

New Evidence for Verb Cluster Formation at PF and a Right-Branching VP*

Martin Salzmann

University of Leipzig

1. Introduction: Verb Clusters in West Germanic

West Germanic OV-languages are famous for their verb clusters, i.e. the phenomenon that the verbal elements of a clause all occur together clause-finally (under verb second, where the finite verb moves to C, only the non-finite verbs occur together). In what follows, numbers indicate the embedding relations, i.e. 1 stands for the highest, i.e. the embedding verb, 2 for the immediately embedded verb etc.:

- (1) dass man darüber [reden₃ können₂ sollte₁] 321
that one about.it talk.INF can.INF should
'that one should be able to talk about it' (Standard German)

Importantly, this sequence of verbs forms is impenetrable (hence the term *cluster*): if one tries to extrapose a complement of the lexical verb, the extraposee has to target the higher VP, i.e. VP1, it cannot attach to the lower one, i.e. VP2 (cf. Haider (2003: 92ff.)):

- (2) dass man [_{VP1} [_{VP1} [_{VP2} [_{VP2} ___₁ reden] *darüber₁] kann] √ darüber₁]
that one talk.INF about.it can.INF about.it
'that one can talk about it' (Standard German)

However, the cluster property is limited to descending orders (i.e. where the dependent verbs precede the selecting verbs). With the exception of Standard Dutch (cf. fn. 4), sequences of verbs in ascending order (where the dependent verbs follow the selecting verbs) can be interrupted, instantiating so-called Verb-Projection Raising (VPR):

- (3) dass de Hans wett₁ es Buech läse₂
that the John wants a book read.INF
'that John wants to read a book' (Swiss German)

* Versions of parts of this work were presented at NELS 42 (University of Toronto), at the University of Leipzig, and at the DGfS 2012 (University of Frankfurt). I am grateful to the audiences at these occasions for helpful discussion, especially to Fabian Heck, Gereon Müller, and Hubert Haider. The paper has also benefitted from discussions with Doreen Georgi and Peter Culicover.

The second striking fact about verb clusters is the massive cross-linguistic, dialectal and also inter-speaker variation concerning the possible orders of the verbal elements and the fact that in many cases, several orders are grammatical for a single speaker. In the case of 3-verb-clusters, for instance, of the 6 logically possible orders, 123, 132, 321 and 312 occur quite frequently while 231 is quite rare and 213 is often argued not to exist. For reasons I do not have space to discuss, I will assume that one cannot rule out any of the orders (for instance, 213 is fully grammatical with Swiss German perception and benefactive verbs, cf. Lötscher (1978: 3, fn. 2)). The factors governing the variation are still poorly understood and will not be discussed here (see Culicover (2013) for a recent processing-based account). The literature contains various proposals to account for the cluster property. In many, it is the consequence of complex head formation. A complex head can arise through direct base-generation (e.g. Bader and Schmid (2009), Haider (2003)), via syntactic V-incorporation (e.g. Evers (1975)) or via reanalysis (Haegeman and van Riemsdijk (1986)). In other accounts, there is no complex head formation; the illusion of a cluster arises because constituents of the embedded VPs must be moved out to reach a position where they are licensed (e.g. den Dikken (1995)). Various mechanisms have been proposed to derive the different orders: base-generation (e.g. Bader and Schmid (2009)), V-incorporation (e.g. Evers (1975)), VP-movement (Barbiers (2005)), or inversion at PF (e.g. Haegeman and van Riemsdijk (1986), Wurmbrand (2004a)). In this paper, I will argue that there is complex head formation. 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. The paper is organized as follows: In section two, I introduce the empirical domain where the crucial evidence can be found, viz. the Swiss German motion verb construction. Section three presents a crucial asymmetry between ascending and descending orders in this construction. Section four shows that a left-branching approach cannot derive the asymmetry. Section five contains the proposal. In section six I show that post-syntactic cluster formation provides a new solution to what I refer to as the cluster paradox. Section seven concludes.

