cluster with VPR and 312 is a two-verb cluster with a nominal object in regular object position (231 involves a slightly more complex account). In Salzmann (2012) I show that this proposal cannot be extended to German varieties.

New arguments for verb cluster formation at PF

As pointed out in 3.4.1 above, 213 allows for non-verbal material between V1 and V3, viz. 21X3. This is expected because only V1 and V2 are involved in cluster formation while VP3 remains in its base position. Importantly, though, V1 and V2 are impenetrable, extraposition to VP2 is impossible, the extraposee has to occur at the right edge of the cluster:

- (67) a. *dass er eme Maa ghulffe₂, won er nett find, hät₁, sini wonig striiche₃
that he a.DAT man helped C he nice finds has his apartment paint.INF
b. dass er eme Maa ghulffe₂ hät₁ sini wonig striiche₃, won er nett findt
that he a.DAT man helped has his apartment paint.INF C he nice finds
'that he helped a man paint his apartment that he likes' (Swiss German)

It has already been shown in 2.2 above that 231 disallows VPR in Swiss German (on 23X1 in West-Flemish, see 3.4.4 below). Again, this follows from the cluster forming mechanism as V2 and V3 first have to form a unit before they can undergo cluster formation with V1. The approach additionally correctly predicts the impossibility of extraposition to VP3. Again, extraposition has to target the right edge of the cluster:

- (68) a. *dass er go₂ spile₃ demit gaat₁ b. dass er go₂ spile₃ gaat₁ demit
that he PRT play.INF it.with goes that he PRT play.INF goes it.with
'that he goes to play with it' (Swiss German)

Turning finally to 312, we can observe that the order is completely impenetrable. First, extraposition to a position following V3 is impossible.

- (69)a. *dass er rede₃ *drüber hät₁ wele₂ b. dass er rede₃ hät₁ wele₂ √drüber
that he talk.INF it.about has wanted that he talk.INF has wanted it.about
'that he wanted to talk about it' (Swiss German)

Furthermore, it also disallows VPR. For instance, adverbials modifying VP2 cannot occur between V1 and V2 (something that is possible under a 123 order):

- (70) *dass geschter öppert verbichoo₃ het₁ unbedingt söle₂
that yesterday someone drop.by.INF had.SUBJ definitely should.INF
'that yesterday someone definitely should have dropped by' (Swiss German)

We can thus conclude that the post-syntactic approach makes the correct predictions both with respect to the attested orders cross-linguistically as well as the penetrability of the various orders.

3.4.3 Previous approaches and their predictions with respect to order and penetrability

There are other approaches to verb clusters that also generate all six logically possible orders, viz. Haegeman and van Riemsdijk (1986), Koopman and Szabolcsi (2000), Haider (2003), and Bader and Schmid (2009a). In this subsection, I will show that they make incorrect predictions with respect to

penetrability, i.e. they predict penetrability for certain orders which, however, are always compact cross-linguistically. The post-syntactic approach proposed here instead correctly limits penetrability to those orders where one finds penetrability within West-Germanic (with one exception, cf. 3.4.4 below). I will start with the base-generation approach by Bader and Schmid (2009a) (which is based in part on Williams (2004)). I will first illustrate its major shortcoming on the basis of the facts from the motion verb construction before showing that it generally makes incorrect predictions w.r.t. penetrability. I repeat the crucial contrast from above: Adjacency between *go* and the infinitive is required only in the 231 order while VPR is possible under 123:

- (71) a. *dass i [**go**₂ de Muetter es Buech chauffe₃] gang₁
 b. *?dass i [de Muetter **go**₂ es Buech chauffe₃] gang₁
 c. dass i [de Muetter es Buech **go**₂ chauffe₃] gang₁
 that I PRT the.DAT mother PRT a book PRT buy.INF go.1s
- (72) a. dass i gang₁ [**go**₂ de Muetter es Buech chauffe₃]
 b. dass i gang₁ [de Muetter **go**₂ es Buech chauffe₃]
 c. dass i gang₁ [de Muetter es Buech **go**₂ chauffe₃]
 that I go.1s PRT the.DAT mother PRT a book PRT buy.INF

In this approach, different cluster orders and VR- and VPR-structures are directly generated as such in the syntax, they are neither transformationally related to each other nor to some putative base-structure. Their system is based on an elaborate subcategorization system that specifies not only the category of the complement of some head, but also directionality, i.e. whether it is selected to the left or to the right, and the projection level of that complement, i.e. whether it is an X^o or an XP. Concerning projection level, the possibility of X^o-complements is needed to base-generate complex heads as in VR-structures. XP-complements, on the other hand, are needed for VPR-structures. Under this approach, the asymmetry in the motion verb construction can be derived relatively easily by stipulating a disjunction in the lexical entry of the motion verb: When it selects its *go*-complement to the right, *go* can be either X or XP (73a), but when it selects *go* to the left, *go* has to be a (possibly complex) head (73b). The entry for *go* is very simple. It can take either an X^o- or an XP-complement to its right (73c):

