

On the limits of variation in West-Germanic verb clusters

Evidence from displaced morphology and extraposition for the existence of clusters with 213 order

Martin Salzmann – University of Leipzig – martin.salzmann@uni-leipzig.de

April 1st, 2016

Abstract:

Recent work on verb clusters within West-Germanic has argued in favor of restrictive theories of cluster formation that only generate a subset of the logically possible orders in three-verb clusters, explicitly ruling out the 213 order. In this context it is remarkable that Swiss German features a verb cluster-like construction with an unmarked 213 order. I will argue that this construction indeed represents a proper verb cluster and not an instance of the 3rd Construction, which also allows for the 213 order. Based on new diagnostics, viz. displaced *zu* and relative clause extraposition, I will show that verb clusters/Verb Projection Raising differ from the 3rd Construction with respect to a fundamental structural relation: while the lexical VP is a structural complement in the former, it is in a non-complementation relation in the latter. Applying the diagnostics to the Swiss German 213 construction delivers a clear result: the construction has the hallmarks of complementation and thus represents a genuine verb cluster. I conclude from this that theories of verb clusters must be more powerful than previously claimed in that they must be able to generate all six logically possible orders, including the 213 order.

1 Introduction: verb cluster orders

One prominent feature of West-Germanic OV-languages like Dutch and German is the clustering of verbal elements at the end of the clause in V-final structures, as in the following example (under verb second, where the finite verb moves to C, only the non-finite verbs occur together):¹

- (1) dass er das Buch lesen₃ können₂ muss₁
that he the book read.INF can.INF must.3SG
'that he must be able to read the book'

Standard German

Such sequences of verbs are referred to as verb clusters (for a detailed overview, cf. Wurmbrand 2005, to appear). Verb clusters usually involve the combination of a lexical verb with one or several functional verbs (auxiliaries, modals) as in (1) or combine several (more or less) lexical verbs: some of these take a bare infinitival complement (e.g. perception verbs like 'see' or benefactives like 'help'), cf. (2-a), while others select a so-called *zu*-infinitive, where the dependent verb is accompanied by the particle *zu* (in German) or *te* (in Dutch), see (2-b):

- (2) a. dass er sie das Buch lesen sah
that he her the book read.INF see.PST.3SG
'that he saw her read the book'
b. dass er das Buch zu lesen versprach
that he the book to read.INF promise.PST.3SG
'that he promised to read the book'

Standard German

Apart from the clustering, there are two further striking properties that are characteristic of West-Germanic verb clusters: they display restructuring effects and show massive cross-linguistic/dialectal and even intra-speaker variation with respect to the possible orders of the verbal elements.

¹The glosses follow the Leipzig glossing rules, see <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>. The number indices on the verbs indicate the hierarchical relations, i.e. 1 stands for the highest verb in the government sequence, 2 for the immediately dependent verb, etc.

Since this paper will focus on the second aspect, I will be very brief concerning restructuring effects: despite their sometimes bi- or multi-clausal semantics (cf. the translation of (1) and (2)), verb clusters behave like a monoclausal unit for syntactic operations, i.e. they show so-called restructuring effects and thus instantiate what is often called a coherent construction. While verbs taking a bare infinitival or a participial clause as their complement are obligatorily restructuring, verbs taking a *zu*-infinitive form different classes: some are obligatorily restructuring (e.g. *scheinen* ‘seem’ in Standard German), some are optionally restructuring (e.g. *versuchen* ‘try’ in Standard German) and others do not allow for restructuring (e.g. *bedauern* ‘regret’ in Standard German).

Research on verb clusters has accumulated a plethora of restructuring effects (see e.g. Haider 2010: 310-321 for a list of such effects in German, Wurmbrand 2001 for a more fine-grained classification into different degrees of restructuring and Reis and Sternefeld (2004) for a critique thereof). I will illustrate coherence by means of scrambling and pronoun fronting.² It is well-known that German prohibits scrambling/pronoun fronting across a finite clause-boundary:

- (3) *dass [den Roman]₁/ihn₁ keiner will, [CP dass ich ___₁ lese]
 that the novel/him no one want.3SG that I read.1SG
 ‘that nobody wants that I read the novel/it’ *Standard German*

However, scrambling/pronoun fronting from the projection of the lowest verb in a verb cluster is unproblematic (I label the projection of the lexical verb as XP to remain theory-neutral at this point):

- (4) dass [den Roman]₁/ihn₁ keiner [XP ___₁ lesen] wollte
 that the novel/him no one read.INF want.PST.3SG
 ‘that nobody wanted to read the novel/it’

Since the embedded object precedes the matrix subject, it is clear that it has left the embedded clause.

I now turn to variation in the possible orders of the verbal elements in the cluster. While in 2-verb clusters, there are only two possible orders, both of which are instantiated across West-Germanic as well as within a dialect/a single variety (e.g. in Standard Dutch or in certain Swiss German dialects), in 3-verb clusters, there are six logically possible orders. Focusing on verb clusters with functional verbs (modals and auxiliaries) as V1 and V2 for the moment, the existence of the orders 123, 132, 321, 312 and 231 within West-Germanic is undisputed (while largely absent in German varieties, 231 occurs frequently in West-Flemish and Afrikaans, cf. Wurmbrand 2004, Barbiers 2005, Biberauer 2013). The following examples illustrate the various orders:

- (5) a. dat Hans het boek heeft₁ willen₂ lezen₃
 that John the book have.3SG want.INF read.INF
 ‘that John wanted to read the book’ *123 Standard Dutch*
- b. dass Hans das Buch hat₁ lesen₃ wollen₂
 that John the book have.3SG read.INF want.INF
 ‘that John wanted to read the book’ *132 Standard German*

²These are the diagnostics that most of the literature considers solid. Pronoun fronting is sometimes taken to be more liberal than scrambling, see Wurmbrand (2001: 267f.); similarly, scrambling has been claimed to be possible with non-restructuring verbs if the scrambled DP receives a focus interpretation (Wurmbrand 2001: 269f.). The relevant scrambling examples in this paper are all acceptable without a focus interpretation.

I will use these diagnostics because they also work if the transparent XP contains more structure, as is often the case in Verb Projection Raising (VPR), i.e. if the verbal complex contains non-verbal material. In such configurations, some clause-union phenomena such as the lack of an independent tense or negation domain may be lacking while scrambling or pronoun fronting is still possible.

- c. dass Peter das Buch gelesen₃ haben₂ muss₁
 'that Peter the book read.PTCP have.INF must.3SG
 'that Peter must have read the book' 321 *Standard German*
- d. dass Hans das Buch lesen₃ hat₁ wollen₂
 that John the book read.INF have.3SG want.INF
 'that John wanted to read the book' 312 *Colloquial German*
- e. dat hy die medisyne kon₂ drink₃ het₁
 that he the medicine could.INF drink.INF have.3SG
 'that he could drink the medicine' 231 *Afrikaans*

Importantly, the general availability of these five orders does not imply that a given variety will instantiate all these orders; in fact, most varieties will only allow for a subset of them. Furthermore, the availability of a certain cluster order often depends on the cluster type, i.e. the type of functional verbs involved. For instance, in Standard German, the 132 order is generally only possible if V1 is an (perfective or future) auxiliary but not if it is a modal.

Crucially, the 213 order has been claimed to be absent in the major cluster types, see e.g. Zwart (1996), Seiler (2004), Wurmbrand (2004), Barbiers (2005) and Abels (2011, 2015). The major cluster types are Aux-Mod-Inf (e.g. 'has want read' and 'will want read'), Mod-Mod-Inf (e.g. 'must can read') and Mod-Aux-Ptcp (e.g. 'must have read'), but the generalization arguably also holds for Aux-Aux-Ptcp (e.g. 'is been read' as in the passive, or 'has had read' as in the double perfect, see Author xxxx).³

Against this background, it is remarkable that 213 orders are absolutely unmarked in certain Swiss German cluster types, viz., those with perception verbs (6-a), (6-b), phasal verbs (6-c), (6-d) and benefactives (6-e), (6-f) as V2 taking a bare infinitive as V3 ((6-a/c/e) are from Lötscher 1978: 3, 9; cf. Schallert 2012: 260f., 278f., Schallert 2014: 205, 227f. for equivalent data from Vorarlberg German):⁴

- (6) a. wil er en ghöört₂ hät₁ chooo₃
 because he him hear.PTCP have.3SG come.INF
 'because he heard him come'
- b. Han s voll lustig gfunde det z stah wo ich scho Stars gseh₂
 have.1SG it very funny find.PTCP there to stand.INF where I already stars see.PTCP
 han₁ stah₃
 have.1SG stand.INF
 'I found it funny to stand there where I have seen stars stand.'
<http://schnalletagebuech.blogspot.fr/2013/09/volleyball-hollywood.html>, accessed March 21, 2016
- c. Wo s aagfange₂ hät₁ rägne₃, simer i d Beiz
 when it start.PTCP have.3SG rain.INF are.1PL in the pub
 'When it started to rain, we went to the pub.'

³There are a few claims to the contrary: Cooper (1995: 154) reports 213 for Zurich German Mod-Mod-Inf clusters; Schmid and Vogel (2004) report it for Aux-Mod-Inf clusters in Rheiderländer Platt (Low German), the dialect of St. Gall and the dialect of Meran; Schwalm (2013: 65f., 69f., 81f., 86f.) reports 213 in Hessian Mod-Aux-Ptcp, Mod-Mod-Inf and Aux-Mod-Inf clusters; Schallert (2012: 285) reports 213 in Vorarlberg German Aux-Mod-Inf clusters; finally, den Dikken (1994: 82f.) reports 213 in Middle English clusters with Fut-Mod-Inf. Although not all sources are explicit about this, the 213 order does not seem to constitute the default order in any of these cases.

⁴V1 is normally the perfective auxiliary; examples where V1 is a modal are attested as well but are much less frequent (cf. Schallert 2012: 267 for a similar observation in Vorarlberg German):

- (i) jetzt weiss i won i afange₂ muess₁ sueche₃
 now know.1SG I where I start.INF must.1SG search.INF
 'Now I know where I have to start searching.'

<http://www.chefkoch.de/forum/2,50,204376/Pfund-um-Pfund-am-31-1-2006.html>, accessed March 15, 2016

- d. leider han i den recht zuegnoh wo i ufghört₂ ha₁
 unfortunately have.1SG I then much put on weight.PTCP when I stop.PTCP have.1SG
 schaffe₃
 work.INF

‘Unfortunately, I put on quite some weight after I stopped working.’ <http://www.babywelten.ch/commun6&Page=2>, accessed March 21, 2016

- e. das er mer ghulffe₂ hät₁ abwäsche₃
 that he me.DAT help.PTCP have.3SG do.the.dishes.INF
 ‘that he helped me do the dishes’
- f. d Nacht, wo de Kobi glehrt₂ hät₁ flüüge₃
 the night, when the Jacob learn.PTCP have.3SG fly.INF
 ‘the night when Jacob learned to fly’

<http://www.godybodmer.info/dasbuch.htm>, accessed March 22, 2016

The grammaticality of these orders is uncontroversial: not only can examples be found in more traditional descriptions like Lötscher (1978) and on the Internet, the grammaticality of 213 orders has also been verified in recent empirical work: according to Glaser (in preparation), the order 213 in a cluster with ‘has learned X-inf’ is accepted by 75% out of nearly 3000 Swiss German speakers (and thus about as many as the 123 order). Note also that 213 is an unmarked order in this cluster type; it freely alternates with 123 and according to Lötscher (1978: 3, fn. 2) and Schmid (2005) also allows for 231, although the latter clearly constitutes a more marked order. 231 orders do not seem to occur elsewhere in Swiss German, except in clusters with motion verbs, see Author (xxxx). The 213 order in this cluster type thus needs to be distinguished from the residual instances of 213 listed in fn. 3 that can be found in other cluster types, where they constitute apparently possible but marked orders. The cluster type in (6) thus represents the only one where 213 represents the default order/an unmarked order.

The special status of some of these cluster types is not a peculiarity of Swiss German, but has been noted in other varieties as well: Zwart (1995) reports 213 with perception verbs in the dialect of Stellingwerf; according to Schmid (2005: 64-72), 213 is found with inchoatives and control verbs (without *te*) in Afrikaans; similarly, Zwart (2007: 80) reports 213 with ‘has learned to do’ in Luxembourgish; finally, Louden (2011: 169, 179) reports 213 with perception verbs and benefactives in Pennsylvania German (as well as with V2 being a causative verb ‘let’ or ‘make’, a motion verb or ‘need’).

While little information is available about the properties of the 213 orders in these other varieties, the Swiss German 213 construction can easily be shown to represent a restructuring configuration. Both scrambling and pronoun fronting are readily possible:^{5, 6}

- (7) dass er {d Chuchi | si} aagfange₃ hät₁ _{t_d Chuchi/si} putze₃
 that he the kitchen her start.PTCP have.3SG clean.INF
 ‘that he started to clean the kitchen/it’

Swiss German

This is unsurprising given that verbs taking bare infinitival complements are usually regarded as restructuring verbs. In Standard German, for instance, restructuring is obligatory with bare infinitives (cf. Reis and Sternefeld 2004). Given that Swiss German allows for Verb Projection Raising,

⁵Examples without explicit references are taken from Zurich German. To my knowledge, the facts discussed in this paper hold for all Swiss German dialects.

⁶Pronoun fronting in Swiss German has to be handled with care: Swiss German has three series of pronouns, viz., clitic, weak and strong versions, which are not distinguished in all person/number/gender/case combinations, though, see Weber (1987) for details. Crucially, the clitic versions seem generally unacceptable in infinitival complements, even if the matrix verb is non-restructuring (arguably because there is no proper host). The pronouns of the weak series, however, can remain inside infinitival complements and will therefore be used in what follows.

viz., non-verbal material within the verb cluster, demonstrating the obligatoriness of restructuring in the 213 construction in (6) cannot be done by means of scrambling; the scrambled DP in (7) can thus also occur within the lexical VP, see (8-a). Pronoun fronting, however, is obligatory, showing that we are dealing with obligatory restructuring, see (8-b):⁷

- (8) a. dass er aagfange₃ hät₁ d Chuchi putze₃
 that he start.PTCP have.3SG the kitchen clean.INF
 ‘that he started to clean the kitchen’
 b. dass er {✓si} aagfange₃ hät₁ {*si} putze₃
 that he her start.PTCP have.3SG her clean.INF
 ‘that he started to clean it’

Swiss German

I will argue in this paper that the existence of these 213 orders has important ramifications for the theory of verb clusters because they imply that verb cluster theories have to be powerful enough to generate all six logically possible orders, contrary to recent claims in the literature.

The paper is organized as follows: in section two, I will introduce various theories of verb clusters and their predictions with respect to the possible orders. In section three, I will compare the Swiss German 213-construction with another construction featuring 213 orders, viz., the 3rd Construction. In section four, I will introduce new diagnostics to distinguish between verb clusters proper/Verb Projection Raising on the one hand and the 3rd Construction, viz., displaced morphology and relative clause extraposition. In section five, I will apply the new diagnostics to the Swiss German 213 construction and show that it behaves like a proper verb cluster rather than the 3rd Construction. In section six, I discuss the implications for the theory of verb clusters, in section seven I address the rarity of 213 orders, and section eight concludes. The appendix in section nine briefly discusses diagnostics from the previous literature to distinguish verb clusters/VPR from the 3rd Construction that are inconclusive.

2 Theories of verb clusters

In this section, I will briefly discuss previous theories of verb clusters and their predictions with respect to possible orders in three-verb clusters. Since not all approaches are explicit about the orders they can generate, the discussion below will be partly based on my interpretation of the mechanisms.

Importantly, this discussion is based on the presupposition that verb clusters are the same phenomenon across West-Germanic and that all varieties make use of essentially the same cluster-forming/-ordering mechanism. Note that this is not a necessity as it is in principle conceivable that different varieties have fundamentally different verb cluster grammars. However, all West-Germanic varieties with verb clusters share certain important properties: they always show transparency/restructuring effects, the various orders do not affect the semantics and only (partially) ascending orders allow for Verb Projection Raising. Given these shared properties, it is much more economical and conceptually more attractive to posit just one basic cluster-forming mechanism across the West-Germanic dialects. The cross-linguistic differences in the possible cluster orders will then not be due to very fundamental properties of grammar but rather to more surface-oriented or even extra-grammatical constraints (see e.g. Barbiers 2005, Bader and Schmid 2009, Abels 2015).

⁷For some speakers (arguably of the younger generation), the phasal verbs and the benefactives can also be used with a *z*-infinitive, which entails optional restructuring. This is close to the Standard German pattern where the phasal verbs require a *zu*-infinitive while the benefactives only require them if the infinitival complement is extraposed; in intraposed position, *zu* is optional.

In what follows, I will distinguish between restrictive and powerful theories. Restrictive theories are those in which only a subset of the logically possible orders can be generated. Powerful theories allow for the generation of all six logically possible orders.