2. Swiss German Motion Verbs

Motion verbs in Swiss German take an infinitival complement which (unlike in the standard language) is obligatorily preceded by the particle *go*, cf. Lötscher (1993), Brandner and Salzmann (2012):

- (4) Ich gang **go** bügle.
 I go.1s PRT iron.INF
 'I'll go iron.' (Swiss German)

The fact that *go* is obligatory and is only selected by motion verbs suggests that it is a head. What is crucial for our purposes here is that there are good reasons to believe that *go* is a non-finite verbal element that heads its own VP-projection (cf. the proposals in Schönenberger and Penner (1995) and van Riemsdijk (2002)): First, like modal verbs, it shows obligatory restructuring effects (the verbal elements behave like a mono-clausal unit); for instance, weak pronouns that are arguments of the infinitive have to be fronted to a high position in the matrix clause:

- (5) a. *Ich gang jetzt [go s hole]. b. Ich gang s₁ jetzt [go __ hole].
 I go.1s now PRT it get.INF I go.1s it now PRT get.INF
 'I'll go get it now.' (Swiss German)
- (6) 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. Dutch *om*). Second, as with non-finite modals, the position of *go* is variable within the VP, both under verb second, (7)/(9), and V-final, (8)/(10), thus instantiating Verb Projection Raising (VPR: the a and b examples) and Verb Raising (VR: the c-examples)

- (7) 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.'
- (8) 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
- (9) 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.'
- (10) 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

Evidence that *go* heads its own projection can be seen in the following example where *go* is topicalized with elements preceding it:

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

I thus adopt the conclusion of the above-mentioned sources and treat *go* as a non-finite verb heading its own VP-projection.¹

¹ At first sight, it may seem tempting to analyze *go* as the spell-out of a lower copy of the matrix motion verb, especially because it is phonetically similar to the infinitive of 'go'. However, there are good arguments against pursuing such an analysis: First, *go* occurs after various motion verbs, including *räne*

3. The Puzzle: Descending Orders

The motion verb construction behaves in a special way in descending orders (i.e. where the finite verb occurs at the end of the verb cluster, unlike in the ascending order in (8)): the particle must precede the infinitive, leading to what seems to constitute a 231 order:

- (12) *dass* *i* [**go**₂ *poschte*₃ *gang*₁] 231
 that I PRT do.shopping.INF go.1sg
 ‘that I go to do shopping’ (Swiss German)

This is surprising given that most 3-verb-clusters prohibit the 231 order in Swiss German; in fact, apart from certain perception and benefactive verbs (Löttscher (1978: 3, fn. 2)), the motion verb construction is the only one where the finite verb can occur at the end of the cluster; most other cluster types are either 123 or 132.

- (13) * *wele*₂ *läse*₃ *ha*₁ 231 (Aux-Mod-Inf-cluster)
 wanted read.INF have.1s (Swiss German)

Importantly, *go* must immediately precede the infinitive in descending orders, VPR is not a possibility; there is thus an asymmetry between ascending and descending orders.²

- (14) 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

4. Why a Left-Branching Approach Does Not Work

I repeat the crucial contrast from above: Adjacency between *go* and the infinitive is required only in the descending order:

- (15) 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

‘run’, *lauffe* ‘run’, *springe* ‘run’ or *schicke* ‘send’. 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 motivation in Swiss German. Second, in many dialects (e.g. Zurich German), the form of *go* does not correspond to any of the forms of the verb ‘go’, it is e.g. not identical to the infinitive (which would be *gaa* in ZG). Third, and most importantly, as pointed out in van Riemsdijk (2002: 160, fn. 22), it is unclear which intermediate positions would be spelled-out, especially in V-final structures like (8) where the finite verb is normally assumed to be in its base-position so that no intermediate copies would arise.

² The crucial empirical data are drawn from a questionnaire study that I carried out at the University of Zurich in spring 2008. The informants were students of German linguistics in their second semester.

The facts described for *go* (the restructuring properties, the positional variability and the asymmetry between ascending and descending orders) hold more generally for Swiss German verb doubling, i.e. also for the particles *cho* ‘come’ and *la* ‘let’; cf. e.g. Löttscher (1993) and Brandner and Salzmann (2012).

verbs. Without any further qualifications, this mechanism can derive the following orders (apart from 123): 132 (V_3 inverts with V_2 : 1[32]), 321 (V_3 inverts with V_2 and [V_3V_2] invert with V_1 : [[32]1]), 213 (V_2 inverts with V_1 : [21]3). 231 and 312 can only be derived if there is rebracketing without inversion: In 231, V_2 and V_3 rebracket, and then [V_2V_3] invert with V_1 , resulting in [[23]1]; similarly, in 312, V_1 and V_2 rebracket [V_1V_2] and are then inverted with V_3 : [3[12]]. This is a very powerful mechanism in that it derives all six logically possible orders some of which are quite rare. But given that I do not think that one should rule out any of the orders (cf. section 1), this nevertheless seems to be the right way to go. I assume that all West-Germanic languages make use of this mechanism.