- (73) a. $V_{\text{motion}} \rightarrow V_{\text{go}}/VP_{\text{go}}$ c. $V_{\text{go}} \rightarrow V(P)_{\text{inf}}$
 b. $V_{\text{go}} \leftarrow V_{\text{motion}}$

The possibility of VPR in 123 orders like (72) follows from the fact that both the complement of the motion verb and of *go* can be XP (due to functional composition, arguments of the lexical verb can be merged anywhere in the projection of the complex head). The impossibility of VPR in 231 orders as in (71) simply follows from the fact that when the motion verb takes its complement to the left, it has to be X^o. This has the automatic consequence that *go* itself also can only take an X^o-complement. Admittedly, this approach can derive the asymmetry, but this is only possible at a very high cost, namely by essentially reformulating the descriptive generalization. Given this formalism, it would be

just as easy to describe a language which has the opposite pattern, i.e. VPR only in 231 orders but only VR in 123 orders. Given that the 123 order is the one most systematically associated with VPR cross-linguistically (while VPR is usually impossible in 231 orders), this points towards a systematic weakness of the formalism.⁴² Put more generally, the approach does not make any predictions in which orders penetrability is possible, in fact it should be possible in all orders: Since a V (be it V1 or V2) can in principle always select a VP and nothing precludes selecting a VP to its left, the approach wrongly predicts orders that are at least partially descending to be penetrable, contrary to fact (recall 3.4.2.). For instance, it wrongly predicts the following structures (with X marking a possible extraposition site): 3X2X1, 3X12, 2X13, 23X1, 13X2. Furthermore, it wrongly predicts the possibility of 31X2 (e.g. with an adverbial, if V1 takes a VP-complement to the right) as in (70) above.

Haider (2003) proposes a hybrid account where (descending) VR-structures are base-generated as complex heads while (ascending) VPR structures involve excorporation of a verb of the complex head and reprojecting movement (the excorporated verb projects an additional VP-shell). The starting point is always a 321 order in form of a base-generated complex head. A 132 order is then derived as follows:

$$(74) \text{ a. } [_{VP} [_{V^{\circ}} [_{V^{\circ}} [_{V^{\circ}} V3] V2] V1]] \quad \rightarrow \quad \text{ b. } [_{VP} V1_1 [_{VP} [_{V^{\circ}} [_{V^{\circ}} [_{V^{\circ}} V3] V2] __1]]]]$$

To derive the 123 order, both V2 and V1 have to reproject. This correctly derives the possibility of VPR in ascending structures like (72):

$$(75) \text{ a. } [_{VP} [_{V^{\circ}} [_{V^{\circ}} [_{V^{\circ}} V3] V2] V1]] \quad \rightarrow \quad \text{ b. } [_{VP} V1_2 [_{VP} V2_1 [_{VP} [_{V^{\circ}} [_{V^{\circ}} [_{V^{\circ}} V3] __1] __2]]]]]$$

Importantly, 231 orders are an intermediate step of a 123 derivation, i.e. they only involve reprojecting of V2:

$$(76) \text{ a. } [_{VP} [_{V^{\circ}} [_{V^{\circ}} [_{V^{\circ}} V3] V2] V1]] \quad \rightarrow \quad \text{ b. } [_{VP} V2_1 [_{VP} [_{V^{\circ}} [_{V^{\circ}} [_{V^{\circ}} V3] __1] V1]]]]$$

Since projection/shell-formation is correlated with VPR, we expect VPR to be possible with 231 orders, contrary to what was found in (71). It seems that non-verbal material intervening between V2 and V3 has to be scrambled out if V1 fails to undergo reprojecting. But this is a case of look-ahead and it is unclear how scrambling could be enforced in the right configuration. The other 2 orders can also be generated with reprojecting movement: 213 involves excorporation of V1 followed by excorporation of V2, 312 involves excorporation of V1 followed by excorporation of V3. As for general predictions w.r.t. penetrability, the approach correctly rules out incorrect extrapositions in the following structures: In *3X2X1, *23X1, and *13X2 the relevant verbs are still part of the cluster. In *21X3 and *3X12 projecting movement does not generate any new adjunction sites to the left, extraposition automatically targets the right edge of the cluster. However, there are problems with VPR if adverbials are taken into account. Through shell-formation, adverbials should be licensed in the following structures *2X13 and *3X1X2 because they are possible under X1X23 which also involves reprojecting. To summarize, then,

⁴² A similar problem obtains for Abels (2011), and Abels (2013) where 123, 132, 231 and 321 result from flexible base-generation/linearization of VPs. It is not obvious why the orders should differ with respect to VPR.