2.1 Restrictive theories

A number of theories that have recently been proposed are designed in a way that they cannot generate the 213 order while allowing for the five other orders.⁸

One very prominent proposal is Barbiers (2005), who derives the five orders by means of VP-movement, starting from a right-branching base order: 123 thus arises in the absence of any movement operations:

(9) 123: [_{VP1} V1 [_{VP2} V2 [_{VP3} V3]]]

132 involves movement of VP3 to SpecVP2:

(10) 132: [_{VP1} V1 [_{VP2} [_{VP3} V3] V2 t_{VP3}]]

321 involves VP3 to SpecVP2 movement followed by movement of VP2 (containing VP3) to SpecVP1:

(11) 321: [_{VP1} [_{VP2} [_{VP3} V3] V2 t_{VP3}] V1 t_{VP2}]

321 involves successive-cyclic movement of VP3; it first moves to SpecVP2 and then moves on to SpecVP1:

(12) 312: [_{VP1} [_{VP3} V3] V1 [_{VP2} t_{VP3} V2 t_{VP3}]]

231, finally, involves movement of VP2 containing VP3 to SpecVP1:

(13) 231: [_{VP1} [_{VP2} V2 [_{VP3} V3]] V1 t_{VP2}]

213 is ruled out under this approach as a matter of principle because VP2 cannot move without pied-piping VP3. Note that it is presupposed that remnant movement of VP2 (following movement of VP3 to some functional head between V1 and V2) is not possible. Barbiers (2008) is a variant of this proposal; it is based on the idea that there is always VP-movement as in the derivation of 321 orders. The various surface orders then arise via different realization options at PF:

(14) [_{VP1} [_{VP2} [_{VP3} V3]₁ V2 [_{VP3} V3]₁]₂ V1 [_{VP2} [_{VP3} V3]₁ V2 [_{VP3} V3]₁]₂]

Abstracting away from the details, what is relevant in the case at hand is that 213 is ruled out by a ban on partial spell-out: this would involve spelling-out parts of the moved VP2 in different copies (i.e. an instance of distributed deletion with V2 spelled out in the top copy and V3 in the bottom copy).

Another type of approach that limits the cluster orders to five out of six and explicitly excludes 213 is based on flexible linearization of sister nodes/VP-inversion and VP3-movement. To my knowledge, it was first proposed in Wurmbrand (2004), who argues that by means of VP-inversion, one can generate the four orders 123, 132, 321 and 231. The fifth one, 312, is argued to involve VP3-movement. Basically the same idea is proposed in Abels (2011, 2015), who embeds his approach in

⁸I will not discuss the approach by Barbiers and Bennis (2010), who argue that in Dutch there are only two cluster orders, viz., 123 and 321. All other orders are reanalyzed as involving a non-verbal V3 so that these are in fact two-verb clusters, some of which may involve Verb Projection Raising. While this is certainly a very radical and thus interesting proposal, I will not discuss it in any detail here because its focus is orthogonal to the question pursued in this paper, viz., the existence/non-existence of the 213 order.

a more general theory of neutral word order (building upon Cinque 2005 and Abels and Neeleman 2012). Under this approach, we obtain the following representations for the five different word orders:

- (15) 123: [_{VP1} V1 [_{VP2} V2 [_{VP3} V3]]]
 132: [_{VP1} V1 [_{VP2} [_{VP3} V3] V2]]
 321: [_{VP1} [_{VP2} [_{VP3} V3] V2] V1]
 231: [_{VP1} [_{VP2} V2 [_{VP3} V3]] V1]
 312: [_{VP1} [_{VP3} V3] V1 [_{VP2} V2 t_{VP3}]]

213 is ruled out in this type of system because it can neither be derived by means of flexible linearization nor by movement since only movement of VP3 (viz. the lexical projection) is possible.⁹

2.2 Powerful theories

There are several theories of verb clusters that can generate all six logically possible orders. In the OT-approach by Schmid and Vogel (2004), the mechanism that generates the six orders is left unspecified so that I cannot provide any representations. In Bader and Schmid (2009), complex heads can be directly generated in different orders based on the formal language CAT:

- (16) 123: [V1+[V2+V3]] 132: [V1+[V3+V2]] 321: [[V3+V2]+V1]
 312: [V3+[V1+V2]] 231: [[V2+V3]+V1] 213: [[V2+V1]+V3]

213 is possible in this theory because the selectional properties of V2 can be inherited by the complex consisting of V1 and V2 so that V3 can be merged as a complement.¹⁰

The hybrid theory proposed in Haider (2003) can arguably also generate all six orders (the author does not address this issue, unfortunately). In his system, descending VR-structures are base-generated as complex heads while structures deviating from the strict 321 order 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 by reprojection of V1:

⁹These restrictive theories also make predictions about larger clusters, of course. Abels (2015), for instance, predicts the existence of only 14 out of 24 logically possible orders in 4-verb clusters. As he demonstrates convincingly, there is a very close match with what is attested in the major cluster types. I will not discuss clusters consisting of four or more verbs in this paper in much detail because the structures become too complex to apply the diagnostics that I will use below to test whether the 213 order constitutes a proper verb cluster. What is crucial in the current context, though, is that 4-verb clusters containing one of the verbs that allow for the 213 order also allow for orders that are predicted to be impossible by the restrictive theories. For instance, in a cluster consisting of Mod-Aux_{perf}-begin/stop/help/learn/see/hear-Inf, the unexpected order 1324 is just as unmarked as the predictably grammatical 1234 order:

- (i) dass er s Buech sött₁ aagfange₃ ha₂ läse₄
 that he the book should.3SG begin.PTCP have.INF read.INF
 'that he should have started to read the book' Swiss German

The unexpected orders 3214, 3241 and 3124 also seem acceptable with this cluster type, though they are certainly more marked. I take this to be sufficient to show that clusters containing this special set of verbs generally allow orders that are not covered by the restrictive theories.

The clusters from the dialects in Steinach and Wasungen with 1324/3241 order discussed in Höhle (2006: 74) and the 2143 clusters from Lindhorster Platt described in Bölsing (2011: 211-217) and discussed in detail in Abels (2015) may constitute further counter-examples.

¹⁰A very similar base-generation account is Sternefeld (2006: 619ff.). Unfortunately, he does not discuss 231 and 213 orders.

- (17) a. $[_{VP} [_{V^o} [_{V^o} [V3] V2] V1]] \rightarrow$
 b. $[_{VP} V1 [_{VP} [_{V^o} [_{V^o} [V3] V2] t_{V1}]]]$

To derive the 123 order, both V2 and V1 have to reproject:

- (18) a. $[_{VP} [_{V^o} [_{V^o} [V3] V2] V1]] \rightarrow$
 b. $[_{VP} V1 [_{VP} V2 [_{VP} [_{V^o} [_{V^o} [V3] t_{V2}] t_{V1}]]]]]$

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

- (19) a. $[_{VP} [_{V^o} [_{V^o} [V3] V2] V1]] \rightarrow$
 b. $[_{VP} V2 [_{VP} [_{V^o} [_{V^o} [V3] t_{V2}] V1]]]$

213 involves excorporation of V1 followed by excorporation of V2:

- (20) a. $[_{VP} [_{V^o} [_{V^o} [V3] V2] V1]] \rightarrow$
 b. $[_{VP} V2 [_{VP} V1 [_{VP} [_{V^o} [_{V^o} [V3] t_{V2}] t_{V1}]]]]]$

312, finally, involves excorporation of V1 followed by excorporation of V3:¹¹

- (21) a. $[_{VP} [_{V^o} [_{V^o} [V3] V2] V1]] \rightarrow$
 b. $[_{VP} V3 [_{VP} V1 [_{VP} [_{V^o} [_{V^o} [t_{V3}] V2] t_{V1}]]]]]$

All six orders can arguably also be generated by the approach by Haegeman and van Riemsdijk (1986), which is based on reanalysis in syntax plus inversion of sister nodes at PF (the authors do not discuss all orders). The base is taken to be a descending 321 order. If V3 is reanalyzed with V2, and the resulting [V3V2]-node is reanalyzed with V1, we obtain four possible orders, depending on which node inversion is applied to, viz. 123, 132, 321 and 231 (as in the approaches by Wurmbrand and Abels discussed above). 312 requires reanalysis and inversion between V2 and V1 (perhaps followed by reanalysis of the complex [V2V1] with V3). 213, finally, requires reanalysis of V2 and V1, followed by reanalysis and inversion of the complex [V2V1] with V3.¹²

Similarly powerful is the remnant movement approach of Koopman and Szabolcsi (2000). For instance, 213 can be generated if a constituent containing V3 is moved to a position between V1 and V2 (SpecFP1 in (22)); thereafter, remnant VP2 is moved to a position above V1 (note that the actual derivations in this work are much more complex; the derivations of the other orders will be similar in spirit to those of Barbiers 2005):

- (22) $[_{VP1} [_{VP2} \mathbf{V2} t_{VP3}] [_{V1'} \mathbf{V1} [_{FP1} \mathbf{VP3} [_{F1'} F1] t_{VP2}]]]]]$

The two types of theories thus only differ with respect to the generation of the 213 order. This implies that the Swiss German construction with 213 is crucial for determining the adequate degree of restrictiveness of cluster theories. Given that this construction allows for 213, one may be tempted to conclude that the powerful theories are on the right track. However, there may be an alternative analysis of the Swiss German construction that I will discuss in the next section.¹³

¹¹Haider (2003: 118) entertains a second possibility to generate the different orders, a process which he calls 'cliticization', but which basically amounts to head-adjunction, either to the left or to the right, again starting from a base-generated complex head in 321 order.

¹²One might try to rule out 213 by reference to cyclicity since the second reanalysis operation targets a more embedded structure rather than applying at the root, but as far as I can tell given the representations in Haegeman and van Riemsdijk (1986), there is no such restriction.

¹³Traditional approaches based on head-adjunction as in Evers (1975) can also generate all six orders, at least under certain assumptions: under a descending base-order, 321 involves no head-movement, 231 involves right-adjunction of V3 to V2, 123 involves right-adjunction of the complex [V2V3] to V1; 312 involves right-adjunction of V2 to V1. 132 requires left-adjunction of V3 to V2, followed by right-adjunction of the complex [V3V2] to V1. 213 requires left-

3 Clusters with 213 order as an instance of the 3rd Construction?

3.1 Another construction with an unmarked 213 order

Interestingly, unmarked 213 orders are also found in the so-called 3rd Construction: this is a coherent construction where the restructuring predicate takes a *zu*-infinitive (rather than a bare infinitive, as in the Swiss German construction) and where part of the non-finite clause occurs in postverbal position (scrambling indicates that we are dealing with a coherent construction):¹⁴

- (23) dass er dem Hans versucht₂ hat₁ t_{demHans} zu helfen₃
 that he the.DAT John try.PTCP have.3SG to help.INF
 ‘that he tried to help John’ 3rdC Standard German

There are a number of important similarities between the two constructions: First, V2 is (more or less) lexical. Second, V2 appears as a participle.¹⁵ Third, both 213 orders are penetrable, i.e. they allow for non-verbal material between V1 and V3 (again, scrambling in (24-b/c) indicates coherence):

- (24) a. wo s mer ghulffe₂ händ₁ s **Gschier** abwäsche₃
 when they me.DAT help.PTCP have.1P the dishes wash.INF
 ‘when they helped me wash the dishes’ Swiss German
Lötscher (1978: 2)
- b. wo s mer s Gschier ghulffe₃ händ₁ **gründlich** t_{s Gschier} abwäsche₃
 when they me the dishes help.PTCP have.3PL thoroughly wash.INF
 ‘when they helped me wash the dishes thoroughly’ Swiss G.
- c. dass er dem Hans versucht₂ hat₁ t_{dem Hans} **die Uhr** zu stehlen₃
 that he the.DAT John tried has the watch to steal.INF
 ‘that he tried to steal John’s watch’ 3rdC Standard German

Given these similarities, it may be possible to unify the Swiss German 213 construction with the 3rd Construction.¹⁶ The Swiss German 213 construction would then not constitute a proper verb cluster so that we could reduce the number of possible cluster orders to 5 out of 6. This would argue in favor of the restrictive verb cluster theories.

3.2 Swiss German 213 orders – verb cluster or 3rd Construction?

How can it be determined whether the Swiss German 213 construction is a proper verb cluster or an instance of the 3rd Construction? Before approaching this question, a refinement is in order: the penetrability argument in (24) is inconclusive because non-verbal material inside the cluster is also a hallmark of Verb Projection Raising, which is readily available in Swiss German (cf. Lötscher

adjunction of V2 to V1 followed by right-adjunction of V3 to the complex [V2V1]. Under a right-branching base as in Zwart (1996), 123 involves no adjunction; 132 requires left-adjunction of V3 to V2; 321 involves left-adjunction of V3 to V2 followed by left-adjunction of the complex [V3V2] to V1. 231 requires right-adjunction of V3 to V2, followed by left-adjunction of the complex [V2V3] to V1; 213 can be derived quite easily, viz., by left-adjunction of V2 to V1; 312 requires right-adjunction of V2 to V1, followed by left-adjunction of V3 to the complex [V1V2].

¹⁴The translational equivalent of the 3rd Construction also exists in Swiss German. However, since it differs from the Standard German construction in some important respects (see 4.5, 5.1.1 below), I will use Standard German for comparison.

¹⁵In the Swiss German 213 construction, V2 sometimes appears in the infinitive, instantiating the IPP-effect. See 9.1 for discussion.

¹⁶Zwart (2007: 80f.) seems to have something along these lines in mind although the passage in the text is not fully clear to me. Louden (2011) assumes without much argument that 213 orders in Pennsylvania Dutch involve extraposition. Cf. also Kroch and Santorini (1991: 321) for a similar idea in a very different framework.

1978). As shown in (25), both VPR and the 3rd Construction are restructuring configurations (cf. scrambling) and occur in (partially) ascending order:

- (25) a. dass er em Hans wett₁ t_{em Hans} d Uhr stäle₂
 that he the.DAT John want.3SG the watch steal.INF
 ‘that he wants to steal John’s watch’ VPR, *SwG*
- b. dass er dem Hans versucht₁ t_{dem Hans} die Uhr zu stehlen₂
 that he the John try.3SG the watch to steal.INF
 ‘that he tries to steal John’s watch’ 3rdC, *StG*

As far as I can tell, there is a near-consensus in the literature (pace Haegeman 1992) that verb clusters and VPR are structurally very similar, see e.g. Haegeman and van Riemsdijk (1986), Broekhuis (1993), den Dikken (1994, 1995, 1996), Zwart (1996). Especially in more recent work the prevailing assumption seems to be that the major difference consists in whether the non-verbal material has to evacuate the lexical VP or not (or whether the lexical VP contains enough structure to host non-verbal material).

What is still an open question, though, is whether VPR and the 3rd Construction should at all be distinguished structurally. Obviously, only if there is such a difference does it make sense to investigate whether the Swiss German 213 orders instantiate a verb cluster or the 3rd Construction.

While some have argued that the two constructions should be distinguished, see e.g. Haegeman and van Riemsdijk (1986), Geilfuß-Wolfgang (1991), others have argued that the basic structure is the same, see e.g. Wyngaerd (1989) and ter Beek (2008). As far as I can tell, the mainly semantic arguments that have been advanced in the literature are eventually inconclusive; although there is a tendency that XPs which are extracted from the lexical VP can reconstruct in VPR but not in the 3rd Construction, there remain counter-examples and the contrasts are very subtle. I will not review these arguments here; the interested reader is instead referred to the appendix in 9.2. The IPP-effect, which has often been argued to be a diagnostic for proper verb clusters, is similarly inconclusive, see the appendix in 9.1 for discussion.

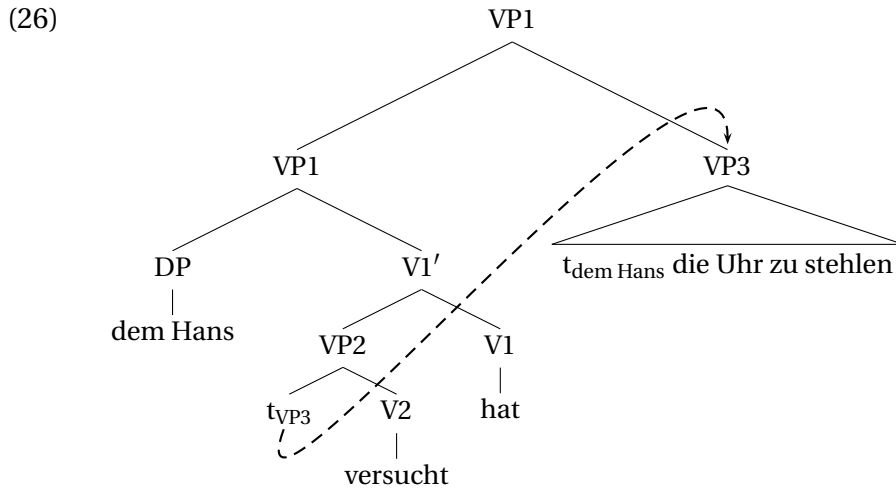
Based on new structural diagnostics, I will argue below that Verb clusters/VPR should indeed be distinguished from the 3rd Construction. Before one can assess the implications of the diagnostic, I will address the analytical options for 213 orders.

3.3 Analytical options for the 213 order: +/- complementation

The possible derivations of 213 orders can lead to two fundamentally different structural relationships within the verb cluster: the lexical VP either ends up as a complement of the functional verb(s) or as an adjunct/non-complement. I will discuss the two options in turn.