The orders that actually occur in a given variety are due to language-/dialect-particular linearization constraints for verbal elements that limit the power of PF-cluster formation, for instance, that a participle has to precede the auxiliary.

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 (5)–(6). Rather, such effects follow from the small size of a complement (as e.g. in den Dikken (1995)). For instance, a modal only takes a 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. 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, are taken to be larger in size so that they can host their own negation, adverbials etc. As a consequence, they behave more like independent clauses.

The workings of the approach can be illustrated with a simple example: The starting point will be the following structure (Note that linear order does not play a role at this point):

- (18)

dass er	$[_{VP1}$	wett ₁	$[_{VP2}$	es Buech	läse ₂]]	VPR
that he		wants		a book	read.INF	(Swiss German)

In some varieties, this structure can be directly linearized as a right-branching structure (e.g. in Swiss German), in others it cannot (e.g. Standard German, see below). VR-structures (where the verbal elements are adjacent) are derived by (syntactically) moving the XPs out of the embedded VP, cf. (19a) (see below for qualification). Depending on the variety, the verbal elements form a complex head and invert at PF, cf. (19b) (in Swiss German, this is generally optional in 2-V-clusters with V_1 = modal):

- (19)

a.	dass er	$[_{VP1}$	[es Buech] ₁	wett ₁	$[_{VP2}$	__ ₁	läse ₂]]	movement in syntax
	that he		a book	wants			read.INF	
b.	dass er	es	Buech	[läse ₂ +wett ₁]	inversion at PF			
	that he	a	book	read.INF+wants	VR (21)			

The trigger for inversion in this example will be a linearization statement according to which infinitives can (optionally) precede modals. In Standard German, where only the descending variant in (19b) is grammatical, 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 the linearization constraint Inf >> Mod.³

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:

- (20) [Ein Buch lesen₂] will₁ 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 (19) or Verb Projection Raising (18). 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.⁴

The nature of the movement operation needs to be made more precise. It cannot be equated with scrambling because given that VP-complements are linearized to the right of the embedding modal and that the modal eventually occurs clause-finally (in Standard German), movement must not only affect 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. Fronting of those, however, cannot be subsumed under scrambling. If one wants movement to be triggered, then a more general evacuation mechanism is necessary. A possibility would be licensing movement as in Hinterhölzl (2006) where for each type of constituent in the lower VP there is a designated landing site in the projection of the higher verb (I am simplifying Hinterhölzl's proposal for my purposes). While this may strike one as a bit mechanical, it would derive the empirical facts. An alternative would be to assume that constituents related to the lower VP do not have to be generated within that VP. Rather, they can be directly merged (= base-generated) in the higher VP. There are various mechanisms that can be used to implement this, e.g. argument composition and percolation of theta-roles, cf. e.g. Sternefeld (2006: 622ff.). This will not be sufficient for non-arguments, though. A possible solution is suggested in Salzmann (2011), who proposes a mechanism (covert predicate raising) that can be used to base-generate both arguments and adjuncts outside the projection they are semantically related to. I will not choose between the movement or the base-generation option here. For reasons of

³ 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 given that restructuring effects also obtain without cluster formation, namely in VPR (cf. Haegeman (1992: 109ff.), Salzmann (2011)), it seems preferable to separate the syntax that leads to monoclausal behavior and the cluster-forming operation.

⁴ Standard Dutch, which has ascending orders but no VPR, raises interesting questions in this context. One way of deriving the facts is to say that the verbs do form a cluster, but that they rebracket without inversion. Since rebracketing without inversion may be needed independently (cf. below), this is probably not detrimental. Alternatively, one might argue that Dutch is subject to stricter linearization requirements in that it requires all material related to the lexical VP to precede all verbal elements. Possible evidence for the second position comes from the fact that Standard Dutch does allow certain elements in the cluster, e.g. resultative adjectives.

simplicity, I will formulate the analysis in the following sections in terms of movement, but it should be borne in mind that the base-generation alternative is just as viable.⁵

5.2 Applying the Model to the Motion Verb Construction

We will start with the derivation of the ascending orders in (8), repeated for convenience:

- (21) 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 starting point will look as follows:

- (22) 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 (21a) is a direct linearization of (22): neither XP-movement nor cluster-formation is necessary. (21b/c) are derived by moving XPs from the lowest VP to the projection of *go* (alternatively, they are directly generated in the higher projection, cf. the previous subsection):

- (23) 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
 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 (14), repeated from above?