Haider's approach cannot derive the asymmetry in the motion verb construction and incorrectly predicts penetrability in two further configurations.⁴³

I now turn to XP-movement approaches. I will illustrate the inherent difficulties of this type of approach w.r.t. penetrability on the basis of Barbiers (2005), but they also extend to Barbiers (2008), which is a variant of the 2005 article based on flexible copy realization, and to radical remnant movement approaches like Koopman and Szabolcsi (2000) and Hinterhölzl (2006). Barbiers (2005) adopts an ascending, right-branching structure in the base with verbs taking VP-complements. As a consequence, a 123 order with VPR is the default. Other cluster orders are derived by means of VP-movement. 231 orders are derived by moving VP2 (pied-piping VP3) into the specifier of V1:

(77) [_{VP1} [_{VP2} V2 [_{VP3} V3]] V1 _{VP2}]

In this type of approach, VPR-structures are the default, VR-structures obtain if the complements of the lexical verb move to a position above VP1 (e.g. some agreement projection as in Zwart (1994)). Given that what is moved is a full VP, we expect VPR to be possible in 231 orders, contrary to fact (at least in Swiss German). Evacuation of VP3 is thus necessary if it is pied-piped by V2, but not when it is left in-situ. There is no simple way of capturing the VPR-asymmetry in this system. On a more general level, the approach makes incorrect predictions w.r.t. penetrability in the following cases: Since only VPs move and no complex head is formed, there should be additional adjunction sites. The approach thus incorrectly predicts the following structures (with X = extraposee) to be possible: *3X12 (VP3 moves to Spec, V1), *23X1 (as in (77)), *13X2 (VP3 moves to SpecV2), *3X2X1. Furthermore, the approach fails to rule out VPR-cases with adverbials in *3X1X2 (as in (70)) given that adverbials can occur in these positions in the 123 order. Koopman and Szabolcsi (2000: 193f.) resort to specific filters to enforce impenetrability of certain clusters. Given that the possibilities of penetrability converge to a very large extent cross-linguistically, an approach like the one proposed here that directly derives these possibilities is certainly preferable.

The approach proposed here bears a certain similarity to the proposal in Haegeman and van Riemsdijk (1986) where complex heads, transparency effects and the various orders are derived by means of reanalysis, i.e. rebracketing in syntax, and reordering of sister nodes at PF. VR structures obtain if reanalysis involves heads, VPR obtains if it involves projections of V (V' and VP). Suppose we want to derive the asymmetry in (71)-(72) given the following left-branching basis (for simplicity's sake, I do not indicate the structure beyond VP):

⁴³ Haider (2003: 118) entertains another option for certain German verb clusters, viz. cliticization (which is basically adjunction of a head to a higher head, either to the left or to the right). 231 orders could then be derived by adjoining V3 to the right of V2 ([[V3+V2] V1] → [[₁ V2+V3₁] V1]). This would derive a complex head, as is necessary for (71). The same operation would also derive a fully impenetrable 312 order (with V2 adjoining to V1). The problem that remains, though, is to force cliticization and simultaneously prevent reprojection for exactly these orders. One could, of course, stipulate that certain verbs cannot undergo reprojection; blocking lexical verbs from reprojection would rule out *3X1X2. Ruling out reprojection of inchoatives, benefactives, and perception verbs as V2 would avoid the problems with *2X31 and *2X13. However, this would have to be restricted to cases where it is not followed by reprojection of V1 since 123 is, of course, penetrable. This cannot be done in a straightforward way and shows again that the approach does not correctly capture the cross-linguistic generalizations about penetrability.

3.4.4 *Overgeneration and undergeneration*

I have opted for a mechanism that is powerful enough to generate all orders that are attested in West-Germanic. This will invariably overgenerate for individual dialects, varieties and idiolects since they generally only allow for a (often different) subset of the six orders. I nevertheless think that this approach is on the right track because verb clusters have very similar properties cross-linguistically. For instance, they always show transparency effects, the order variation is never associated with semantic differences and VPR is only possible in ascending clusters. In a sense then, verb clusters are the same phenomenon cross-linguistically so that they should be accounted for with the same mechanism in all varieties unless there is clear evidence to the contrary (see Barbiers (2005) for a similar view).⁴⁶

The present account thus overgenerates with respect to the orders acceptable in individual varieties and the question arises what the grammar of such varieties should look like. I will largely follow Barbiers (2005) who proposes that the grammar of every speaker has access to all the orders that can be generated by the grammar and that the orders that the speakers actually use and consider acceptable are determined by extragrammatical factors such as geography, frequency, register, social class, age etc. The parsing preferences ('biases') discussed in Culicover (2013) can be added to this list. Such a perspective has the immediate advantage that it avoids the pervasive gradience issue observed in most empirical work on verb clusters (cf. e.g. Seiler (2004), Barbiers (2005), Bader and Schmid (2009a)). Speakers usually allow several orders for a given cluster type, often to variable degrees. In many cases the boundary between grammatical and ungrammatical can only be drawn on an arbitrary basis and it is often that case that all orders receive at least some degree of acceptability. For instance, Glaser (in preparation) found each of the six possible orders in Swiss German Aux-Mod-Inf clusters to be accepted by at least 2.5% of the informants with 231 orders receiving the lowest ratings, 123 orders receiving almost 100% acceptance and 312 orders being of intermediate acceptability (35%). Treating such cases as involving different degrees of markedness that result from the interaction of the various soft factors mentioned above certainly does better justice to the facts than classifying some orders as grammatical and some as ungrammatical.⁴⁷ Given this view, the linearization constraints mentioned in

Compared to cluster formation at PF, such an approach has a number of disadvantages: First, it requires optional movement and variable adjunction sites for the same movement operation. Second, like all approaches involving the formation of a complex head in syntax, it is confronted with the cluster puzzle discussed in 3.3. Third, since reordering is not subject to adjacency, it could wrongly derive 132 orders where arguments of V3 are stranded as in **dass er hat lesen wollen ein Buch* 'that he has read wanted a book'.