3.3.1 The 213 order with non-complementation

One possibility to obtain a 213 order starts out from a descending/OV-base and applies (remnant) extraposition of the lexical VP. This is the classical derivation of the 3rd Construction, see e.g. Broekhuis et al. (1995) for an example like (24-c):



Under a VO/anti-symmetric approach (Kayne 1994), this structure can be obtained in two steps. In an initial step, we derive the 213 order: first, VP3 moves to FP1 between V1 and V2 followed by remnant movement of VP2 to SpecVP1:

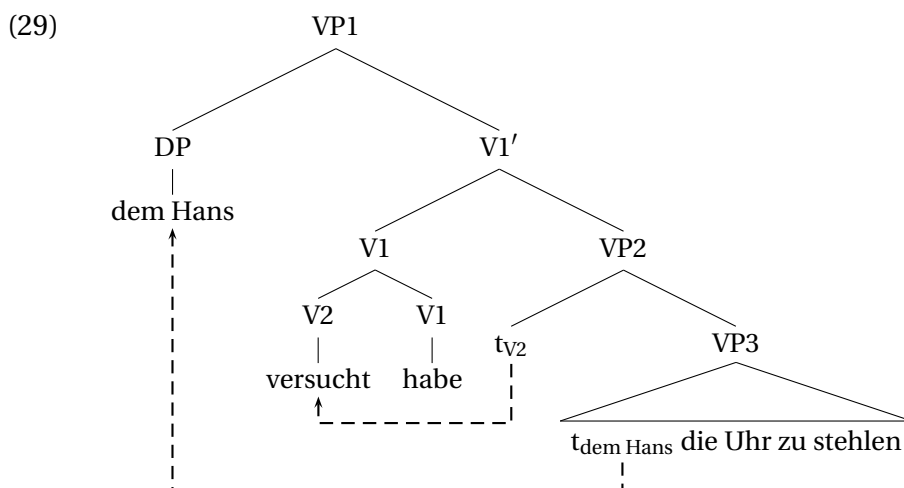
(27) $[_{VP1} [_{VP2} \mathbf{V2} t_{VP3}] [_{V1'} \mathbf{V1} [_{FP1} \mathbf{VP3} [_{F1'} F1 t_{VP2}]]]]$

In a second step, VP3 moves to FP2 above VP1 followed by remnant movement of VP1 to FP3. As a consequence, VP1 no longer contains VP3:

(28) $[_{FP3} [_{VP1} [_{VP2} \mathbf{V2} t_{VP3}] [_{V1'} \mathbf{V1} [_{FP1} t_{VP3} [_{F1'} F1 t_{VP2}]]]] [_{F3'} F3 [_{FP2} \mathbf{VP3} [_{F2'} F2 t_{VP1}]]]]$

3.3.2 The 213 order with complementation

Under a VO/anti-symmetric approach, 213 can be obtained if V2 and V1 invert/form a cluster (e.g. by head-movement, reanalysis or some other operation) while VP3 remains in its base-position, cf. e.g. ter Beek (2008):



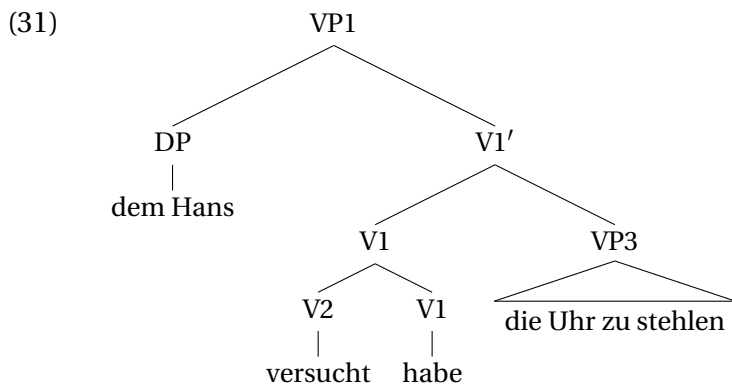
The lexical VP3 is thus a complement of V1/V2.

An alternative derivation is the remnant movement analysis from the previous subsection: VP3 moves to FP1 between V1 and V2 followed by remnant movement of VP2 to SpecVP1, see Koopman and Szabolcsi (2000):

(30) $[_{VP1} [_{VP2} \mathbf{V2} t_{VP3}] [_{V1'} \mathbf{V1} [_{FP1} \mathbf{VP3} [_{F1'} F1 t_{VP2}]]]]$

Under an OV-approach, 213 with complementation can be obtained if e.g. V1 inverts with VP2 after V2 has inverted with V1 by head-movement or reanalysis, if V1 undergoes reprojection followed by reprojection of V2 as in Haider (2003) or if there is reanalysis between V2 and V1 followed by inversion of this complex with V3 as in Haegeman and van Riemsdijk (1986), as discussed above.

In more surface-oriented approaches (e.g. Bayer and Kornfilt 1994, Sternefeld 2006, Bader and Schmid 2009 and HPSG-approaches quite generally), the complementation structure is directly generated:



Importantly, while there are thus various technical means to obtain a 213 order, the resulting structures can be classified according to whether they involve complementation or non-complementation/adjunction. The diagnostics that I am about to introduce will crucially differentiate between these two structural relationships.

4 New diagnostics to distinguish verb clusters/VPR from the 3rd Construction

In this section, I will introduce several diagnostics that systematically differentiate between verb clusters and VPR on the one hand and the 3rd Construction on the other: displaced *zu*, haplology effects with *zu*, relative clause extraposition and stranding of the lexical VP. These diagnostics will show that while verb clusters and VPR involve a complementation structure, the 3rd Construction is characterized by adjunction/non-complementation.

4.1 The phenomenon of displaced *zu*

It is standardly assumed that morphological selection requires a head-complement relationship (but see Wurmbrand 2012 for an approach based on upward Agree). Canonically, the non-finite morphology selected by a verb V_n is realized on the immediately subordinate verb, viz. V_{n+1} (if the non-finite morphology is selected by a noun or a complementizer, the non-finite morphology is realized on the highest verbal element in the relevant domain).

Interestingly, in German varieties, non-finite morphology often does not follow this transparent pattern. Rather, the morphological exponents appear to be displaced. Consider the following triple from Standard German illustrating an Aux-Mod-Inf-cluster with three different serializations (the complementizer *ohne* selects a *zu*-infinitive):

- (32)
- | | | | | | | |
|----|------|-----------------------------|---|--|---|-----|
| a. | ohne | das Buch lesen ₃ | gekonnt ₂ zu haben ₁ | without the book read.INF can.PTCP to have.INF | ‘without having been able to read the book’ | 321 |
| b. | ohne | das Buch haben ₁ | lesen ₃ zu können ₂ | without the book have.INF read.INF to can.INF | ‘without having been able to read the book’ | 132 |

- c. ohne das Buch lesen₃ haben₁ **zu** können₂
 without the book read.INF have.INF to can.INF
 ‘without having been able to read the book’ 312

In (32-a), which involves a 321-order, the *zu*-infinitive appears on the hierarchically highest verb of the cluster, viz. V1, as we would expect given the standard assumptions about morphological selection. In (32-b/c), however, which involve a 132 and 312 order, respectively, *zu* does not occur on V1 but rather on V2. It thus seems to be displaced. As an initial generalization, *zu* always attaches to the last verb of the complement of the *zu*-selector. In a configuration as in (32) with the *zu*-selector outside the cluster, displacement occurs once V1 is not cluster-final (i.e. in all orders except 321 and 231). Crucially, if *zu* occurs on V1 in (32-b/c), the result is sharply ungrammatical, as (33) shows for (32-b):¹⁷

- (33) *ohne das Buch **zu** haben₁ lesen₃ können₂
 without the book to have.INF read.INF can.INF
 ‘without having been able to read the book’ 132

The grammatical status of displaced *zu* in the standard language has been controversial ever since Grimm (1837:949). This is related to two facts: first, it appears to violate the canonical rule of morphological selection. Second, subjects show a significant degree of uncertainty and variability in empirical tests according to Reis (1979) and Haider (2011).

While Merkes (1895) and Haider (2011) consider displacement ungrammatical, Meurers (2000) and Vogel (2009) provide arguments in favor of its grammatical status. I will not review this discussion here as I have done so elsewhere, see Author (xxxx). Furthermore, the status of displacement in the standard language is orthogonal to the current discussion as the argument will be based on Swiss German data. I will instead briefly summarize my arguments in favor of the grammatical status of displacement even in the standard language. First, treating displaced *zu* as ungrammatical fails to account for the significant contrast between (33) and (32-b). Second, once the placement rule for non-finite morphology is reconsidered, no rule of grammar is violated anymore, see 4.3.2 below. In addition to these more theory-internal arguments there are also empirical arguments in favor of treating displaced *zu* as grammatical: first, it is attested in careful sources. Second, displaced *zu* is well-attested in German dialects without there being any indication that the construction should be marked/unacceptable (see section 4.2). Third, displaced *zu* is part of a larger phenomenon of displacement: in earlier stages of German we find displacement of participial forms, and in certain East Middle German varieties there is displacement of a wealth of non-finite forms including gerunds and special infinitives. Consider the following examples((34-a) is from Fleischer and Schallert 2011: 185, (34-b) from Höhle 2006: 68):

- (34) a. dez han₁ wir unser kunichlich Insigel an disen breiff haissen₂ gehenket₃
 therefore have.1PL we our royal seal to this letter let.INF attach.PTCP
 ‘Therefore we had our royal seal attached to this letter.’
Middle High German, 1286
- b. kâsd₁ mə heləf₂ gəschri:₃
 can.2SG me.DAT help.INF GE.write.INF
 ‘Can you help me write?’
dialect of Kleinschmalkalden

In (34-a), the participle selected by V1 is realized on V3 rather than on V2, which appears in the bare infinitive, thereby instantiating the Participio pro Infinitivo (PPI) construction. In (34-b), V1 selects

¹⁷Interestingly, the version in (33) represents the grammatical Dutch pattern; the corresponding particle *te* always occurs in the expected place, displacement is not a possibility. This implies that *te* reaches its surface position in a different way than proposed below for German. See Author (xxxx) for discussion.

a a so-called *ge*-infinitive, viz., an infinitive with a *ge*-prefix. However, the *ge*-infinitive occurs on V3 while V2 (which selects a bare infinitive) occurs in the infinitive. Displacement of these forms follows the same pattern as displaced *zu*: the exponents attach to the last verb of the complement of the selector of the non-finite morphology; this leads to displacement once the verb immediately dependent on the *zu*-selector is not last in the complement (which is regularly the case in these varieties as they predominantly use ascending orders).

4.2 Displaced *zu* in German dialects

Displaced *zu* is more prominent in German dialects because they often have (strictly) ascending orders. Interestingly, in all of the literature, displaced *zu* is described as the canonical realization of non-finite morphology in (partially) ascending orders; there are no comments suggesting that it may be a marked or deviant phenomenon. The grammaticality of displacement is thus undisputed. This consensus comprises traditional grammars (Hodler 1969:560, Weber 1987:244 and especially the works cited in Höhle 2006), more descriptive treatments (Comrie and Frauenfelder 1992) as well as formal approaches (Bader 1995:22 and Cooper 1995:188f.). Furthermore, displaced *zu* can be heard on the radio (Cooper 1995) and be found on the internet. The following examples are a small selection. (35-a) is from Weber (1987:244,fn.1), (35-c) is from Comrie and Frauenfelder (1992:1059), and (35-d) is from Weise (1900:154):

- (35) a. Er schiint₁ nüüt wele₂ z wüsse₃ dervoo.
 He seem.3SG nothing want.INF to know.INF about.it
 ‘He does not seem to be interested in it.’ 1 ... 23 Zurich G.
- b. Ich liebe d freiheit, selber de tag chöne₁ z bestimme₂.
 I love.1SG the freedom self the day can.INF to determine.INF
 ‘I love the freedom to determine my schedule.’ 12 Swiss G.
 cf. <http://badoo.com/de-ch/0279246484/>, accessed March 11, 2013
- c. Ech ha ts Büach kchöiff, fer dam Marco cheni₁ z sägan₂, ...
 I have.1SG the book buy.PTCP for the.DAT Marco can.INF to say.GER
 ‘I bought the book to be able to tell Marco ...’ 12 Bosco Gurin
- d. weil er sich nicht von ihm braucht₁ lassen₂ anzuschnauzen₃
 because he self not by him need.3SG let.INF rant.at.INF 123
 ‘because he does not need to be ranted at by him’ Altenburg

A fact that will play an important role below is that displacement is also attested with Verb Projection Raising:

- (36) ohni mi welle₁ uf d bullesite z stelle₂, im gegeteil
 without me want.INF on the cops.side to put.INF on.the contrary
 ‘without wanting to side with the cops, on the contrary, but ...’ 1X2 Swiss German
<http://www.fcbforum.ch/forum/showthread.php?4328-usschritige-nachem-spiel-!/page4>; accessed March 11, 2013

The empirical situation can thus be summarized as follows: *z(u)* always attaches to the last verb of the complement of the *zu*-selector; this will lead to displacement if in that complement the hierarchically highest verb does not occur last (as will be regularly the case in ascending orders).¹⁸ Displacement as in (35-d) can be schematically illustrated as in (37) (displacement can, of course, also originate on a complementizer rather than V1 as in (32) above). Under displacement, the verb selected by the *zu*-selector appears in a default form, usually in the infinitive; the selectional requirements of V2 are suppressed or alternatively are satisfied by the displaced morphology, see

¹⁸For very rare cases where *zu* is displaced to the left, see Schallert (2012:252).

fn. 24 and Author (xxxx) for discussion of these issues:¹⁹



4.3 The derivation of displaced *zu*

I will now propose an analysis of displaced *zu* that is based on two essentially independent components: a theory of verb cluster formation where complex heads arise at PF and a post-syntactic approach to the placement of non-finite morphology. Importantly, the facts can probably also be derived with other verb cluster theories and different approaches to displacement, although not all combinations will work. I will briefly come back to alternatives in section 6 below. I will first introduce the presupposed theory of verb cluster formation before combining it with displacement.

4.3.1 Verb cluster formation at PF

In Author (xxxx) I have argued in favor a new theory of verb cluster formation. Like the classical verb raising approach by Evers (1975) and later implementations, it involves the formation of complex heads. It crucially differs from previous approaches, though, in that the complex heads arise after syntax, viz. at PF, via Local Dislocation, i.e. through affixation and reordering under adjacency (cf. Embick and Noyer 2001).

One central component of this approach is that the order of the verbal elements in the cluster is determined differently than the order of heads, specifiers and complements. For the latter, I assume that while specifiers are systematically ordered before their heads, complements can be linearized both before or after the head, i.e. I do not adopt a rigid anti-symmetric system à la Kayne (1994) but a more flexible system like that of Abels and Neeleman (2012) that incorporates ordering statements for all sisterhood relations. Furthermore, the ordering statements can refer to properties of the head and the non-head. This allows the complements of a certain head to be linearized differently depending on properties of the non-head. Concretely, I assume that non-verbal complements are linearized to the left of the verb while verbal complements are linearized to the right. This produces a mixed system similar to that in Cooper (1995), Barbiers (2000) and Schmid and Vogel (2004). The consequence for cluster orders is that if nothing else happens, a strictly ascending 123 order obtains.²⁰

Orders that deviate from 123 are derived by means of a late PF-operation, viz., Local Dislocation, which involves rebracketing and (in the standard case) inversion of the two heads; it applies

¹⁹Displaced morphology in German has to be distinguished from so-called parasitic morphology in Norwegian/Swedish and Frisian, which is essentially a Participio pro Infinitivo construction, see Wurmbrand (2012: 132):

- (i) Ik ben tankber dat ik sa folle dien₃ kinnen₂ haw₁.
I am thankful that I so much do.PTCP can.PTCP have
'I am grateful that I could do so much.'

Frisian

Unlike displaced morphology in German, parasitic morphology is restricted to participles and there are no default forms on V2; rather, the participle appears twice. Furthermore, parasitic morphology allows for multiple displacement, i.e. spreading of participial forms onto several dependent verbs. Frisian additionally differs from German in that parasitic morphology only occurs in left-branching/descending clusters (Frisian only has right-branching/ascending clusters in the 3rd Construction); finally, Frisian also has upward displacement (the requirements of V3 are realized on V2).

²⁰For the purposes of this paper it does not matter whether linearization of specifier, head and complement takes place at Merge or post-syntactically. Given that this ordering requires reference to hierarchical structure, it will invariably take place before the reordering in the verbal complex, which applies to linear structure and thus at a later point of the derivation given a PF-architecture as in Embick and Noyer (2001).

to linear structure and is constrained by adjacency. This cluster-forming mechanism can derive all six logically possible orders. Next to 123, it can derive 132, which involves complex head formation between V2 and V3, leading to 1[32], 321 (complex head formation V3+V2 followed by complex head formation between V1 and [V3+V2]), viz., [[32]1], 213 (complex head formation between V1 and V2), viz. [21]3. 312 and 231 involve string-vacuous cluster formation, viz., rebracketing without inversion: In 312, there is first rebracketing between V1 and V2 followed by complex head formation + inversion with V3: [3[12]]. In 231, finally, V2 and V3 rebracket before the newly formed complex undergoes complex head formation + inversion with V1: [[23]1].²¹

While the grammar in principle always allows the generation of all six orders, the possible orders in a given variety can be determined by linearization parameters (a point I come back to in 7 below). Finally, the coherence/restructuring effects we find with verb clusters are due to the fact that the relevant verbal projections contain less structure, viz. lack a CP- (and possibly a TP-) layer, cf. e.g. Wurmbrand (2007).²² In what follows, I will label all verbal projections as VPs for ease of readability even though some may better be classified as functional; furthermore, I will also omit the vP-layer.

In Author (xxxx) I have presented two major arguments in favor of the PF-perspective which for reasons of space I will only summarize very briefly here: first, the approach makes cross-linguistically correct predictions about which orders are penetrable (i.e. allow for VPR) in which positions. For instance, while 132 orders allow for non-verbal material between V1 and V3, 312 orders are completely impenetrable because they consist of a complex head (a fact that is more difficult to capture if for instance cluster reordering involves VP-movement). Second, cluster formation at PF solves the so-called cluster puzzle: clause-final verbs in descending order form an impenetrable unit, which follows from their forming a complex head, see (38-a). Consequently, extraposition from the lexical VP has to target the projection of VP1. However, in verb-second structures, which are derived from the verb-final structure, parts of the cluster can suddenly be moved, viz. entire VPs with extraposed elements that would not be well-formed clause-finally, and the finite V1 itself, see (38-b):

- (38) a. dass man *t_{darüber}* reden₃ *darüber können₂ *darüber sollte₁ ✓darüber
 that one talk.INF about it can.INF about it should.3SG about it
 ‘that one should be able to talk about it’ *Standard German*
- b. [_{VP} Reden₃ darüber] sollte₁ man schon *t_{VP}* können₂ *t_{V1}*.
 talk.INF about it should.3SG one indeed can.INF
 ‘One should indeed be able to talk about it.’ *Standard German*

This is puzzling if complex head formation takes place in syntax – one seems to be dealing with a movement paradox and is forced to adopt excorporation (but only in the derivation of verb second structures). A syntactic account to complex head formation is also at pains to explain why verb second movement never involves a complex head but just the hierarchically highest element of the cluster. These problems disappear once cluster formation takes place post-syntactically: it comes too late to block movement in verb-second structures and to allow fronting of the entire complex head to C while it correctly rules out extraposition to VP3 or VP2 in (38-a) as this would block cluster formation and thus the descending order.