- (24) 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 basis is again as in (22). In the derivation of (24c) the arguments of the lexical verb move to (or are base-generated in) the projection of the matrix motion verb.⁶

- (25) 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

⁵ 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. I take this to be an advantage since these varieties have been largely unaffected by prescriptive pressure and thus arguably represent more natural systems.

⁶ Importantly, (25) is also a possible surface form. Movement to VP₁ is thus not restricted to descending orders.

To derive the apparent 231 order [*go chauffe gang*], one could be tempted to invert the matrix motion verb *gang* 'go.1s' with its syntactic sister [*go chauffe*] 'PRT buy.INF'. But once it is possible to invert *gang* with its sister, one could, given the structures in (22) or (23a), also derive the ungrammatical orders in (24a/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)). Importantly, without any restrictions, inversion will overgenerate, i.e. it might invert only *gang* + *go* in (25), leading to the ungrammatical (26):

- (26) *dass i de Muetter es Buech go+gang chauffe] 213
 that I the mother a book PRT+go.1s buy.INF (Swiss German)

What we need to achieve is that the motion verb inverts with [*go chauffe*]. Given that inversion only affects terminals, *go* + *chauffe* first need to rebracket for this to become possible. Crucially, rebracketing does not have to be stipulated here. Rather, it follows from an independent property of *go* (Lötscher (1993)): it is a (category-insensitive) proclitic (see 5.3 for evidence) and thus cliticizes onto the infinitive, forms a unit with it (= rebracketing) so that the entire [*go+inf*] complex can be inverted with the matrix verb:

- (27) 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 fact that a prosodic property of *go* is involved not only explains why we obtain an order that is very restricted otherwise; 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* just leans onto the following constituent, it does not target a particular syntactic position). Furthermore, the facts show that cyclicity is crucial: cliticization of *go* (= rebracketing/string-vacuous local dislocation) precedes inversion (the derivation can be compared with the placement of Latin *-que* 'and' in Embick and Noyer (2001: 575f.)).

The trigger for the inversion will be an optional linearization statement: The infinitive can optionally precede the motion verb (motion V >> V; V >> motion V). Since *go* forms a unit with the infinitive, it is as if we are dealing with a 2-V-cluster where inversion between main verb and infinitive is usually optional in Swiss German, cf. (19).

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 (22) or (23a)). Inversion is blocked in both cases since we do not have the required input structure with adjacent verbs as in (25). As a consequence, only an ascending order is possible, viz. (21a/b). One may wonder what happens to *go* if it precedes a DP as in the ascending orders (21a/b). I will assume that since *go* is not category-sensitive, it can also cliticize onto a determiner/noun (the indefinite article is pro-clitic as well):

- (28) 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 [go+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. Inversion is blocked because it only applies to verbal elements; [*go+es+Buech*], however, is not obviously verbal: in Embick and Noyer (2001) it is assumed that – as in syntax – elements that are adjoined at PF are not the head of the constituent. Consequently, [*go+es+Buech*] counts as nominal (*go* and the indefinite article are cliticized onto the noun, not vice versa). [*go+chauffe*] in (27), however, certainly is verbal and thus allows for inversion.

Things are different with non-referential/incorporated nouns: they become part of the non-finite verb (they cannot be separated from it); in those cases, *go* cliticizes onto the [N+V]-complex, and finally, there is inversion between the matrix motion verb and the [*go+N+V*]-complex:

- (29) a. *dass i gang [go=[Fläisch+ chauffe]]* → inversion:
 that I go.1s PRT meat buy.INF
 ‘that I go to buy meat’
 b. *dass i [[go=[Fläisch+ chauffe]]+ gang]*
 that I PRT meat buy.INF go.1s (Swiss German)

5.3 Evidence for the Clitic Nature of *Go*

Evidence for the clitic nature of *go* comes from two facts: First, it is phonetically reduced; 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):⁷

- (30) a. *[*De Muetter hälffe*]₁ *gang I nöd go* ___₁ .
 the mother help.INF go.1sg I not PRT
 b. ?[*De Muetter hälffe*]₁ *han i nöd wele* ___₁ .
 the mother help.INF have.1s I not wanted

This concludes my analysis of the orders in the *go*-construction.⁸

⁷ As pointed out by a NELS-reviewer, this predicts that examples like (30a) should be rescuable by inserting a DP onto which *go* can cliticize (basically as in the ascending orders in (8a/b), (21a/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. I don’t have a good explanation for this fact, but the problem seems to be more general in that the corresponding example with a modal is also 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.