⁴⁶ This argues against approaches that combine several mechanisms, e.g. VP-inversion at PF + VP-movement as in Wurmbrand (2004b) or flexible base-generation + VP-movement as in Abels (2011), Abels (2013). It is sometimes claimed in this context that 312 orders involve focus on V3 (and may therefore justify a different derivation). However, this claim has been falsified in Bader and Schmid (2009a). Quite apart from that this would be an implausible state of affairs because 312 represents the default order in Bavarian, many Austrian varieties and in Eastern Switzerland.

⁴⁷ To give an example: The fact that rebracketing without inversion follows from an independent property of *go* certainly contributes to the unmarkedness of the 231 order in this case. No such mitigating factor is found if V2 is non-clitic. Treating one case as grammatical and the other one as ungrammatical would fail to capture this insight. Note also in this context that according to Abels (2013), the markedness of 231 results from a mismatch between the prosodic structure and the syntactic structure that is not found with other cluster types. The fact that the different prosody of the *go*-construction makes 231 orders more acceptable can be seen as indirect support for this proposal.

3.1.1 above are best understood as markedness constraints that do not determine grammaticality but preference.⁴⁸

While the approach proposed here overgenerates w.r.t. order in individual varieties, the possibilities it generates for penetrability are closely matched by what can be observed cross-linguistically. There are two exceptions, viz. Standard Dutch and West Flemish. Standard Dutch (including many of its dialects, cf. the map in Barbiers et al. (2008b: map 2.3.1.7)) has ascending clusters but disallows VPR (the present account thus overgenerates), and West-Flemish has VPR in 231 orders (the account undergenerates):

- (79)a. dat Jan het meisje {√een kus} heeft proberen {*een kus} te geven.
 that John the girl a kiss has tried.IPP a kiss to give.INF
 that John tried to give the girl a kiss' (Standard Dutch, Broekhuis et al. (1995: 99)
- b. dat Jan het meisje {√een kus} geprobeerd heeft {√een kus} te geven.
 that John the girl a kiss tried has a kiss to give.INF
- (80) da Valère willen₂ **dienen boek** lezen₃ eet₁.
 that V. want.INF this book read.INF has
 'that Valère has wanted to read that book' West Flemish, cf. Haegeman (1998: 634)

The minimal pair in (79) shows that with the IPP effect (79a), the verbal elements are impenetrable (with the exception of incorporated X^o-elements like particles and non-referential nouns), while with a participle (79b) no adjacency is required. The data are usually interpreted as showing that (79a) involves a complex head while (79b) involves extraposition of VP3 (and is thus an instance of the third construction). In the present account, the Dutch facts can be derived if it is assumed that Dutch obligatorily makes use of rebracketing without inversion, i.e. involves the structure [1[23]]. Enforcing this derivation and excluding the structure 123 that does not involve any rebracketing (as in Swiss German) can be done by means of a surface constraint. Instead of such rather mechanical solutions I will follow Hoeksema (1994), who argues that VPR is residually possible also in Standard Dutch. While most of the interveners that one finds can be interpreted as incorporated X^o-elements, there are also cases that resist such an analysis. Hoeksema therefore regards the near absence of VPR as a matter of usage rather than grammar and attributes the near impossibility of VPR to standardization (which makes sense given that VPR was possible in earlier stages of the language and still allows similar orders in the cluster). This solution has the advantage that Dutch can make use of the same possibilities as the other varieties. As with cluster orders, the restrictions are due to extra-grammatical factors.

The West-Flemish case in (80) is more problematic as it cannot be generated by the cluster-formation operation proposed here. To my knowledge, it is the only West-Germanic variety that allows for VPR

⁴⁸ The situation with verb clusters can be compared with that of the order in the German middle field where similar questions arise concerning the boundary between grammaticality and markedness. An constraint-based solution as e.g. in Müller (1999) would be an interesting way of modeling this. For reasons of space, I will leave the development of such an analysis for verb clusters for another occasion.

in 231 orders. Given the assumptions made here, the only possible way of accommodating these facts is to assume that West Flemish can make use of an additional mechanism, i.e. inversion between V1 and VP2 or VP-movement or VP2 to Spec, V1. I am not aware of any decisive evidence in favor of either option; the fact that 231 only occurs if V1 is a perfective auxiliary and bears certain tense/mood features (cf. Haegeman and Oosterhof (2012)) can be implemented under either version. I will nevertheless opt for VP-inversion because this is more in line with the approach to verb clusters pursued here and because VP-movement predicts movement properties (such as freezing) that are never observed in these constructions. West Flemish would thus remain like the other varieties, the only difference being that it has an additional operation in a specific context.