²¹This mechanism can also generate most of the logically possible orders in 4-verb clusters, except for 3142 and 2413. What is crucial in the present context is that unlike the restrictive theories discussed in 2.1 above it can generate the orders 1324, 3214, 3241, and 3124, which are acceptable in Swiss German clusters with V3 belonging to the class of predicates that allow for 213 (recall fn. 9).

²²Alternative mechanisms to derive coherent structures like TP-movement and incorporation as in Grewendorf and Sabel 1994 would work as far as I can tell.

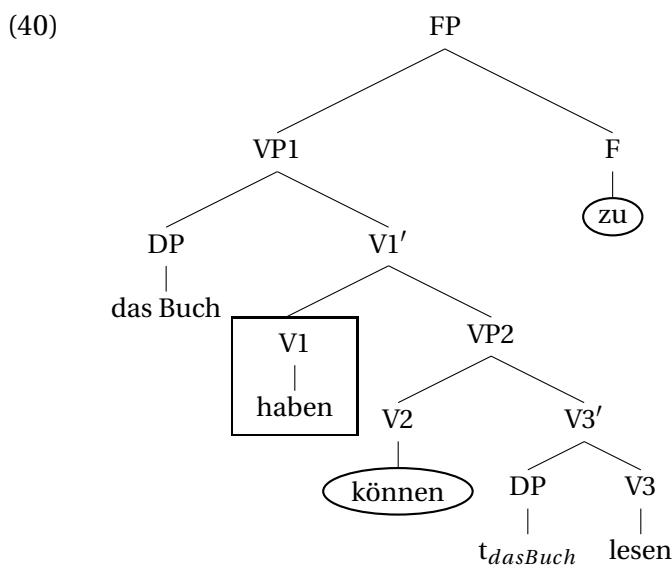
4.3.2 The placement of non-finite morphology

The basic idea underlying *zu*-placement is very simple: the non-finite morphology originates in independent syntactic heads and is associated with its host post-syntactically by means of Local Dislocation. The vocabulary items are inserted into functional heads that ordered after their complement, in accordance with the head-final character of the language. As a consequence, the non-finite morphology always comes last in the complement of the *zu*-selector. This derives the generalization that *zu* always affixes onto the last verb of the complement of the *zu*-selector. The mechanism that associates the morphology with its host is thus always the same, but since Local Dislocation applies to linear structure, it can have very different effects, depending on the order in the verb cluster: if the order is strictly descending, the morphology appears to be well-behaved. If, however, V1 is not final in the cluster, *zu* will appear to be displaced. Crucially, however, there is thus no displacement operation as such; rather, displacement is only a side-effect.

Concretely, I assume that there is a separate functional head F that hosts the features corresponding to *zu* (cf. also Den Dikken and Hoekstra 1997: 1062). This head occurs above VP. Morphological selection is thus checked in syntax: A V1 that takes a *zu*-infinitive is syntactically combined with an FP hosting the relevant syntactic features (given a post-syntactic approach to morphology, cf. Halle and Marantz 1993, the morphological exponents are inserted late). This functional head has another important property: it takes its VP-complement to the left. As a consequence, the non-finite morphology always comes last in the complement of the *zu*-selector. I will now go through the derivations for both well-behaved 321 cases like (32-a) as well as examples with displacement such as (32-b) and (32-c). I repeat the first two for convenience:

- (39) ohne das Buch {lesen₃ gekonnt₂ **zu** haben₁ | haben₁ lesen₃ **zu** können₂ }
 without the book read.INF can.PTCP to have.INF have.INF read.INF to can.INF
 ‘without having been able to read the book’ 321/132 St. German

As discussed above, given that verbs take their verbal complements to the right, we obtain a right-branching cluster with strictly ascending 123 order; furthermore the functional head F or ordered after the verb cluster, see (40). Note that the hierarchically highest verb is boxed while the verb that *zu* will end up on is circled.²³



²³Non-verbal material has been scrambled out of the lexical VP; this is optional in varieties that allow Verb Projection Raising, see e.g. Broekhuis (1993). For a base-generation alternative to scrambling, see e.g. Fanselow (2001) and Author (xxxx).

At vocabulary insertion, the hierarchical structure is converted into a linear structure. Now verb cluster formation comes into play. To derive the examples in (39), we generate 1[32] and [[32]1]. *Zu*-placement is next: Since by assumption *zu* takes its VP-complement to the left, it is linearized after the verb cluster. As it is a dependent element with selectional properties, it needs a host; furthermore, it is specified to attach to the left of the verb. By Local Dislocation, it is affixed onto and inverted with the closest, i.e. linearly adjacent verbal element:²⁴

- (41) a. 1[32] *zu* ⇒ 1[3+*zu*+2]
↑ LD ↓
 b. [[32]1] *zu* ⇒ [[32]+ *zu*+1]
↑ LD ↓

Note that the order of operations follows from cyclicity, given that the PF-derivation proceeds inside-out as proposed in Embick and Noyer (2001: 576, 580).²⁵ Displacement with 312 orders as in (32-c) proceeds similarly: first, a 312 order is formed: [3[12]]. Then, *zu* is affixed onto and inverted with V2:²⁶

- (42) [3[12]] *zu* ⇒ [3[1+*zu*+2]]
↑ LD ↓

Finally, in a variety that allows a strictly ascending 12(3) order as in (35) and (36), things are particularly simple: the 12(3) order arises through the ordering between heads and complements. *Z(u)* is linearized after the verb cluster and then targets the last verb of the cluster:

- (43) 1 2 3 *zu* ⇒ 1 2 *zu*+3
↑ LD ↓

²⁴*Zu* can be shown to be a proper affix as it has selectional properties: it is only compatible with verbs in the bare infinitive. Given its flexible positioning, it has been referred to as a phrasal affix, see e.g. Vogel (2009), Hinterhölzl (2009). In the trees, the non-finite verbs all appear in the infinitive. This is a slight simplification: verbs like V1 that do not receive any functional morphemes due to displacement, are assigned the infinitive feature by default before vocabulary insertion; verbs such as e.g. V3 in the example under discussion that are governed by an infinitive-selecting verb, however, may receive the infinitive morphology via a separate functional head between V2 and V3. Given cyclicity, this head will attach before *zu*. In Author (xxxx), where I discuss the morphological aspects of displacement in detail, I argue instead that the infinitive is never present syntactically; verbs like V3 appear in the infinitive because the vocabulary item for *zu* has a feature that triggers the insertion of the infinitive exponent, cf. Halle and Marantz (1993) for similar cases.

Since for reasons of space I cannot discuss the IPP-effect, I omit the functional projection for the participle in (40).

²⁵Although the full hierarchical structure is no longer available at this point, it is assumed in this model of the PF-architecture that there are still different domains/constituents that determine cyclicity.

²⁶In all the derivations so far, *zu* is inverted with the last verbal terminal of the cluster. This might seem to be in conflict with the generalizations in Embick and Noyer (2001: 577f.) about what type of element can adjoin where. They distinguish between Morphological Words (MWd), which refers to independent heads as well as complex heads, and Subwords (SWd), which refers to terminal nodes of complex heads. Crucially, they argue that MWds can only adjoin to MWds while SWds can only adjoin to SWds. In the case at hand, however, *zu*, an independent head and thus an MWd, adjoins not to the entire complex head but rather only to a segment of the complex head, viz. V2. One possibility to avoid a conflict is to assume a default rule that adjoins stray affixes to the verbal complex so that they become subwords and can subsequently rebracket and invert with one of the cluster's segments. Alternatively, it may be possible to derive the placement possibilities of elements like *zu* from their selectional properties (see also Embick and Noyer 2001: 580): unlike clitics such as the Latin coordinator *-que* 'and', which is category-insensitive, the vocabulary items for non-finite verbal features not only select a category but also specific versions of the category, e.g. a stem or a bare infinitive. This automatically precludes affixation to more complex elements.

To avoid infixation of *zu* into prefix verbs, I will assume that they are combined in syntax (or even before that) and that their internal structure is no longer accessible at PF (cf. *zu verlassen* 'to leave' vs. **ver-zu-lassen*). Particle verbs like *anfangen* 'begin', on the other hand, do not form a complex in syntax (as suggested by the fact that they are separated under verb second). Consequently, affixation of *zu* between the particle and the verb stem is unproblematic: *an-zu-fangen* 'to begin'. See Author (xxxx) for more discussion of the morphological aspects.

Note that since F is always linearized at the end of the complement of the *zu*-selector, *zu* will attach to the last verbal element of the cluster in strictly ascending orders. This holds both if F occurs outside the verb cluster as in the examples just discussed and if F is selected by V1 as in (35-a), (35-d), (34-a), (34-b).

To summarize, *zu* is inserted into a functional head F that is linearized after its VP-complement. *Zu* is placed by Local Dislocation, an operation that is sensitive to linear precedence and adjacency. Consequently, *zu* is always associated with the last verb of the complement of the *zu*-selector. *Zu* appears well-behaved in (3)21 orders like (32-a) because in these orders, the hierarchically highest occurs cluster-final. Displacement, on the other hand, is just a side-effect of cluster orders where V1 is not cluster-final; there is no displacement rule as such. Rather, there is just a single mechanism that associates the non-finite morphology with its host. Put more generally, displacement arises from a conflict between the general head-finality of the language (as expressed by the head-final linearization of the functional heads and their VP-complements) and (partially) ascending verb clusters. The facts thus all fall out from independently motivated principles: the head-finality of the language motivates the head-final linearization of functional heads. The selectional properties of the vocabulary item *zu*, i.e. its prefixal nature, determine its exact position. An explicit rule for the placement of non-finite morphology is thus not necessary. Finally, the various cluster order possibilities are simply independent properties of a given variety.

One of the major advantages of the PF-approach is that no problems arise for semantic interpretation. This is particularly crucial in the case of participle displacement as the participle arguably contributes to the interpretation of the perfect. In the PPI-construction (34-a), it must consequently not be interpreted on V3 but rather on V2. Since the morphology is inserted and placed post-syntactically, no problems arise in the present approach: at LF, which interprets the hierarchical structure produced by narrow syntax, the relevant features are located in the correct position. Concretely, in the PPI-case, there would be an FP above VP2 for the participle so that the interpretative contribution by the participle correctly applies to VP2 and not to VP3.

My post-syntactic approach is inspired by a number of derivational precursors such as von Stechow (1990: 156) and Sternefeld (1990: 251) and especially Hinterhölzl (2009: 208). Since I have pointed out the advantages of my implementation in detail in Author (xxxx), I refrain from doing so here. Non-derivational alternatives will be briefly addressed in section 6 below.

4.4 Displacement as a diagnostic for complementation

Having established the mechanism underlying *zu*-placement, I will now show that it can be used as a diagnostic for complementation, which in turn will help distinguish between Verb clusters/VPR and the 3rd Construction.

4.4.1 VPR: displacement obligatory

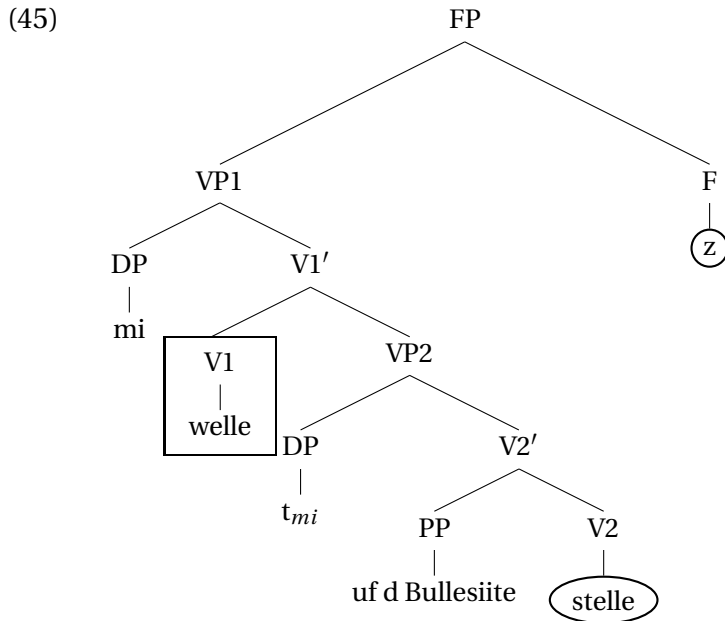
Recall that in both verb clusters and Verb Projection Raising, *zu* occurs on the last verb of the complement of the *zu*-selector. Consequently, *zu* appears displaced if V1 is not the last verb of the complement. I repeat the VPR-example with displacement (36) for convenience:

- (44) ohni mi (*z) welle₁ t_{mi} uf d bullesite z stelle₂
 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>; accessed March 11, 2013

Since the example involves a strictly ascending cluster, its derivation is straightforward. The starting point is a right-branching VP-structure with F being ordered after the VP:



After linearization and vocabulary insertion, *z*-placement applies next and inverts *z* with the final verb of the cluster:

(46) $V_1 PP V_2 z \Rightarrow V_1 PP \underset{\uparrow LD \downarrow}{z+V_2}$

As in the Standard German examples in (32), the lexical VP is a structural complement. Consequently, displacement can be seen as a diagnostic for complementation:²⁷

(47) *zu*-displacement → complementation

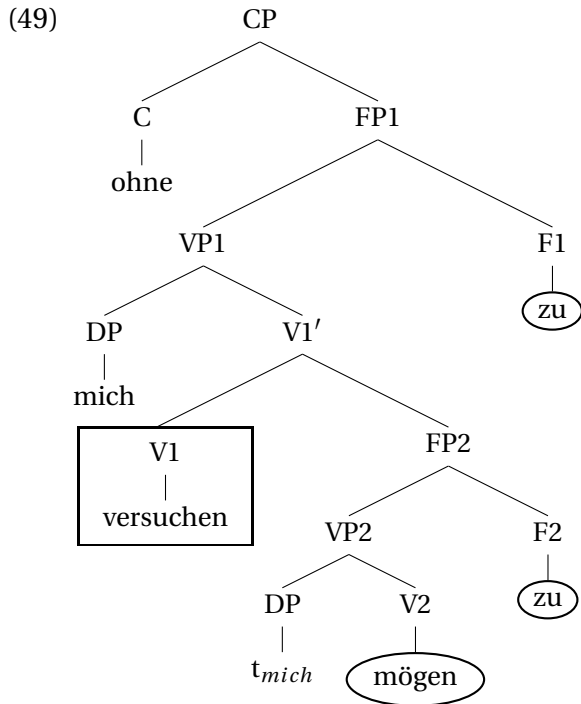
4.4.2 3rd Construction: no displacement

The situation is very different in the 3rd Construction: there is no displacement. Rather, *zu* appears on V1 (since V1 also selects a *zu*-infinitive, there is another *zu* on V2; pronoun fronting makes sure that we are dealing with a 3rd Construction and not a non-restructuring CP-complement):

(48) ohne mich *(**zu**) versuchen₁ t_{mich} zu mögen₂
 without me to try.INF to like.INF
 ‘without trying to like me’ Standard German

Given that verbs take verbal complements to the right, *zu*-placement would arguably apply to an ascending structure as in (49):

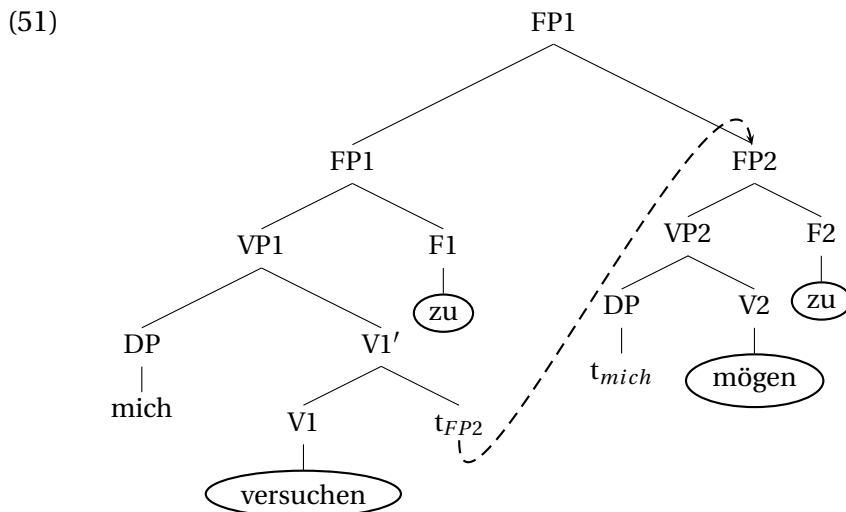
²⁷Note that displacement provides an argument against the extraposition analysis of VPR as in Wyngaerd (1989), Besten and Broekhuis (1992), Haegeman (1992) because extraposition would bleed displacement, see the next subsection.