⁸ Verb doubling in West Flemish (Haegeman (1990)) is very similar to the Swiss German *go*-construction. The major difference is that the double *gon* only cliticizes onto verbs (and verbs with incorporated nouns), implying that the resulting structure will always be a VR-structure, VPR being ruled out:

- (i) *dan ze atent goan *gon en boek gon lezen*
 that they always go.3p PRT a book PRT read.INF
 ‘that they always go read a book’

6. Post-Syntactic Cluster Formation Solves the Cluster Paradox

In this section I will show that the cluster formation approach proposed here generally has desirable consequences for the analysis of verb clusters because it provides a novel solution to what I will call the cluster paradox:⁹ 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. Consider the following example with remnant VP-topicalization and verb second:

- (31) [reden]₂ sollte₁ man darüber ___₂ können ___₁.
 explain should one about.it can.INF
 ‘One should be able to talk about it.’ (Standard German)

Under the standard assumption that V2 is derived from a V-final structure, we would expect this underlying structure to look like the example in (1). If this involves a complex head, it is unclear how (31) can be derived – it seems that both verbal elements have excorporated from the cluster.

One might try to solve the problem by allowing for excorporation or by assuming that cluster formation is optional. The first option is conceptually undesirable, the second option is empirically problematic: There is evidence that cluster-formation is obligatory clause-finally: in the following example (a more complex version of (2)), extraposition cannot target VP₂ or VP₃; rather, the extraposee has to attach to VP₁ (the structures indicated presuppose a left-branching structure for ease of exposition only, for a right-branching structure, cf. below; for the original observation, cf. Haider (2003: 92ff.)):

- (32) dass man [VP₁ [VP₁ [VP₂ [VP₃ [VP₃ ___₁ reden] *darüber₁]
 that one talk.INF about.it
 können] *darüber₁] sollte] √darüber₁]
 can.INF about.it should about.it
 ‘that one should be able to talk about it’ (Standard German)

This result follows if there is cluster-formation = complex head-formation, e.g. by means of V adjunction (e.g. Evers (1975)) or if it is assumed that the entire cluster is base-generated as e.g. in Bader and Schmid (2009). Importantly, however, under VP-topicalization, extraposition *can* target VP₃:

- (33) [VP₃ [VP₃ ___₁ reden] darüber₁]₃ sollte₂ man schon ___₃ können ___₂.
 talk.INF about.it should one indeed can.INF

This constitutes a paradox: cluster-formation must be obligatory to rule out (32), but will simultaneously make (33) impossible. One way out seems to be to opt for obligatory cluster formation + excorporation of the finite verb. However, this will not be sufficient because the non-finite modal *können* was also part of the cluster in (32). To derive (33), it would also have to excorporate from the cluster so that the remnant VP can be

As in Standard German, constituents of the lexical VP will have to move out for the derivation to converge.

⁹ Further evidence for post-syntactic cluster formation is presented in Wurmbrand (2004b) who shows that cluster orders sometimes depend on the morphological form of a verb.

topicalized.¹⁰ However, there is no evidence that a verb in clause-final position would ever undergo movement; consequently, excorporation of the V_2 *können* seems highly unlikely so that (33) cannot be derived.

This implies that neither excorporation nor optionality of cluster formation is sufficient. It thus remains completely unclear how both (32) and (33) can be derived under an approach that assumes a complex head in syntax. Conversely, an approach without complex head formation will be hard pressed to account for (32).¹¹

Under the PF-perspective proposed here, the crucial property that solves the paradox is the timing: In the cases with apparent excorporation, i.e. (31) and (33), V-to-C-movement and remnant topicalization occur in syntax and thus before PF-cluster formation can apply; i.e. the syntactic operations destroy the context for cluster-formation; we are dealing with a bleeding relationship.¹² The derivation of (32) and (33) proceeds as follows: the base-structure will be the following right-branching VPR-structure:

- (34) 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 VP_1 , VP_2 or VP_3 ; this will in any case lead to adjacency between the three verbs so that cluster-formation + inversion is possible (since inversion is necessary in Standard German, extrapolation or, alternatively, movement of the PP to VP_1 /base-generation in VP_1 is obligatory):¹³

- (35) 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} 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

Note that for the derivation of (32), (35a–c) are all possible base structures since adjacency obtains in all three. (33) then simply involves topicalization of VP_3 in (35a) and V-to-C-movement of V_1 . Again, the syntactic operations bleed cluster formation.¹⁴

¹⁰ Given the structure in (32), this would arguably involve topicalization of VP_1 and not of VP_3 .