4 Conclusion and Outlook

I have argued in favor of a post-syntactic treatment of verb cluster formation based on a penetrability asymmetry that arises between ascending and descending orders in Swiss German verb doubling. The ordering restrictions in the descending order require crucial reference to linear order and order-related prosodic properties. The analysis proposed here thus differs from previous post-syntactic approaches in that verbal terminals are inverted and not syntactic sister nodes. This change in perspective has the additional advantage that it provides a new and simple solution to the otherwise quite intractable cluster puzzle found with extraposition and VP-topicalization. Furthermore, I have shown that this mechanism not only generates all of the cross-linguistically attested six orders but also (with one exception) correctly predicts which orders are penetrable in which positions. Taken together this provides new arguments for a derivational approach to verb cluster formation and order variation. Finally, while most verb cluster phenomena are amenable to both a left-branching or a right-branching treatment, the order-dependent penetrability asymmetries in the motion verb construction and beyond constitute clear evidence in favor of a right-branching VP.

5 References

- Abels, Klaus. 2011. Hierarchy-order relations in the Germanic verb cluster and in the noun phrase. *Groninger Arbeiten zur Germanistischen Linguistik* 53:1-28.
- Abels, Klaus, and Neeleman, Ad. 2012. Linear Asymmetries and the LCA. *Syntax* 15:25-74.
- Abels, Klaus. 2013. On 2-3-1: Ms. UCL.
- Assmann, Anke, and Heck, Fabian. 2013. Opaque Intervention. In *Rule Interaction in Grammar (Linguistische Arbeitsberichte 90)*, eds. Fabian Heck and Anke Assmann, 463-509. Leipzig: University of Leipzig.
- Bader, Markus, and Schmid, Tanja. 2009a. Verb clusters in colloquial German. *The Journal of Comparative Germanic Linguistics* 12:175-228.
- Bader, Markus, and Schmid, Tanja. 2009b. Minimality in verb-cluster formation. *Lingua* 119:1458-1481.
- Barbiers, Sjef. 2000. The Right Periphery in SOV Languages. English and Dutch. In *The Derivation of VO and OV*, ed. Peter Svenonius, 181-218. Amsterdam: John Benjamins.
- Barbiers, Sjef. 2005. Word order variation in three verb clusters and the division of labor between generative linguistics and sociolinguistics. In *Syntax and Variation. Reconciling the Biological and the Social*, eds. Leonie Cornips and Karen P. Corrigan, 233-264. Amsterdam: John Benjamins.
- Barbiers, Sjef. 2008. Werkwoordclusters en de grammatica van de rechterperiferie. *Nederlandse Taalkunde* 13:160-188.

- Barbiers, Sjef, Koeneman, Olaf, Lekakou, Marika, and Van der Ham, Margreet eds. 2008a. *Microvariation in syntactic doubling. Syntax and semantics*. Bingley: Emerald.
- Barbiers, Sjef, van der Auwera, Johan, Bennis, Hans, Boef, Eefje, de Vogelaer, Gunther, and Ham, Margreet van der. 2008b. *Syntactic Atlas of the Dutch Dialects, volume II*. Amsterdam: Amsterdam University Press.
- Barbiers, Sjef, and Bennis, Hans. 2010. De plaats van het werkwoord in zuid en noord. In *Voor Magda. Artikelen voor Magda Devos bij haar afscheid van de Universiteit Gent*, eds. Johan De Caluwe and Jacques Van Keymeulen, 25-42. Gent: Academia Press.
- Bayer, Josef, and Kornfilt, Jaklin. 1994. Against Scrambling as an instance of Move-alpha. In *Studies on Scrambling. Movement and Non-Movement Approaches to Free Word-Order Phenomena*, eds. Norbert Corver and Henk van Riemsdijk, 17-60. Berlin: Mouton.
- Bayer, Josef, Schmid, Tanja, and Bader, Markus. 2005. Clause union and clausal position. In *The function of function words and functional categories*, eds. Marcel den Dikken and Christina M. Tortora, 79-113. Amsterdam: John Benjamins.
- Biberauer, Theresa, Holmberg, Anders, and Roberts, Ian. 2008. Structure and Linearization in Disharmonic Word Orders. *Proceedings of WCCFL* 26:96-104.
- Biberauer, Theresa, and Walkden, George. 2010. 231 From A(frikaans) to Z(ürich German): A Challenge to the Final-over-final Constraint. Ms. University of Cambridge.
- Biberauer, Theresa. 2013. Germanic verb clusters again - with Afrikaans center stage. Paper presented at the Comparative Germanic Syntax Workshop 28, October 3-5, Leipzig.
- Bouma, Gosse. 2003. Verb clusters and the scope of adjuncts in Dutch. In *Verb Constructions in German and Dutch*, eds. Pieter A. M. Seuren and Gerard Kempen, 5-42. Amsterdam: John Benjamins.
- Brandner, Ellen, and Salzmann, Martin. 2009. Crossing the lake: Motion verb constructions in Bodensee-Alemannic and Swiss German. *Groninger Arbeiten zur germanistischen Linguistik* 48:81-113.
- Brandner, Ellen, and Salzmann, Martin. 2012. Crossing the Lake: Motion verb constructions in Bodensee-Alemannic and Swiss German. In *Comparative Germanic Syntax: The State of the Art*, eds. Peter Ackema et al., 67-98. Amsterdam: John Benjamins.
- Broekhuis, Hans, Besten, Hans den, Hoekstra, Kees, and Rutten, Jean. 1995. Infinitival complementation in Dutch: On remnant extraposition. *The Linguistic Review* 12.
- Burgmeier, Markus. 2006. I gang go schaffa - Zur Vorkommensweise der Infinitivpartikel ‚go‘ in alemannischen Dialekten, MA thesis, University of Zurich.
- Cooper, Kathrin E. 1995. *Topics in Zurich German syntax*. Groningen: Germanistisch Instituut.
- Culicover, Peter. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.
- De Kuthy, Kordula, and Meurers, Walt. 2001. On Partial Constituent Fronting in German. *The Journal of Comparative Germanic Linguistics* 3:143-205.
- den Besten, Hans, and Edmonson, Jerold, A. 1983. The Verbal Complex in Continental West Germanic. In *On the Formal Syntax of the Westgermania*, ed. Werner Abraham, 155-216. Amsterdam: John Benjamins.
- den Dikken, Marcel. 1994. Minimalist Verb (Projection) Raising. *Groninger Arbeiten zur Germanistischen Linguistik* 37:71-88.
- den Dikken, Marcel. 1995. Verb (Projection) Raising, Scope, and Uniform Phrase Structure. *Proceedings of NELS* 25:95-110.
- den Dikken, Marcel. 1996. The Minimal Links of Verb (Projection) Raising. In *Minimal Ideas. Syntactic Studies in the Minimalist Framework*, eds. Werner Abraham, Samuel David Epstein, Hoskuldur Thrainsson and C. Jan-Wouter Zwart, 67-96. Amsterdam: Benjamins.
- Dobler, Eva, and Rothmayr, Antonia. 2001. The gi+inf construction in Vorarlberg German. *Wiener Linguistische Gazette* 67-69:1-24.
- Embick, David, and Noyer, Rolf. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32:555-595.
- Embick, David. 2007. Linearization and Local Dislocation: Derivational Mechanics and Interactions. *Linguistic Analysis* 33:303-336.
- Evers, Arnold. 1975. The transformational cycle of Dutch and German. Doctoral Dissertation, University of Utrecht.