Since both *zus* are linearized after the verb cluster, they would both attach to V2, leading to the wrong result (perhaps they might be reduced to one *zu* by haplology, see section 4.5 below):



The correct placement facts obtain instead if the 3rd Construction involves (remnant) extraposition; in (51), the lexical VP together with its functional structure is right-adjoined to FP1.²⁸



²⁸Extraposition as rightward movement may seem alien in a system with ascending verb clusters (but see Haegeman 1998: 294 for the same assumption), but since I do not adopt a strict antisymmetric system, nothing should rule this out in principle. I mainly use rightward movement for ease of exposition; leftward movement of FP2 followed by remnant movement of FP1 as e.g. in Hinterhölzl (2006) would work just as well for present purposes. The coordination analysis of extraposition proposed in de Vries (2002: chapter 7) may work as well as long as FP is not outside the coordination hosting antecedent and extraposed clause. Deletion of *zu* in the second conjunct hosting the extraposed clause will prevent a clash in the morphology (*zu* would otherwise attempt to attach to the finite complement clause). I have no new insights to offer as to what triggers extraposition.

After linearization, both *zu*s are adjacent to different verbs and Local Dislocation derives the correct result:

- (52) DP **zu**₁+V₁ **zu**₂+V₂
 ↑ LD ↓ ↑ LD ↓

Unlike in the Standard German verb cluster examples in (32) and in VPR as in (36), the lexical VP is not a structural complement in the 3rd Construction. Consequently, the lack of displacement can be interpreted as a diagnostic for adjunction/non-complementation:

- (53) no displacement → adjunction/non-complementation

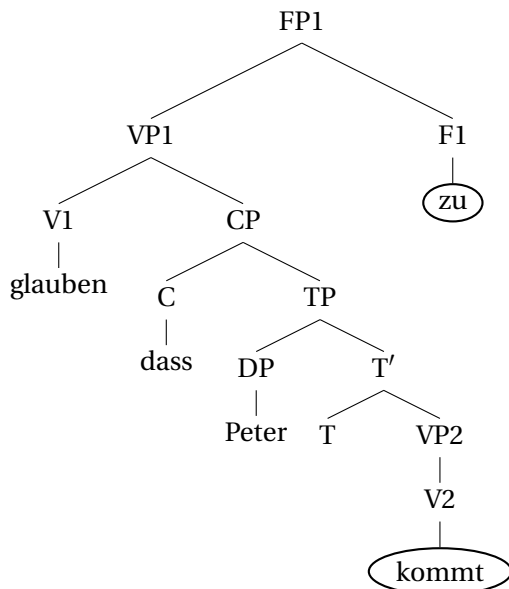
4.4.3 Sentential object CPs: no displacement

Finite CP-complements (as well as non-finite non-restructuring CP-complements) pattern with the 3rd Construction, i.e. there is no displacement, as the following contrast shows:

- (54) a. *ohne glauben, [CP dass Peter **zu** kommt]
 without believe.INF that Peter to come.3SG
 ‘without believing that Peter will come’ Standard German
- b. ohne **zu** glauben, [CP dass Peter kommt]
 without to believe.INF that Peter come.3SG
 ‘without believing that Peter will come’ Standard German

(54-a) results if the CP remains a complement of the matrix verb as in (55):

- (55) CP-object = complement:

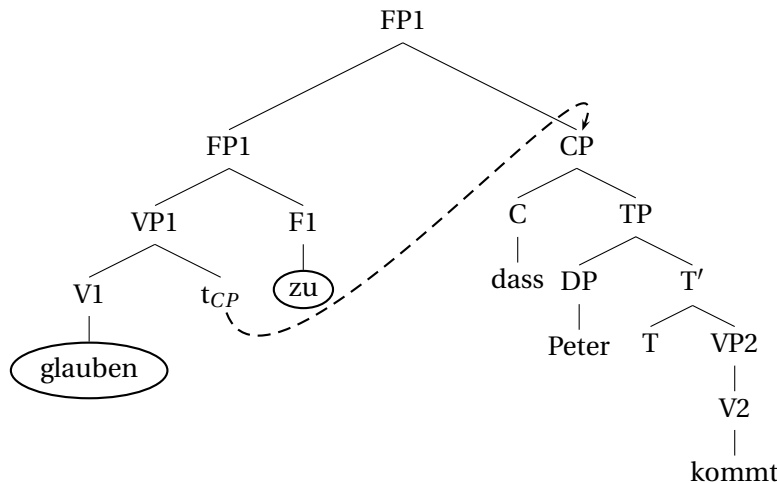


After linearization, *zu* would follow the finite verb *kommt* ‘comes’. Local Dislocation will fail because the selectional properties of *zu* are not respected: it selects a bare infinitive and therefore cannot attach to the finite verb. This derivation thus crashes at PF:

- (56) *V₁ C DP **zu**+V₂
 ↑ ↓

The correct result obtains if the CP is extraposed to FP1 as in (57):

(57) CP-object = extraposed



Zu is now linearized after the matrix verb and can thus felicitously undergo Local Dislocation; the finite CP does not interfere anymore:

(58) **zu**+V₁ C DP V₂

Note that the lack of displacement shows that postverbal finite (and non-finite) object CPs are not sisters of V at surface structure (pace Zwart 1993, Haider 1993, 2010).²⁹

To summarize this section so far, I have established that displacement is a diagnostic for complementation, i.e. the lexical VP is a structural complement. Conversely, absence of displacement is an indication of adjunction/non-complementation. The asymmetry w.r.t. displacement between verb clusters/VPR and the 3rd Construction shows that they involve fundamentally different structural relations. In the rest of this section I will discuss further diagnostics for complementation that all converge on the same result: verb clusters/VPR bear the hallmarks of complementation while the 3rd Construction involves adjunction/non-complementation.

4.5 Another diagnostic for complementation: missing *z* in Swiss German

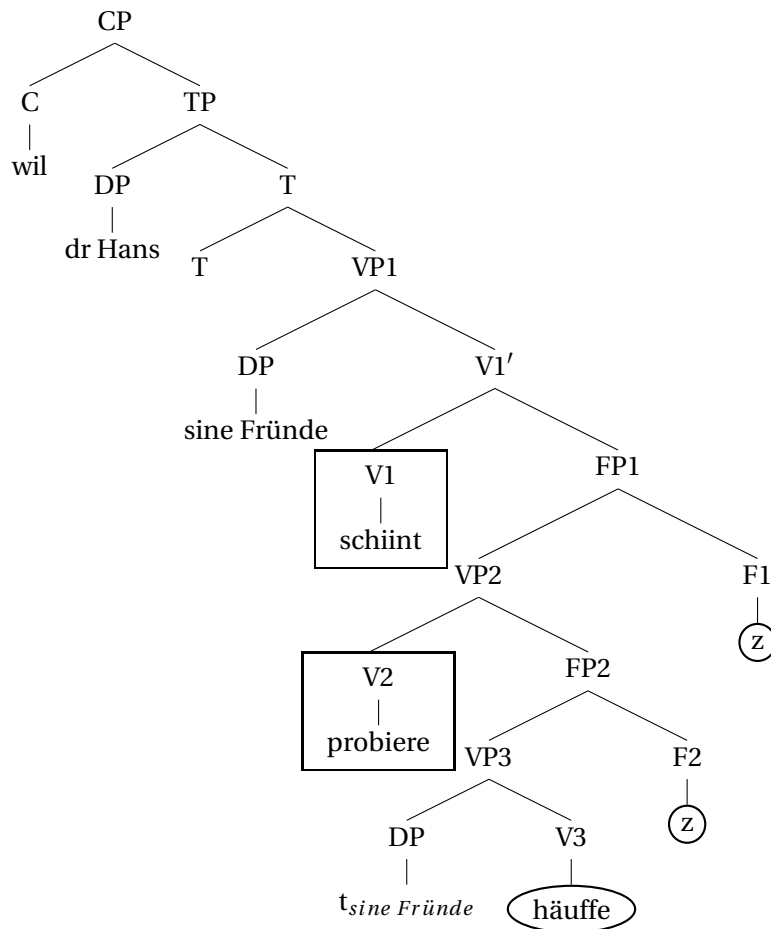
The so-called missing-*z* construction in Swiss German constitutes another diagnostic for complementation. In this construction, there are 2 *z(u)*-selectors and the verbs in the cluster appear in ascending order. Interestingly, there is only one *z*, namely on the last verb of the verb cluster (while the non-final verb appears in the bare infinitive). In the following example from Bernese German (Bader 1995: 22,26), V1 and V2 select a *z*-infinitive, but there is only one *z*, on V2:

(59) wüu dr Hans sine Fründe schiint_{1[zu]} probiere_{2[zu]} **z** häuffe₃
 because the John his.DAT friends seem.3SG try.INF to help.INF
 'because John seems to try to help his friends' *Bernese German*

The missing-*z* construction can be derived as follows: given that verbal complements are ordered to the right of the governing verb, we obtain an ascending VP. Furthermore, by assumption, the functional heads F1/F2 are ordered after their VP-complement:

²⁹Whether finite CP-complements can also be linearized to the left of the matrix verb is a controversial issue. While CPs can indeed occur to the left of the matrix verb in the middle field, this may also constitute a scrambled position. The same goes for intraposed non-finite non-restructuring CPs, see Bayer et al. (2005) for discussion.

(60) Missing *z*:



After linearization, both *zs* follow the verb cluster. I propose that the two *zs* are reduced to one by haplology (which can be understood as deletion under identity). Consequently, only one *z* is affixed onto the last verb of the cluster:

- (61) a. $V_1 V_2 V_3 z z$ both *zs* adjacent \rightarrow haplology: $V_1 V_2 V_3 z z$
 b. $V_1 V_2 z+V_3$ Local Dislocation
 $\quad \uparrow \text{LD} \downarrow$

Note that the haplology effect only obtains if FP2 remains a complement of V1. If it were extraposed, we would expect the same pattern as in the Standard German 3rd Construction. In fact, as far as I can tell, missing *z* is optional, i.e. many speakers also accept a version with two *zs*. This suggests that speakers optionally have access to extraposition of FP2 (to a position above FP1).³⁰

³⁰The missing-*z* construction is difficult to elicit although it can be found on the Internet:

- (i) ... au ohni probiere₁ z wohrsagere₂
 also without try.INF to prophesy.INF
 'without trying to prophesy'

<http://thats-me.ch/forum/em-gewinner/20/31>, accessed March 28, 2013

According to Cooper (1995: 188f.), missing *z* is limited to Verb Raising cases and is blocked in Verb Projection Raising. However, this claim could not be verified in an informal survey. Furthermore, a quick google search delivers two counter-examples, see (ii):

- (ii) a. ... ohni öpe jeh mau säuber probiere₁, Dütsch z rede₂
 without PRT ever once self try.INF German to speak.INF
 'without ever trying to speak German oneself'

But what is crucial for what follows is that missing *z* is another diagnostic for complementation:

(62) missing *z* → complementation

Verb sequences with missing-*z* thus represent proper verb clusters.

Missing *z* is not just a quirk of Swiss German. Similar haplology effects can be found in other West-Germanic varieties as well:

In Frisian, we find a missing *te*-construction: while the cluster order is normally strictly descending in this language, the 3rd construction allows for (partially) ascending 312 orders. In (63), both the complementizer *om* and V1 *hoeve* ‘need’ select a *te*-infinitive. Interestingly, we only find one *te*, crucially before the cluster-final verb V2, see Den Dikken and Hoekstra (1997: 1062):

(63) ... om net kontrolearre₃ hoeve₁ te wurden₂
to not check.PTCP need.INF to become.GER
‘in order not having to be checked’ 312 *Frisian*

This pattern obtains if there is displacement and haplology. The fact that V1 appears as a bare infinitive and not as a gerund, which is normally selected by *te*, suggests that there has been no *te*-deletion (Eric Hoekstra p.c.).

Missing-*zu* also seems to be residually possible in Standard German. Consider the following examples where *ohne* ‘without’ selects a *zu*-infinitive as does the V1 of the verb cluster *glauben* ‘believe’ (scrambling of the object ‘the book’ ensures that we are dealing with a coherent construction):

(64) a. ohne das Buch verstehen₃ (??zu) glauben₁ zu können₂
without the book understand.INF to believe.INF to can.INF 312

b. ohne das Buch (??zu) glauben₁ verstehen₃ zu können₂
without the book to believe.INF understand.INF to can.INF
‘without believing to be able to understand the book’ 132
Standard German

<http://www.chefkoch.de/forum/2,22,296109/An-alle-CHer-Wir-zelebrieren-den-Kantoenligeist.html>, accessed March 28, 2013.

b. S Ziel isch nid blibe z' stah sondern versueche₁ glich z bliebe₂
the goal is not stay.INF to stand.INF but try.INF same to stay.INF
‘The goal is not to make no progress but to try to remain the same’

<http://www.mosiweb.ch/maennerriege/maennerriege.htm>, accessed March 18, 2016

Another counter-example can be found in the description of the dialect of Bosco Gurin, see Comrie and Frauenfelder (1992: 1058) (the complementizer *fer* selects a *z* as does *tia/tian*; the infinitive of causative ‘do’ always appears as a gerund):

(iii) Ech ha ts Büach kchöiff, fer ts Chenn tian₁ waldsch z leeran₂.
I have.1SG the book bought for the child make.GER Italian to learn.GER
‘I bought the book in order to make the child learn Italian.’ *dialect of Bosco Gurin*

It seems to me that under the intended interpretation, a second *zu* leads to strong degradation.³¹ Attested examples are few, for what it is worth, here are two instances found on the Internet:

- (65) a. Also schreibt einfach normale Bewertungen, ohne versuchen zu betrügen
 So write.IMP simply normaly reviews without try.INF to cheat.INF
 ‘So just write normal reviews without trying to cheat.’
<http://www.elitevpers.com/forum/tutorials/2512465-geld-verdienen-durch-billiger-de-599.html>, accessed March 18, 2016
- b. Aber Windows scheint ja versuchen zu starten (siehe Fotos).
 but windows seem.3SG indeed try.INF to start.INF see pictures
 ‘but Windows indeed seems to try to start (see pictures).’
<http://www.informationsarchiv.net/topics/45648/>, accessed March 18, 2016

This suggests that at least for some speakers extraposition of the complement of the restructuring verb is optional in Standard German as well.

Finally, haplology effects are also found with other types of displaced morphology. For instance, Höhle (2006: 70) provides an example from the dialect of Barchfeld where both V1 and V2 select a *ge*-infinitive, but only one *ge*-infinitive is found, on V3.

4.6 Further evidence for the split complementation vs. adjunction

If verb clusters/VPR and the 3rd Construction indeed differ w.r.t. such a fundamental structural relationship, i.e. complementation vs. adjunction, we expect to find further reflexes of this distinction. In this subsection, I will provide evidence for two such reflexes.

4.6.1 Relative clause extraposition

Like other languages, German allows for simultaneous extraposition of (finite and non-finite) complement/argument clauses and relative clauses:

- (66) a. dass Maria zu einem Mann sagte, [den sie nicht kannte], dass er ihr
 that Mary to a man say.PST.3SG who she not know.PST.3SG that he her.DAT
 gefalle
 please.SUBJ.3SG
 ‘that Mary said to a man she didn’t know that she likes him’
- b. dass sie einem Mann versprach, [den sie nicht kannte], ihm beim
 that she a.DAT man promise.PTCP who she not know.PST.3SG he.DAT at the
 Umzug zu helfen
 move to help.INF
 ‘that she promised a man she didn’t know to help with the move’

Standard German

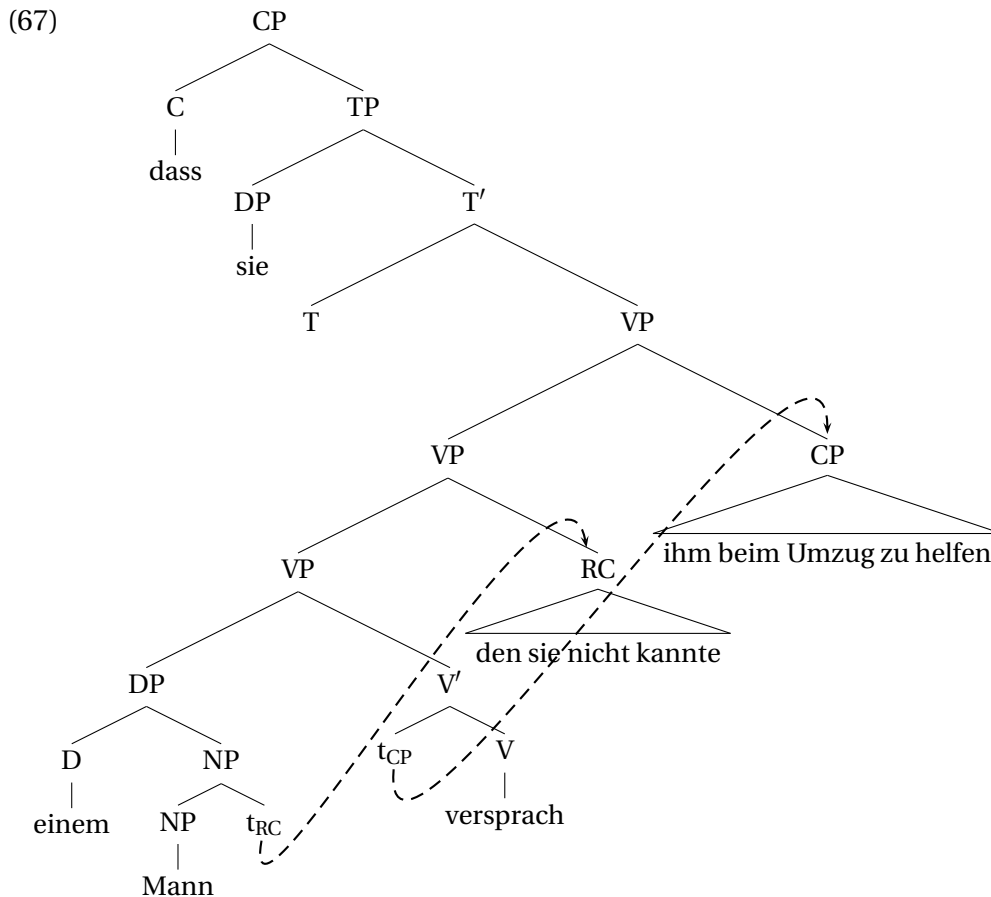
There has been some debate about the possible orders. While Haider (2010: 199f.) argues that the order is fixed with the RC necessarily preceding the complement clause, Sternefeld (2006: 783)