¹¹ Wurmbrand (2007) proposes that extrapolation can in principle always target any of the VPs (VP_3 , VP_2 , VP_1); whether the extraposed copy can be phonetically realized at PF is determined by prosodic principles (extraposed material must occur at the edge of a prosodic constituent); in (32), this is only the case if extrapolation targets VP_1 ; if it targets VP_2 or VP_3 , the in-situ copy must be realized; in (33), however, the extraposed copy adjoined to VP_3 can be realized since it is at the edge of a prosodic constituent.

¹² 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). The fact that the finite modal precedes the non-finite modal (which is ruled out in the verb-final structure) at surface structure in (31) and (33) therefore does not lead to a clash.

¹³ There will thus be successive cluster-formation leading to [[32]1]. In both instances, cluster formation + inversion is triggered by a linearization constraint that requires infinitives to precede modals.

¹⁴ The cluster-paradox also obtains under a right-branching VP + *syntactic* verb cluster-formation (= complex head formation). To derive (32), one can use either (34) or (35a–c) as a basis. As long as there is successive head-incorporation, we get a grammatical result. Since verb-final structures always involve a descending (i.e. 321) order in Standard German, verb-incorporation must be taken to be obligatory.

7. Conclusion and Outlook

I have argued that the ordering facts in the Swiss German motion verb construction provide new evidence for a post-syntactic treatment of verb cluster formation and for a right-branching VP. More specifically, the ordering restrictions on the particle *go* in descending orders require an analysis that makes crucial reference to linear order and order-related prosodic properties. The analysis proposed here thus differs from previous post-syntactic approaches in that 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 paradox found with extraposition and VP-topicalization. On a more general level, the analysis can be considered hybrid in that it derives the syntactic and order-related properties of verb cluster constructions by independent means, viz. VPs with limited structure and complex-head formation at PF.

References

- Bader, Markus, and Schmid, Tanja. 2009. Verb clusters in colloquial German. *The Journal of Comparative Germanic Linguistics* 12:175-228.
- 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.
- 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.
- Culicover, Peter. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.
- den Dikken, Marcel. 1995. Verb (Projection) Raising, Scope, and Uniform Phrase Structure. *Proceedings of NELS* 25:95-110.
- Embick, David, and Noyer, Rolf. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32:555-595.
- Evers, Arnold. 1975. The transformational cycle of Dutch and German. Doctoral Dissertation, University of Utrecht.
- 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 *go* in West Flemish. In *Linguistics in the Netherlands 1990*, eds. Reineke Bok-Benema and Peter Coopmans, 81-90. Amsterdam: John Benjamins.

However, at the same time, it must be blocked in cases like (31) and (33). It is not clear how this could be ensured: if incorporation is triggered by an attracting feature on the immediately higher verb, this feature must not be present if a) the finite verb undergoes V2 and b) there is VP 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 move to C. In the case of VP-topicalization, however, the selecting verb would have to lack an attracting feature if its VP-complement is marked for undergoing topicalization. This obviously cannot be done in a straightforward way.

- Haegeman, Liliane. 1992. *Theory and Description in Generative Syntax. A case study in West Flemish*. Cambridge u.a.: Cambridge University Press.
- 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.
- Hinterhölzl, Roland. 2006. *Scrambling, remnant movement, and restructuring in west Germanic*. Oxford ; New York, NY u.a.: Oxford Univ. 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.
- Richards, Marc. 2008. Desymmetrization: Parametric Variation at the PF-interface. *Canadian Journal of Linguistics* 53:275-300.
- 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.
- 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.
- Schönenberger, Manuela, and Penner, Zvi. 1995. 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.
- Sternfeld, Wolfgang. 2006. *Syntax: eine morphologisch motivierte generative Beschreibung des Deutschen*. Tübingen: Stauffenburg Verlag.
- 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.
- 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. 2007. How Complex Are Complex Predicates? *Syntax* 10:243-288.

Martin Salzmann
University of Leipzig
Department of Linguistics
Beethovenstraße 15
D-04107 Leipzig

martin.salzmann@uni-leipzig.de