- Fanselow, Gisbert. 2002. Against Remnant VP-Movement. In *Dimensions of Movement. From Features to Remnants*, eds. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers and Hans-Martin Gärtner, 91-125. Amsterdam: John Benjamins.
- Gallmann, Peter. 1999. Wortbegriff und Nomen-Verb-Verbindungen. *Zeitschrift für Sprachwissenschaft* 18:269-304.
- Geilfuss-Wolfgang, Jochen. 1991. Scrambling und Pseudoscambling. In *Verb- und Verbphrasensyntax. Arbeitspapier 11 des Sonderforschungsbereichs 340 Sprachtheoretische Grundlagen für die Computerlinguistik*, ed. Jochen Geilfuss-Wolfgang, 19-57. Tübingen: University of Tübingen.
- Glaser, Elvira. in preparation. Dialektsyntax des Schweizerdeutschen. Syntaktischer Atlas der Deutschen Schweiz (SADS). Zurich: University of Zurich.
- Grosse, Julia. 2005. *Zu Kohärenz und Kontrolle in infiniten Konstruktionen des Deutschen*. Marburg: Tectum.
- Haegeman, Liliane, and van Riemsdijk, Henk. 1986. Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs. *Linguistic Inquiry* 17:417-466.
- Haegeman, Liliane. 1990. The syntax of motional goan in West Flemish. In *Linguistics in the Netherlands 1990*, eds. Reineke Bok-Benema and Peter Coopmans, 81-90. Amsterdam: John Benjamins.
- Haegeman, Liliane. 1992. *Theory and Description in Generative Syntax. A case study in West Flemish*. Cambridge u.a.: Cambridge University Press.
- Haegeman, Liliane. 1998. Verb Movement in Embedded Clauses in West Flemish. *Linguistic Inquiry* 29:631-656.
- Haegeman, Liliane, and Oosterhof, Albert. 2012. Effecten van finietheid, tempus en irrealis op de werkwoordsvolgorde in Vlaamse IPP-constructies. *TABU: Taalkundig bulletin van het Nederlands instituut van de Rijksuniversiteit Groningen*:23-46.
- Haider, Hubert. 1990. Topicalization and Other Puzzles of German Syntax. In *Scrambling and Barriers*, eds. Gunther Grewendorf and Wolfgang Sternefeld, 93-112. Amsterdam: Benjamins.
- Haider, Hubert. 2003. V-clustering and clause union: Causes and effects. In *Verb constructions in German and Dutch*, eds. Pieter Seuren and Gerard Kempen, 91-126. Amsterdam: John Benjamins.
- Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press.
- Hinterhölzl, Roland. 2006. *Scrambling, remnant movement, and restructuring in west Germanic*. Oxford ; New York, NY u.a.: Oxford Univ. Press.
- Hodler, Werner. 1969. *Berndeutsche Syntax*. Bern: Francke Verlag.
- Hoeksema, Jack. 1988. A constraint on governors in the West Germanic verb cluster. In *Morphology and Modularity: In Honour of H.Schultink*, eds. Martin Everaert, Arnold Evers, Riny Huybregts and Mieke Trommelen, 147-161.
- Hoeksema, Jack. 1994. The history of Dutch Verb Projection Raising. Ms. University of Groningen.
- Hoekstra, Eric. 2009. On the interruption of Verb-Raising clusters by nonverbal material. In *Structure Preserved. Studies in syntax for Jan Koster*, eds. C. Jan-Wouter Zwart and Mark De Vries, 175-183. Amsterdam: John Benjamins.
- Koopman, Hilda, and Szabolcsi, Anna. 2000. *Verbal complexes*. Cambridge, Mass: MIT Press.
- Lötscher, Andreas. 1978. Zur Verbstellung im Zürichdeutschen und in anderen Varianten des Deutschen. *Zeitschrift für Dialektologie und Linguistik* 45:1-29.
- Lötscher, Andreas. 1993. Zur Genese der Verbverdoppelung bei gaa, choo, laa, aafaa ("gehen", "kommen", "lassen", "anfangen") im Schweizerdeutschen. In *Dialektsyntax*, eds. Werner Abraham and Josef Bayer, 180-200. Opladen: Westdeutscher Verlag.
- Müller, Gereon. 1999. Optimality, Markedness, and Word Order in German. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 37:777-818.
- Müller, Stefan. 2002. *Complex predicates: verbal complexes, resultative constructions and particle verbs in German*. Stanford, Calif.: CSLI Publications.
- Neeleman, Ad, and Weerman, Fred. 1993. The balance between syntax and morphology: Dutch particles and resultatives. *Natural Language & Linguistic Theory* 11:433-475.
- Neeleman, Ad. 1994. Complex Predicates. PhD Dissertation, OTS, University of Utrecht.
- Nübling, Damaris. 1995. Die Kurzverben im Schweizerdeutschen. In der Kürze liegt die Würze oder Im Spannungsfeld zwischen Reduktion und Differenzierung. In *Alemannische Dialektforschung. Bilanz und*