³¹Note that a *zu* before V1 in (64-b) is fully acceptable (and arguably preferred) once the object ‘the book’ remains inside the lexical VP:

- (i) Ohne *(zu) glauben₁, das Buch verstehen₃ zu können₂.
 without to believe.INF the book understand.INF to can.INF *St. German*

(i) arguably represents a non-restructuring configuration with the complement of V1 being extraposed, basically as with finite CP-complements discussed in section 4.4.3.

proposes that both orders are possible and that preferences for one or the other order depend on heaviness. I will not enter this debate here because the position of CP-argument clauses is orthogonal to the questions pursued here; but as we will see below, transparent non-finite complement clauses can precede as well as follow the extraposed RC. What is crucial for present purposes is that short RC-extrapolation presupposes an adjunction site between the main clause and the complement clause. Consequently, short RC extrapolation is only possible if the complement clause is extraposed as well (note that this also holds if extraposed RCs are base-generated in adjoined position), see (67), the structure of (66-b) (again mostly for representational simplicity, I use the classical right-adjunction analysis of extrapolation):



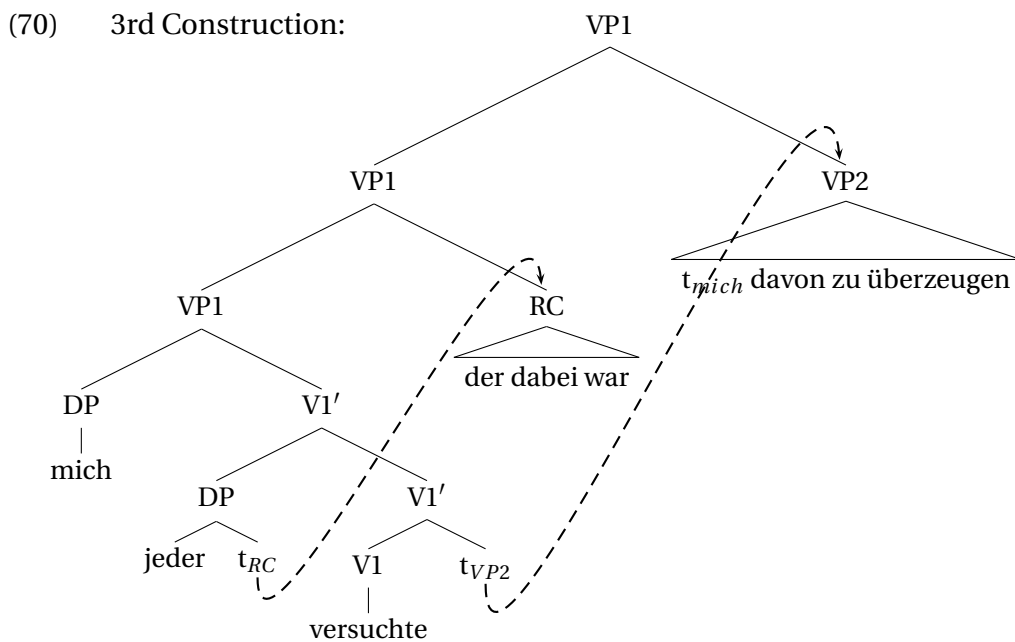
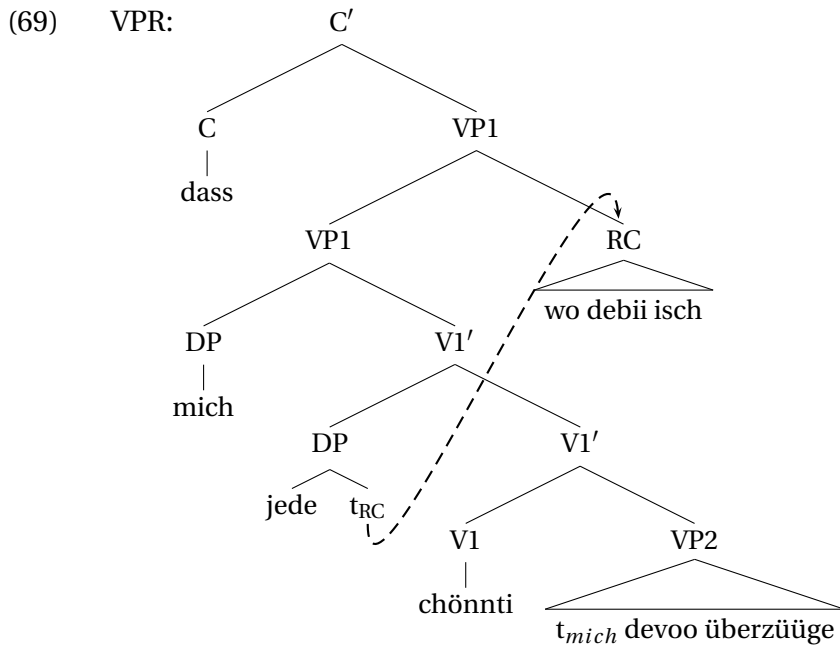
Since VPR and the 3rd Construction differ from each other in that the former involves a complementation structure while the latter is based on a non-complementation structure, we expect an asymmetry with respect to RC-extrapolation: the extraposed RC should be able to precede the VP-complement in the 3rd Construction but not in VPR. This prediction is borne out: while VPR requires long extrapolation, the RC can both precede and follow the complement in the 3rd Construction (pronoun fronting in both examples ensures that we are dealing with a coherent construction):

- (68) a. dass mich jede chönnti₁, [~~X~~ wo debii isch], t_{mich} devoo überzüüge₂, [~~✓~~
 that me everyone could.3SG C present be.3SG of.it convince.INF
 wo debii isch]
 C present is
 ‘that everyone who is present could convince me of it’ VPR; Swiss German

- b. dass mich jeder versuchte₁, [✓ der dabei war], t_{mich} davon zu
 that me everyone try.PST.3SG who present be.PST.3SG of.it to
 überzeugen₂, [✓ der dabei war]
 convince.INF who present be.PST.3SG
 ‘that everyone who was present tried to convince me of it’ 3rdC; *Standard German*

Although the test is difficult to apply in the case at hand because transparent non-finite complements involve less structure and will consequently not be very heavy, the contrast is nevertheless clear.

The structures of the two examples consequently look as follows:



Since the lexical VP is a complement in VPR, the RC will invariably follow it, irrespective of where it is adjoined. In the 3rd Construction, however, both the RC and FP are adjoined to VP1. Since nothing regulates their order, we correctly expect both options to be possible. RC-extrapolation is

thus another diagnostic to distinguish between complementation and non-complementation:³²

- (71) short RC extraposition → adjunction
 obligatory long RC extraposition → complementation

4.6.2 Stranding of the lexical VP

The split complementation vs. non-complementation also predicts an asymmetry w.r.t. stranding if in a three-verb-cluster VP2 undergoes topicalization: it should be possible to strand VP3 in the 3rd Construction since it undergoes extraposition so that topicalization of VP2 is remnant movement. In Verb clusters/VPR, however, where the lexical VP remains a complement, topicalization of VP2 cannot strand VP3. This prediction is borne out (again, pronoun fronting ensures that we are dealing with coherent constructions):

- (72) a. [_{VP2} versucht₂ t_{VP3}] hat₁ er sie nicht t_{V1} t_{VP2}, [_{VP3} seinen Eltern t_{si} vorzustellen₃]
 tried has he her not his parents introduce.to.INF
 ‘He didn’t try to introduce her to his parents.’ 3rdC StG
- b. * [_{VP2} welle₂] het₁ er si scho t_{V1} t_{VP2} [_{VP3} siine Eltere t_{si} vorstelle₃]
 want.INF had.SUBJ he her indeed his parents introduce.to.INF
 ‘He would have liked to introduce her to his parents.’ VPR SwG
- c. * [_{VP2} wollen₂] hat₁ er si seinen Eltern schon t_{V1} t_{VP2} [_{VP3} t_{seinen Eltern} t_{si}
 want.INF have.3sg he her his parents indeed
 vorstellen₃].
 introduce.INF
 ‘He would have liked to introduce her to his parents.’ VC StG

Note that in (72-a), since VP3 is clause-final and follows negation, it cannot have undergone scrambling, which would be an alternative to extraposition to form a remnant; under scrambling, VP3 would have to precede negation. The ungrammaticality of (72-b/c) confirms the general assumption that non-finite VPs without *zu* do not scramble in German, see e.g. Müller (1995: 154).

4.7 Intermediate summary

The previous sections have provided new evidence that Verb clusters/VPR and the 3rd Construction must be distinguished because the lexical VP is in fundamentally different structural relations with respect to the governing verb(s) in the two constructions: while verb clusters/VPR involve a complementation relation, the lexical VP is in a non-complementation/adjunction relation in the 3rd Construction. This structural difference triggers reflexes in various areas of grammar: in displaced morphology, haplology effects with *zu*, relative clause extraposition and stranding of the lexical VP.

Having established that the two constructions need to be distinguished, we can now apply the diagnostics to the Swiss German 213 construction.

³²The contrast between VPR and the 3rd Construction provides an argument for a crucial syntactic component in extraposition: since the two constructions do not differ prosodically, the RC-extraposition asymmetry must be related to syntactic structure, viz., the (un-)availability of an adjunction site (or whatever derives the complementation/-non-complementation contrast). Consequently, the non-finite VP in the 3rd Construction must undergo extraposition in syntax; purely phonological placement will not be sufficient. Haider (2010: 221) claims that a movement account wrongly predicts the order to be invariably complement > RC. However, this is an incorrect interpretation of cyclicity; cyclicity only pertains to the root, it does not prevent complement extraposition from following RC-extraposition as long as the complement attaches to a higher position than the RC. Note that the facts discussed here only imply that the complement undergoes syntactic movement; the RC-extraposition facts can also be captured if the RC is base-generated in adjoined position.

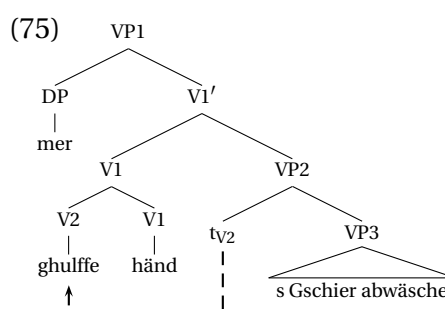
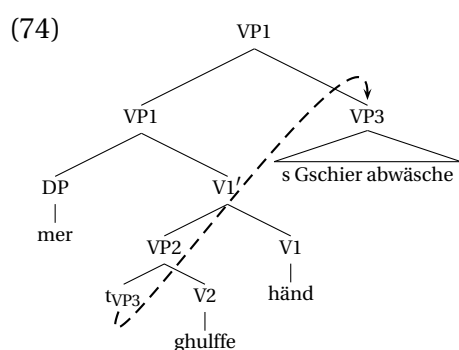
5 Applying the complementation diagnostics to Swiss German

Recall first the analytical possibilities for an example like (24-a), repeated in (73) for convenience:

- (73) wo s mer ghulffe₂ händ₁ s Gschier abwäsche₃
 when they me.DAT help.PTCP have.1P the dishes wash.INF
 ‘when they helped me wash the dishes’

Swiss German

The crucial difference between verb clusters/VP_R and the 3rd Construction consists in the integration of the lexical VP. It is a complement in the former but a non-complement in the latter. The Swiss German 213 construction can thus potentially involve a structure as in (74-a) or (74-b) (for reasons of space, I only provide two structures, one based on a left-branching and one based on a right-branching base order):



I will now apply the diagnostics introduced in the previous section to both the Swiss German 213 construction as well as the Standard German 3rd Construction in 213 order to find out whether they pattern differently or not.

5.1 Displaced *zu*

Recall that displaced *zu* is a diagnostic for complementation while non-displacement indicates non-complementation/adjunction. I will first apply displaced *zu* to 213 orders with V2-selecting a *zu*-infinitive before addressing the Swiss German 213 construction.

5.1.1 *z*-placement in 213 orders with V2 selecting a *z(u)*-infinitive

Looking first at the Standard German 3rd Construction in a 213 order, the result is very clear and as expected: there is no displacement, confirming that it involves non-complementation/adjunction of VP3 (scrambling of *dem Hans* shows that we are dealing with a coherent construction):

- (76) ohne dem Hans versucht₂ *(**zu**) haben₁ zu helfen₃
 without the.DAT John tried to have.INF to help.INF
 ‘without having tried to help John’

3rC Standard German

Things are different in Swiss German: Recall from section 4.5 that complements of restructuring verbs (i.e. FPs) can be treated as either adjuncts (no displacement) or complements (missing *z*). Consequently, in (77), *z* on V1 is optional (the absence of *z* on V2 is, of course, ungrammatical):

- (77) ohni em (z) probiere₁ *(z) hälffe₂
 without he.DAT to try.INF to help.INF
 ‘without trying to help him’

Swiss German

Turning to a variant of (77) in 213 order, we find the same result: both complementation (missing *z*) or adjunction (2x *z*) are possible (while no *z* on V2 leads to ungrammaticality):

- (78) a. ohni s versuecht₂ **z** ha₁ **z** läse₃ adjunction
 b. ?ohni s versuecht₂ ha₁ **z** läse₃ complementation
 c. *ohni s versuecht₂ **z** ha₁ läse₃
 without it try.PTCP to have.INF read.INF
 ‘without having tried to read it’ Swiss German

This is already an important result because (78-b) shows that 213 orders in Swiss German *can* behave like proper verb clusters/VPR! This is thus the first piece of evidence for the existence of proper verb clusters with 213 order.

5.1.2 *z*-placement in SwG 213 clusters with V2 selecting a bare infinitive

I now turn to the Swiss German 213 construction where V3 appears in the bare infinitive. Before tackling three-verb clusters in 213 order it is instructive to have a look at 2-verb clusters with V1 belonging to the class of predicates that allow for 213. Here, the result is very clear: there is displacement of *z*, showing that we are dealing with a complementation structure and thus a proper verb cluster. Here are some attested examples from the Internet (note that since some speakers optionally allow a *z*-infinitive with ‘help’, ‘teach’ and ‘begin, (79-b-d) may be instances of missing *z*):

- (79) a. soooo schön, di wieder mal ghöre₁ **z** singe₂
 so nice you again once hear.INF to sing.INF
 ‘so nice to hear you sing again’
<https://www.facebook.com/video/video.php?v=10200666450684322>, accessed March 28, 2016
 b. ... mithälfe, e Teil vo de neue Bronceglogge hälfe₁ **z** finanziere₂.
 aid.INF a part of the new bronze bell help.INF to finance.INF
 ‘to contribute to help finance part of the new bronze bell’
<https://www.google.ch/#psj=1&q=%22h%C3%A4lfe+z+%22&start=20>, accessed March 28, 2016
 c. Es esch cool gsi bi der lehre₁ **z** fahre₂.
 it be.3SG cool be.PTCP at you.DAT learn.INF to drive.INF
 ‘It was cool to learn to drive with you.’
<http://www.fahrlehrervergleich.ch/bewertungen.php?irat=7442>, accessed March 28, 2016
 d. ich liieb es mit wildfrämde lüt afange₁ **z** rede₂!
 i love it with strange people start.INF to talk.INF
 ‘I love it to start talking to complete strangers’
<http://giannaferrari.blogspot.de/2012/05/sommarya.html>, accessed March 28, 2016

With a 213 order, the result is the same: there is displacement and thus complementation:

- (80) a. *ohni en ghört₂ **z** ha₁ singe₃ extraposition
 b. ?ohni en ghört₂ ha₁ **z** singe₃
 without him heard have.INF to sing.INF
 ‘without having heard him sing’ complementation
Swiss German

It is difficult to find naturally occurring examples. Here is one from the internet (note that since some speakers can optionally use *aafange* ‘begin’ with a *z*-infinitive, this could also be an instance of missing *z*):

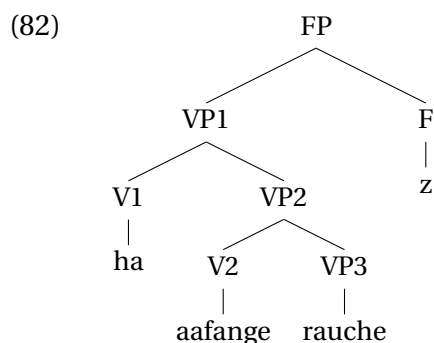
- (81) Wieder en grund meh zum glücklich drüber sii, niä agfange₂ ha₁ z
 again a reason more to happy about.it be.INF never begin.PTCP have.INF to
 rauche₃!
 smoke.INF

‘Another reason to be happy to have never started smoking!’

<https://www.facebook.com/Radio24/posts/10151574652070814>, accessed March 28, 2016

This implies that these Swiss 213 clusters represent a proper verb cluster structure; they are not an instance of the 3rd Construction. Note that the Standard German equivalent of (81) would require a *zu* before V1, in conformity with our expectations: While the special set of verbs in Swiss German that allows for the 213 order selects a bare infinitive, all these verbs require a *zu*-infinitive in the standard language and thus are expected to pattern like other clusters where the restructuring verb selects a *zu*-infinitive (e.g. like *versuchen* ‘try’).

The derivation of (81) in the framework adopted here proceeds as follows: given that verbal complements are ordered to the right of the verbal head while the functional head F is ordered to the right of its VP-complement, the input to the PF-derivation is as follows:



After linearization, there is cluster formation between V2 and V1, leading to a [21]3 order. In a final step, *z* undergoes Local Dislocation and is affixed onto the adjacent V3:

- (83) [V₂+V₁] z+V₃ Local Dislocation
 ↑ LD ↓

5.2 Relative clause extraposition

Recall that while short relative clause extraposition indicates adjunction/non-complementation, obligatory long extraposition is a diagnostic for complementation. The Standard German 3rd Construction patterns as expected, short RC extraposition is possible in 213, thus providing evidence for adjunction/non-complementation:

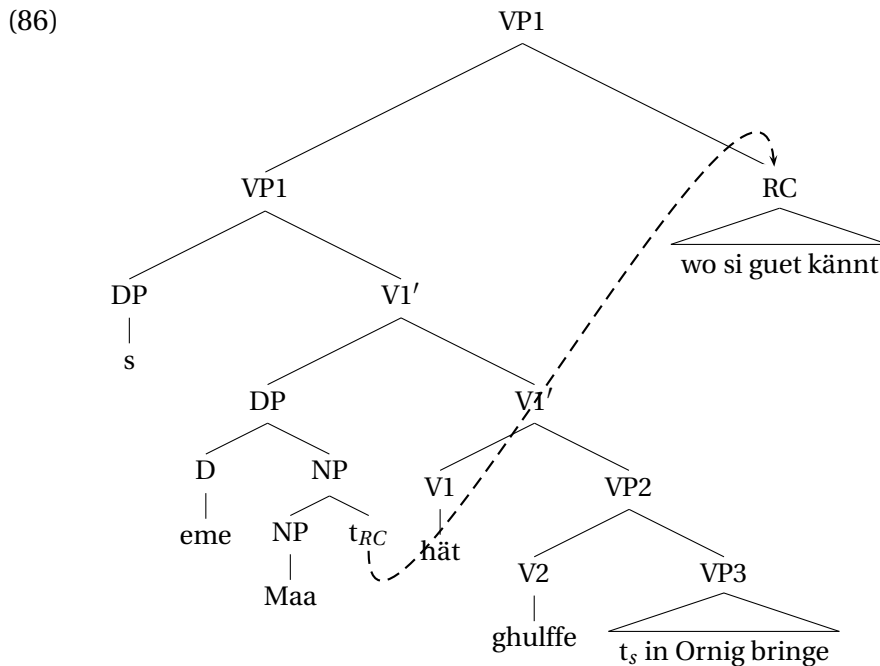
- (84) dass mich jeder versucht₂ hat₁, [✓ der dabei war], t_{mich} davon zu überzeugen₃, [✓
 that me everyone tried has who present was of.it to convince.INF
 der dabei war]
 who present was
 ‘that everyone who was present tried to convince me of it’ 3rdC StG

Things are different with the Swiss German 213 construction: short RC extraposition is strongly degraded, only extraposition to the end of the clause is acceptable:

- (85) dass si s eme Maa ghulffe₂ hät₁, [~~X~~ wo si guet käänt], t_s in Ornig bringe₃, [~~✓~~ wo
 that she it a.DAT man helped has C she well knows in order bring.INF C
 si guet käänt]
 she well knows
 ‘that she helped a man who she knows well to bring it in order’ SwG

This is thus further evidence that the Swiss German 213 construction involves complementation.

The structure of (85) looks as follows (since this is a syntactic representation, cluster formation between V1 and V2 has not yet taken place):



The facts from displaced *z* and relative clause extraposition thus show very clearly that the Swiss German 213 construction involves complementation. Since complementation is a hallmark of proper verb clusters/VPR, I conclude that the 213 orders represent relevant cluster orders. In the next section, I discuss the implications of this finding for the theory of verb clusters.³³

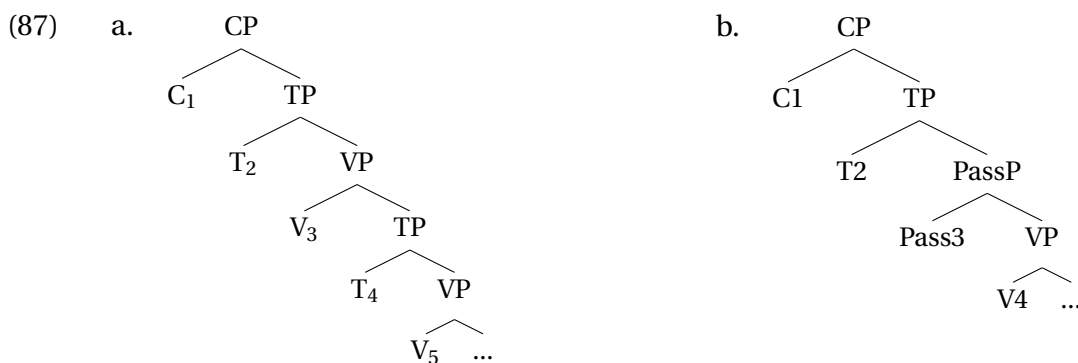
6 213 orders and the theory of verb clusters

As the previous section has shown, the Swiss German 213 clusters behave like proper verb clusters with respect to the complementation diagnostics displaced *zu* and relative clause extraposition. This strongly suggests that all six logically possible cluster orders exist and that theories of verb clusters must be more powerful than has been claimed in some of the (recent) literature.

³³The test involving stranding of the lexical VP requires 4-verb clusters. Again, remnant topicalization of VP2 is only possible in the 3rd Construction but not in the 213-construction ((i-b), which can be based on a 1324 or 3214 order, is grammatical if VP4 is part of the topicalized phrase):

- (i) a. [versucht₃ haben₂] sollte₁ er sie schon, [seinen Eltern vorzustellen₄].
 try.PTCP have.INF should.3sg he her indeed his parents introduce.to.INF
 ‘He should have indeed tried to introduce her to his parents.’ St. German
- b. *[gsee₃ haa₂] sött₁ er si scho [luut singe₄]
 see.PTCP have.INF should.3SG he her indeed loud sing.INF
 ‘He should have indeed seen her sing loudly.’ Swiss German

Before adopting this conclusion, there is one obvious objection that needs to be addressed, see Abels (2015): the clusters under discussion involve a V2 that is much more lexical than the elements involved in the major cluster types, viz. modals and auxiliaries (and perhaps causative ‘let’). Consequently, the Swiss German clusters might instantiate lexical restructuring rather than functional restructuring. At least under the perspective taken in Abels (2015) this would imply that they do not constitute evidence against the neutral theory of word order developed in Cinque (2005), which Abels elegantly extends in slightly modified form to verb clusters: this theory (which I do not have space to lay out in detail) only applies to domains that contain a lexical head and its modifiers, which additionally have to belong to the same class. Verb clusters consisting of just one lexical and one or several functional verbs thus fall under this theory because the functional verbs are analyzed as satellites of the verb, concretely as functional heads in the extended projection of the lexical verb. The clusters I have been dealing with in this paper, however, may not fall under this theory because there are two lexical elements and thus two independent domains. Given the evidence that the two classes indeed pattern differently in some respects (cf. Wurmbrand 2004), this seems a reasonable conclusion (while both classes are lumped together as functional elements in Cinque 2006). The difference between lexical and functional restructuring can be illustrated schematically by the following structures taken from Abels (2015: 19):³⁴



As Abels points out, while C1, T2 and Pass3 can straightforwardly be analyzed as modifiers of V4 in the representation of functional restructuring in (87-b), C1, T2 and V3 cannot easily be treated as modifiers of V5 in lexical restructuring in (87-a).

While I will not dispute the basic logic of the argument, the result strikes me as unsatisfactory for two major reasons: first, the verbs in the Swiss German 213 construction do not behave consistently with respect to the lexical/functional split. Second, treating the Swiss German verbs as a separate class implies that we need another cluster-forming or at least another ordering mechanism to deal with the Swiss German 213 construction.

I will now go through these points in more detail: Concerning the lexical/functional split, of the six verbs that participate in the Swiss German 213 construction, two would have to be classified as functional according to the semantic criteria in Wurmbrand (2004), viz. the phasal verbs *aafange* ‘begin’ and *ufhöre* ‘stop’: basically like their Standard German equivalents, they allow for a weather-*it* and inanimate subjects but do not take internal arguments. The other four verbs of the 213 class *gsee* ‘see’, *ghöre* ‘hear’, *hälffe* ‘help’ and *lehre* ‘learn/teach’ display the inverse pattern. Furthermore, some of these six verbs may also be classified as semi-functional, e.g. the perception verbs, see Wurmbrand (2001: 215ff.). At least on semantic grounds, the verbs occurring in the

³⁴I should point out here that the lexical/functional divide has been called into question by Reis and Sternefeld (2004). Furthermore, the lexical VP in VPR, which certainly represents a coherent construction, involves more structure than just a VP according to most of the literature. See e.g. den Dikken (1996) for the claim that it corresponds to a TP. Consequently, the major difference between lexical and functional restructuring should arguably be located in the projection of the restructuring verb rather than in its complement.

213 construction may thus belong to up to three different groups.³⁵ Crucially though, there is no indication that the verbs occurring in the 213 order behave differently in the relevant respects, i.e. with respect to *zu*-displacement and relative clause-extrapolation. This suggests that the lexical vs. functional divide does not necessarily affect the ordering possibilities.³⁶ As for the second point: if we were to accept the conclusion that the Swiss German 213 construction involves lexical restructuring and therefore does not fall under Cinque's theory of unmarked word order, we would need an additional mechanism to generate the 213 order. Given that the obvious possibility, viz. extrapolation, is not an option, it remains completely unclear what this mechanism would be. What is clear is that this mechanism would have to be flexible enough to accommodate 123 and 231 orders for these verbs (in Swiss German) as well, but the different orders should not differ in interpretation. Furthermore, it would have to generate structures that display coherence effects and involve complementation. Of course, these are the hallmarks of verb clusters. In other words, it seems that adopting Abel's conclusion implies that one has to postulate a second cluster-forming mechanism that essentially only differs from the existing one in that it can also generate 213 orders. While not impossible, this strikes me as an extremely uneconomical and undesirable solution.

I consequently adhere to the assumption that all verb cluster phenomena displaying coherence and complementation properties should be handled by the same cluster-forming mechanism (not ruling out the possibility that the lexical/functional distinction plays a role for other aspects of the grammar of verb clusters).

In the remainder of this section, I will briefly address alternative theories of cluster formation and of *zu*-placement and discuss to what extent they can accommodate the Swiss German 213 construction. Starting with cluster theories, it is obvious that only theories that can generate 213 orders will be considered. Remnant movement based theories like Koopman and Szabolcsi (2000) can be combined with the approach to *zu*-placement proposed above (see also Hinterhölzl 2009). In this type of approach, the constituent containing the three verbal elements (VP1) will move into the specifier of the functional head into which *zu* will be inserted, viz. FP2 in (88) (as before, this is a strongly simplified derivation that is only meant to convey the basic idea of how remnant movement approaches work):

(88) [FP2 [VP1 [VP2 **V2** t_{VP3}] [V1' **V1** [FP1 **VP3** [F1' F1 t_{VP2}]]]] F2 t_{VP1}]

The *zu*-placement mechanism can arguably also be combined with Haegeman and van Riemsdijk (1986), provided that the late inserted morphology can interact with the reanalyzed structure; but given that this is where inversion takes place, this seems unproblematic.

The theory of *zu*-placement proposed above is not fully compatible with cluster theories that involve base-generated complex heads as in Haider (2003) or Bader and Schmid (2009). There are no problems if the functional head hosting the displaced morphology is above the verb cluster (e.g. if *zu* is selected by a noun or a complementizer). However, if one of the verb selects a *zu*, there

³⁵Wurmbrand (2004) proposes another criterion that sets apart functional and lexical restructuring, viz., the optionality of restructuring with lexical restructuring verbs. The obligatoriness of pronoun fronting (see (8-b)) suggests that the 213 construction involves obligatory restructuring and thus patterns with functional restructuring, contrary to Abel's claims. Unfortunately, the other diagnostics that Wurmbrand discusses cannot be applied to the verbs at hand because the context that separates restructuring from non-restructuring is the long-distance passive according to Wurmbrand, which is not compatible with the verbs occurring with 213 (the phasal verbs 'begin' and 'stop' are intransitive, the perception verbs cannot be passivized either when they take an infinitival complement, and with 'help' and 'learn/teach', the long-distance passive is ungrammatical, arguably like in the standard language). Other optionality criteria like the possibility of both intra- and extrapolation do not work either because there is arguably no extrapolation with bare infinitival clauses; rather, if VP3 contains lexical material in the 213 construction, we are simply dealing with VPR.

³⁶Note also that there is no perfect match between the ordering possibilities and the putative lexical status of a restructuring verb. While verbs like *probiere/versueche* 'try' can occur in 321 and 123 order next to the 213 order resulting from extrapolation, the verbs that occur in the 213 construction additionally allow the orders 123 and 231.

simply is no space for that functional head inside the complex head. However, I believe that these theories can derive the placement facts if they are combined with a representational approach to *zu*-placement, perhaps along the lines proposed in Bader (1995) and Vogel (2009). For reasons of space, I will focus on Vogel's approach. He proposes that *zu* is a feature that is assigned to the entire infinitival complement and is crucially not borne by the head of the verb phrase. The morphological realization of the feature is the result of the interaction of alignment constraints that force it to be realized in second to last position within a certain domain. The domain can be defined as follows:³⁷

(89) The domain of *zu*-placement is the XP bearing the *zu*-feature

This works for both well-behaved *zu* in descending orders as well as for displacement in (partially) ascending orders: the feature is realized on the right-most terminal of the relevant phrase. It thus captures the intuition that the placement of *zu* depends on the surface order within the verbal complex and not the hierarchical relations.

This approach can arguably be applied to all cluster theories discussed above, including those that posit a base-generated complex head (as long as the higher head c-commands the lower head): it is conceivable that e.g. V1 imposes a *zu*-infinitive on its complement, a complex head. The alignment constraints together with the domain-definition in (89) will make sure that *zu* is realized on the last verb of the complement of V1, irrespective of the cluster order (to be compatible with Haider's approach, one has to assume that the assignment of the *zu*-feature takes place after reprojection).³⁸

In conclusion, then, the Swiss German 213 construction is in principle compatible with several theories of verb clusters and different theories of *zu*-placement. I take this to be a positive result because it shows that the new diagnostics capture a very fundamental structural relation.

7 On the rarity of 213

The last sections have established that verb clusters with 213 order do exist. This raises the obvious question of why this order is so rare (if not inexistent) in other cluster types, both within Swiss German and beyond. I believe that the rarity of 213 can be (partially) motivated by functional considerations along the lines of Culicover (2013: 270-281): he proposes two biases in the processing of verb clusters: first, verbs are preferentially linearized according to their relative scope. This favors strictly ascending orders, viz. 12(3). The scope-bias is counter-balanced by what he calls a dependency bias: keeping an argument in memory until its predicate is encountered incurs a certain computational cost. In verb clusters where the lexical verb is usually the V2 or the V3, this cost is higher in ascending structures than in descending structures. Consequently, the dependency bias favors descending orders like (3)21. Given that both biases are present simultaneously, we also expect serializations that constitute compromises between the two biases, i.e. that are only partially ascending like 132, 312. Given the two biases, the 213 order emerges as the worst solution as it is clearly disfavored by both. Against this background, the rarity of 213 orders in most cluster

³⁷(89) is a modified version of Vogel's definition. His original formulation on p. 329, which defines *zu*-placement w.r.t. the extended projection of the phrase bearing the *zu*-feature, derives the wrong result in a number of cases, see Haider (2011: 250) and Author (xxxx) for discussion.

³⁸Vogel's approach is critically discussed in Author (xxxx). Apart from increasing the number of feature types, perhaps its most significant drawback is the fact that it is incompatible with Bare Phrase Structure. Furthermore, the constraint ranking essentially incorporates the descriptive generalization; by re-ranking the alignment constraints Vogel uses, one could easily model a variety where *zu* has the converse property, i.e. attaches to the leftmost verb of the relevant domain. It therefore misses a crucial property of displacement: it is related to the head-finality of the language.

types does not come as a surprise.

The fact that 213 is unmarked (in Swiss German and perhaps beyond) if V2 has more lexical content can perhaps be motivated by the dependency bias, at least with perception verbs and ‘help/teach/learn’: by placing V2 at the beginning of the cluster, it becomes closer to its arguments.³⁹ Importantly, there are additional processing-related facts that may favor 213 with the six special verbs in Swiss German: five out of these six verbs (all except for ‘see’) have separate participial forms while modals only have infinitival forms. Crucially, participial forms facilitate the parsing of verb clusters because the dependencies within the cluster can be determined more easily: the morphology makes it clear that V2 depends on V1. If V2 appears as an infinitive like V3, determining the relative dependencies is more complex. The fact that the V2s in this construction have more lexical content (than e.g. modals) will have a similar beneficial effect on parsing. Given these factors, the acceptability of the Swiss German 213 construction can be made sense of.⁴⁰

Before concluding, let me emphasize that these functional factors only provide motivation for the acceptability of the Swiss German 213 orders, but no explanation. The grammar often does not worry about functional factors and may thus grammaticalize structures that may seem suboptimal from a functional point of view. A good example of this is the 231 order in verb clusters. According to Culicover’s metric, it should be just as marked as 213 orders.⁴¹ Indeed, it is very marginal in German varieties and also in most of the Dutch language area. However, in Flemish varieties (cf. Haegeman 1998: 260, 273 and Schmid 2005: 78) and even more so in Afrikaans (Biberauer 2013), 231 orders are unmarked in Aux-Mod-Inf clusters, where Mod comprises modal, causative, benefactive, perception, durative and evidential verbs.