New arguments for verb cluster formation at PF

- Perspektiven. Beiträge zur 11. Arbeitstagung alemannischer Dialektologen*, ed. Heinrich Löffler, 165-179. Tübingen, Basel: Francke.
- Ott, Dennis. to appear. An Ellipsis Approach to Contrastive Left-dislocation. *Linguistic Inquiry*.
- Richards, Marc. 2008. Desymmetrization: Parametric Variation at the PF-interface. *Canadian Journal of Linguistics* 53:275-300.
- Robbers, Karin Barbera Maria. 1997. *Non-finite verbal complements in Afrikaans: a comparative approach*. The Hague: Holland Acad. Graphics.
- Roberts, Ian. 1997. Restructuring, Head Movement, and Locality. *Linguistic Inquiry* 28:423-460.
- Salzmann, Martin. 2011. Resolving the movement paradox in Verb Projection Raising. In *Empirical Issues in Syntax and Semantics 8*, eds. Olivier Bonami and Patricia Cabredo Hofherr, 453-485. Paris: CNRS.
- Salzmann, Martin. 2012. Trying to make sense of verb cluster variation. A Swiss German perspective. Paper presented at DGFS 2012, Frankfurt.
- Salzmann, Martin. 2013a. Rule Ordering in Verb Cluster Formation: On the Extraposition Paradox and the Placement of the Infinitival Particle *te/zu*. In *Rule interaction in grammar*, eds. Fabian Heck and Anke Assmann, 65-121. Leipzig: University of Leipzig.
- Salzmann, Martin. 2013b. 213 – 3rd construction or Verb Projection Raising? Paper presented at the Workshop on Swiss German Syntax, Arezzo, September 6, 2013.
- Schallert, Oliver. 2010. Syntax des Vorarlberger Alemannischen: Ergebnisse eines Forschungsprojekts. *Montfort* 62:35-67.
- Schallert, Oliver. 2012. Untersuchungen zu Ersatzinfinitivkonstruktionen in den Vorarlberger und Liechtensteiner Dialekten. Doctoral Dissertation, University of Marburg.
- Schmid, Tanja, and Vogel, Ralf. 2004. Dialectal Variation in German 3-Verb Clusters: A Surface-Oriented Optimality Theoretic Account. *The Journal of Comparative Germanic Linguistics* 7:235-274.
- Schmid, Tanja. 2005. *Infinitival syntax. Infinitivus pro participio as a repair strategy*. Amsterdam: Benjamins.
- Schmid, Tanja, Bader, Markus, and Josef, Bayer. 2005. Coherence - an Experimental Approach. In *Linguistic evidence*, eds. Marga Reis and Stephan Kepser, 435-456. Berlin: De Gruyter.
- Schmidt, Christa. 2000. Die Verbverdoppelung im Zürichdeutschen. MA thesis.
- Schönenberger, Manuela, and Penner, Zvi. 1995a. Cross-Dialectal Variation in Swiss German: Doubling Verbs, Verb-Projection Raising, Barrierhood, and LF Movement. In *Studies in comparative Germanic Syntax*, eds. Hubert Haider, Susan Olsen and Sten Vikner, 285-305. Dordrecht: Kluwer Acad.
- Schönenberger, Manuela, and Penner, Zvi. 1995b. Probing Swiss-German Clause Structure by Means of the Placement of Verbal Expletives: Tun 'Do' Insertion and Verb Doubling. In *Topics in Swiss German Syntax*, eds. Zvi Penner and Thomas Bader, 291-330. Bern, Switzerland: Peter Lang.
- Seiler, Guido. 2004. On Three Types of Dialect Variation and Their Implications for Linguistic Theory: Evidence from Verb Clusters in Swiss German Dialects. In *Dialectology meets Typology. Dialect Grammar from a Cross-Linguistic Perspective*, ed. Bernd Kortmann, 367-399. Berlin, Germany: Mouton de Gruyter.
- Staub, Friedrich, Tobler, Ludwig, Bachmann, Albert, and Antiquarische Gesellschaft (Zürich). 1881. *Schweizerisches Idiotikon : Wörterbuch der schweizerdeutschen Sprache*. Frauenfeld: Huber.
- Sternefeld, Wolfgang. 2006. *Syntax: eine morphologisch motivierte generative Beschreibung des Deutschen*. Tübingen: Stauffenburg Verlag.
- Suter, Rudolf. 1976. *Baseldeutsch-Grammatik*. Basel: Merian.
- ter Beek, Janneke. 2008. *Restructuring and Infinitival Complements in Dutch*. Utrecht: LOT.
- van Riemsdijk, Henk. 1998. Head Movement and Adjacency. *Natural Language & Linguistic Theory* 16:633-679.
- van Riemsdijk, Henk. 2002. The unbearable lightness of GOing: The projection parameter as a pure parameter governing the distribution of elliptic motion verbs in Germanic. *The Journal of Comparative Germanic Linguistics* 5:143-196.
- Wagner, Michael. 2004. Prosody as Diagonalization of Syntax. Evidence from Complex Predicates. *Proceedings of the 34th Annual Meeting of the North East Linguistic Society*.
- Weber, Albert. 1987. *Zürichdeutsche Grammatik: ein Wegweiser zur guten Mundart*. Zürich: Schweizer Spiegel Verlag.
- Williams, Edwin. 2004. The structure of clusters. In *Verb clusters: A study of Hungarian, German, and Dutch*, eds. Katalin É Kiss and Henk van Riemsdijk, 173-201. Amsterdam: John Benjamins.

New arguments for verb cluster formation at PF

- Wöllstein-Leisten, Angelika. 2001. *Die Syntax der dritten Konstruktion : eine repräsentationelle Analyse zur Monosententialität von "zu"-Infinitiven im Deutschen*. Tübingen: Stauffenburg-Verl.
- Wurmbrand, Susanne. 2004a. West Germanic verb clusters: The empirical domain. In *Verb clusters: A study of Hungarian, German, and Dutch*, eds. Katalin É Kiss and Henk van Riemsdijk, 43-85. Amsterdam: John Benjamins.
- Wurmbrand, Susanne. 2004b. Syntactic vs. post-syntactic movement. Proceedings of the 2003 annual conference of the Canadian Linguistic Association. 284-295.
- Wurmbrand, Susanne. 2005. Verb clusters, verb raising, and restructuring. In *The Blackwell companion to syntax. Volume 5*, eds. Martin Everaert and Henk van Riemsdijk, 227-341. Oxford: Blackwell.
- Wurmbrand, Susanne. 2007. How Complex Are Complex Predicates? *Syntax* 10:243-288.
- Wurmbrand, Susi. 2004c. Two types of restructuring--Lexical vs. functional. *Lingua* 114:991-1014.
- Zwart, Jan-Wouter. 1993. Dutch syntax. A Minimalist approach. PhD Dissertation, University of Groningen.
- Zwart, Jan-Wouter. 1994. Dutch is Head Initial. *The Linguistic Review* 11:377-406.
- Zwart, Jan-Wouter. 1995. A Note on Verb Clusters in the Stellingwerf Dialect. In *Linguistics in the Netherlands 1995*, eds. Marcel den Dikken and Kees Hengeveld, 215-226. Amsterdam: John Benjamins.
- Zwart, Jan-Wouter. 1996. Verb clusters in continental West Germanic dialects. In *Microparametric syntax and dialect variation*, eds. James R. Black and Virginia Motapanyane, 229-258. Amsterdam: John Benjamins.
- Zwart, Jan-Wouter. 2007. Some notes on the origin and distribution of the IPP-effect. *Groninger Arbeiten zur Germanistischen Linguistik* 45:77-99.
- Zwart, Jan-Wouter. 2011. *The syntax of Dutch*. Cambridge, UK ; New York: Cambridge University Press.