As a final point, the facts discussed in this paper imply that a theory of verb clusters must be able to generate 213 orders with the six verbs in question. What is not a priori clear is whether the theory should be restricted in such a way that it only generates 213 orders with these particular verbs or whether it should generate 213 orders across the board. The first option seems more attractive in that it avoids overgeneration. Whatever mechanism restricts the orders (linearization/inversion parameters, movement operations) will then have to be made sensitive to the relevant verbs in question. A consequence of this perspective is that occurrences of 213 in other cluster types like Aux-Mod-Inf have to be set aside as noise. This holds for the residual attestations mentioned in fn. 3 as well as examples found on the Internet like the following:

³⁹This is obvious for ‘help/teach’ but less so for the putative ECM-verbs ‘see’ and ‘hear’. Note, though, that these verbs are peculiar in that the ECM-subject must precede them even in 12(3) orders:

- (i) dass i {de Hans} gsee *{de Hans} ässe
 that I the John see.1SG the John eat.INF
 ‘that I see John eat’

Swiss German

This restriction is unexpected if ‘John’ is an argument of the embedded clause as c-command should be sufficient for case-assignment. If, however, it is actually an argument of the perception verb, its position falls out immediately given that arguments are ordered to the left of their predicates. The perception verbs would thus constitute control verbs like ‘help’ and ‘teach’. Note that the same ordering restriction holds for causative ‘let’, see Haegeman and van Riemsdijk (1986), also suggesting that a reanalysis in terms of Control may be in order. An alternative account of the placement facts might appeal to the directionality of Case assignment – along the lines of: accusative must be assigned to the left –, but such notions are not obviously compatible with current conceptions of Case assignment.

⁴⁰There are two further verbs that occur in verb clusters and have separate participial forms, viz. *laa* ‘let’ and *bliibe* ‘stay’. Neither of them allows for the 213 order, though.

⁴¹See in this context also Abels (2013), who argues that the markedness of 231 results from a mismatch between the prosodic and the syntactic structure that is not found in other cluster types.

- (90) a. Genau das was ich wölle₂ ha₁ ghöre₃.
 exactly that what I want.INF have.1SG hear.INF
 ‘Exactly that which I wanted to hear.’
http://457472.forumromanum.com/member/forum/entry_ubb.user_457472.3.1106773115.1106773115.1.ue_titel-swiss_elite_fighters.html, accessed March 22, 2016
- b. und ine verzell was i alles chöne₂ han₁ mache₃
 and them tell.1SG what I all can.INF have.1SG make.INF
 ‘and tell them what all I was able to do’ [http://forum.worldofplayers.de/forum/threads/655762-Dr-Schwizerclub-49-und-scho-sinds-verbii-d-Ferie/pa ge6](http://forum.worldofplayers.de/forum/threads/655762-Dr-Schwizerclub-49-und-scho-sinds-verbii-d-Ferie/pa%20ge6), accessed March 22, 2016

Similar restrictions will be needed for the 231 order. To my knowledge, it only occurs in Aux-Mod-Inf clusters (with Mod interpreted liberally) but not in Mod-Mod-Inf and Mod-Aux-Ptcp clusters.

When we look at the grammar of a single variety (which may be the grammar of a single individual), the same questions about restrictiveness arise. Either, whatever restricts the possible order will be specified even further (e.g. as in Bader and Schmid 2009) or we allow the grammar to over-generate; in the latter case, the orders that the speakers of a given variety consider acceptable may be due to extra-grammatical factors as proposed in Barbiers (2005).⁴²

I will not take a stand on the first issue, which is essentially a question about the interpretation of our data on which no final conclusions have been reached yet. As for the grammar of single varieties, I tend to favor an approach along the lines of Barbiers (2005) because such a perspective is better suited to deal with the pervasive gradience observed in most empirical work on verb clusters (cf. e.g. Seiler 2004, Bader and Schmid 2009). While there is frequently a dominant order for a cluster type, speakers often accept several orders to varying degrees. Classifying some of these orders as grammatical and others as ungrammatical can often only be done on an arbitrary basis, e.g. by means of a grammaticality threshold like 50%. It strikes me more plausible to treat such gradient data in terms of markedness, i.e. orders with different acceptability ratings do not differ from each other in terms of grammaticality but in terms of markedness (with the factors being partly extragrammatical and processing-related, cf. the biases from Culicover 2013 discussed above). Needless to say, these questions remain important issues for future research.⁴³

8 Conclusion

Much research on verb clusters in recent years has focused on the limits of variation with respect to the possible orders. Next to powerful theories that involve mechanisms that generate all six logically possible orders in three-verb clusters, more restrictive theories have been proposed that are designed to generate only five out of the six logically possible orders and categorically rule out the 213 order.

⁴²See Hendriks et al. (2015) for evidence that speaker’s judgments about non-native orders reflect the orders that can be generated by the grammar. This suggests that speakers have unconscious knowledge about more orders than are attested in their variety.

⁴³The theory proposed in Abels (2015) only applies to neutral word orders. At first sight, this may seem to help solve the gradience issue. However, quite apart from difficulties to determine the neutral order in a given cluster type, this view leaves the syntax of the marked orders open if I am not mistaken. If I understand the theory of word order correctly, marked orders are derived by different means than the neutral orders (these alternative means are not dealt with in the paper). Suppose now that in dialect X, there is an unmarked 123 order and a marked 132 order. This would seem to suggest that this 132 order is derived differently than a 132 order in dialect B, where it is the unmarked order. This strikes me as very implausible given that there is no evidence (to my knowledge) that marked 132 orders differ syntactically from unmarked 132 orders. Suppose instead that the marked 132 order is derived by the usual means. However, once marked orders can be derived by regular means, it is no longer clear to me that marked 213 orders (i.e. those residual occurrences in the major cluster types listed in fn. 3) cannot be derived by regular means as well.

Against this background, it is remarkable that Swiss German (and probably some other varieties within West-Germanic) features a verb cluster-like construction with an unmarked 213 order where V2 is a perception verb, a benefactive verb or a phasal verb taking a bare infinitival V3 as its complement. I have shown that this construction displays the familiar restructuring diagnostics like scrambling and pronoun fronting. I have argued that it is not an instance of the 3rd Construction, which features the same restructuring diagnostics and also allows for an unmarked 213 order. While previous work had not reached a consensus whether verb clusters/Verb Projection Raising and the 3rd Construction should be distinguished, I have shown, based on new diagnostics, that the two constructions differ with respect to a fundamental structural property: while the lexical VP is a structural complement of the governing verb in verb clusters/VPR, it is in a non-complementation/adjunction relation in the 3rd Construction. Several diagnostics converge on this result: displaced *zu*, relative clause extraposition and stranding of the lexical VP. Applying displaced *zu* and relative clause extraposition to the Swiss German 213 construction delivers a clear result: it displays the hallmarks of complementation and thus patterns like proper verb clusters. I have concluded from this that verb clusters with 213 order exist and that theories of verb clusters must be powerful enough to generate all six logically possible orders, contrary to some previous claims in the literature. Importantly, the facts as such are compatible with different theories of verb clusters and different theories of *zu*-placement, although not all combinations work. This is a positive result in my view because it shows that the constructions differ along very basic structural notions that any theory must be able to capture.

9 Appendix: Verb clusters/VPR vs. 3rdC: further diagnostics

In this appendix I will briefly discuss two diagnostics that have been proposed in the literature to distinguish between verb clusters/VPR and the 3rd Construction. In my view, they are inconclusive which is why they are not included in the main text. The first is the IPP-effect, the second concerns the semantic properties of XPs displaced from the lexical VP.⁴⁴

9.1 IPP-effect

The original motivation for the IPP-effect as a diagnostic for verb clusters comes from the following contrast in Standard Dutch: IPP is only possible in Verb Raising, i.e. if the verbs are adjacent, but not in the 3rd Construction, where there can be intervening non-verbal material, see Broekhuis et al. (1995: 99):

- (91) a. dat Jan het meisje {een kus} heeft₁ *proberen*₂ {*een kus} te geven₃
 that John the girl a kiss have.3SG try.INF a kiss to give.INF₃
 ‘that John tried to give the girl a kiss’ *Standard Dutch*
- b. dat Jan het meisje {een kus} *geprobeerd*₂ heeft₁ {een kus} te geven₃
 that John the girl a kiss try.PTCP have.3SG a kiss to give.INF₃
 ‘that John tried to give the girl a kiss’ *Standard Dutch*

Given that V2 appears as a participle in the Swiss German 213 construction, one may thus be tempted to take this as an indication that it instantiates the 3rd Construction. However, it is actually not quite clear what the IPP-effect diagnoses because it also occurs in VPR, where there is

⁴⁴I am aware of one further argument, viz. nominalized verb clusters where the direct object of the lexical verb is realized as a PP, see Neeleman (1990: section 5). Neeleman claims that while this is possible with proper verb clusters, it is not with the 3rd Construction. Unfortunately, I have not been able to establish clear contrasts. Once 3-verb clusters are involved, the acceptability is already strongly reduced for proper verb clusters so that unfortunately no conclusions can be drawn.

certainly no complex head, see the following examples from Swiss German (where the IPP-effect is generally optional, cf. Schmid 2005: 22f.), see Lötscher (1978: 3, fn. 2), and West Flemish, see Haegeman (1998: 275f.):

- (92) a. dass i de Hans ha₁ ghöört₂/ghööre₂ en Arie singe₃
 that I the John have.1s hear.PTCP/hear.INF an aria sing.INF
 ‘that I heard John sing an aria’ *Swiss German*
- b. dass i em Hans ha₁ ghulffe₂/hälffe₂ s Gschier abwäsche₃
 that I the.DAT John have.1s help.PTCP/help.INF the dishes wash.INF
 ‘that I helped John do the dishes’ *Swiss German*
- (93) da Valére ee₁ willen₂ Marie dienen boek geven₃
 that Valere has want.INF Mary that book give.INF
 ‘that Valere wanted to give Mary that book’ *West Flemish*

These facts suggest that the IPP-effect correlates with an ascending order. However, even this is not correct cross-linguistically. There are Austrian varieties where it occurs in descending 321 clusters, see e.g. Haider (2003). It thus remains completely unclear what exactly the IPP-effect diagnoses.

More relevant for the case at hand is the fact that the IPP-effect is also found in the Swiss German 213 construction as in (94) (IPP in 213 orders is also attested in Vorarlberg German, see Schallert 2014: 195f. and in earlier Pennsylvania Dutch, cf. Louden 2011: 178):

- (94) ... wo de Alkohol i mim Läbe afange₂ hät₁ e Rolle spile₃
 when the alcohol in my life start.INF have.3SG a role play.INF
 ‘when the alcohol started playing a role in my life’
<http://hpgmuender.blogspot.fr>; blog on September 29, 2007; accessed March 25, 2016

Whatever the IPP-effect indicates, given (94), there is no reason to conclude that the Swiss German 213 construction is an instance of the 3rd Construction.

9.2 Properties of XPs displaced from the lexical VP

Since extraposed complements arguably reconstruct obligatorily, they will not differ much from complements that are in-situ (cf. e.g. Sternefeld 2006: 781). Consequently, to detect differences between Verb clusters/VPR and the 3rd Construction, it is more promising to focus on the properties of XPs that are displaced from the lexical VP (and not on material within the lexical VP). As we will see presently, there are both similarities as well as (non)-systematic differences (in what follows, I omit examples with pure verb clusters, which pattern with VPR).

9.2.1 XPs displaced from the lexical VP: similarities

XPs displaced from the lexical VP behave the same with respect to a number of tests. First, they do not show *freezing effects*: They behave as if they were in their base-position (cf. also Geilfuß-Wolfgang 1991: 49)

- (95) a. Was_k hat₁ Heinrich ____k für einem Kind vergessen₂ die Zebras zu zeigen₃?
 what have.3SG Hendrik for a.DAT child forget.PTCP the zebras to show.INF
 ‘To what kind of child did Hendrik forget to show the zebras?’
Standard German, cf. Bayer and Kornfilt (1994: 45)

- b. Was_k tänsch, dass de Hans hät₁ ____k für Lüüt wele₂ vo siine Idee
 what believe.2SG that the Hans have.3SG for people want.INF of his ideas
 überzüüge₃?
 convince.INF
 ‘What kind of people do you think John wanted to convince of his ideas?’ *Swiss
 German, cf. Author (xxxx)*

Second, *focus projection is possible* (regardless of whether stress falls on the displaced XP or an XP within VP3), cf. Geilfuß-Wolfgang (1991: 25f.), Wöllstein-Leisten (2001: 96) (for Dutch cf. ter Beek 2008: 198f.):

- (96) a. Er hat₁ einem Kind versucht₂ das MÄRCHEN vorzulesen₃.
 he have.3SG a.DAT child try.PTCP the fairy.tale read.to.INF
 ‘He tried to read the fairy tale to a girl.’ *Standard German*
 b. Wenn er einem Kind hätte₁ das MÄRCHEN vorlesen₃ dürfen₂
 if he a.DAT child had.SUBJ.3SG the fairy.tale read.to.INF may.INF
 ‘if he had been allowed to read the fairy tale to a child’
Standard German

Third, the *displaced XP can belong to non-scrambleable categories*, e.g. *wh*-phrases (cf. Bayer and Kornfilt 1994: 45) and directional PPs (cf. Geilfuß-Wolfgang 1991: 31, 44):

- (97) a. ?Ich habe₁ ihm **was** versucht₂ nach Berlin zu schicken₃.
 I have.1SG he.DAT s.thing try.PTCP to Berlin to send.INF
 ‘I tried to send him something to Berlin.’ *Standard German*
 b. wenn ich ihm **was** hätte₁ nach Berlin schicken₃ können₂
 if I he.DAT something had.SUBJ.1SG to Berlin send.INF can.INF
 ‘if I could have sent him something to Berlin’ *Standard German*

These facts suggest that the displacement operation in these constructions differs from regular scrambling. Whether this implies that a different operation like e.g. pseudoscrumbling (Geilfuß-Wolfgang 1991) is involved or no movement whatsoever (cf. e.g. Bayer and Kornfilt 1994, Fanselow 2001, Author xxxx) is an open question; quite probably, the differences follow from independent factors (surface generalizations, freezing restricted to topical XPs etc.)

9.2.2 Semantic differences between the 3rd Construction and VPR

Next to these similarities, a number of systematic semantic asymmetries have been observed: while XPs displaced from the lexical VP are subject to semantic restrictions in the 3rd Construction, no restrictions are found in VPR. This asymmetry is also found in scopal interactions, where reconstruction is blocked in the 3rd Construction but possible in VPR. I will discuss both aspects in turn.

As for the semantic properties of the displaced XPs, it has been claimed that it cannot be a non-specific indefinite in the 3rd Construction, cf. Geilfuß-Wolfgang (1991: 42f.):

- (98) a. ??Peter hat₁ **einen Adventskalender** vergessen₂ zu basteln₃.
 Peter have.3SG a advent calendar forget.PTCP to make.INF
 ‘Peter forgot to make an advent calendar.’ *Standard German*
 b. dass Peter **einen Adventskalender** hat₁ für mich basteln₃ wollen₂
 that Peter a advent calendar have.3SG for me make.INF want.INF
 ‘that Peter wanted to make an advent calendar’ *St. German*

Nor can the displaced XP be an *idiom chunk* in the 3rd Construction (Geilfuß-Wolfgang (1991: 52)), while this is unproblematic in VPR (Author xxxx):

- (99) a. ?*Er hat₁ seinem Onkel **einen Bären** versucht₂ aufzubinden₃.
 he have.3SG his uncle a bear try.PTCP tie.to.INF
 ‘He tried to pull his uncle’s leg.’ *Standard German*
- b. dass er **känere Flüüg** hät₁ chöne₂ öppis z Leid tue₃
 that he no.DAT fly have.3SG can.INF something to suffering do.INF
 ‘that he could not harm anyone’ *Swiss German*

Turning to scopal interactions, scrambling a quantified XP over another normally leads to scope ambiguities and can be found in verb clusters/VPR; however, the ambiguity is lost in the 3rd Construction, cf. Geilfuß-Wolfgang (1991: 39):

- (100) a. Er HAT₁ *mindestens ein Geschenk* versucht₂ *fast jedem Gast*
 he have.3SG at least one present try.PTCP almost every.DAT guest
 t_{mindestens ein Geschenk} zu überreichen₃
 to hand.over.INF
 ‘He tried to hand over at least one present to almost every guest.’ $\exists > \forall$; * $\forall > \exists$;
Standard German
- b. DASS er *mindestens ein Geschenk* hat₁ *fast jedem Gast*
 that he at least one present have.3SG almost every.DAT guest
 t_{mindestens ein Geschenk} überreichen₃ wollen₂
 hand.over.INF want.INF
 ‘that he wanted to hand over at least one present to almost every guest’ $\exists > \forall$; $\forall > \exists$;
Standard German

If a quantified XP interacts with a scopal matrix verb, wide-scope seems to be obligatory in the 3rd Construction but not in VPR, cf. Author (xxxx) for VPR and Bobaljik and Wurmbrand (2005: 810, 831) for the 3rd Construction:

- (101) a. weil er [_{VP1} **alle Fenster** vergass₁ [_{VP2} t_{alleFenster} zu schliessen₂]
 because he all windows forget.PST.3SG to close.INF
 ‘because he forgot to close all the windows’
 $\forall > \text{forget}$; * $\text{forget} > \forall$; *Standard German*
- b. dass er [_{VP1} **2 Manager** wett₁ [_{VP2} t_{2 Manager} vo siine Idee überzüüge₂]]
 that he 2 managers want.3SG of his ideas convince.INF
 ‘that he wants to convince two managers of his ideas’
 $2 > \text{want}$; $\text{want} > 2$; *Swiss German*

9.2.3 A seemingly straightforward solution

The empirical facts can be summarized as follows: the displaced XP obligatorily takes wide scope in the 3rd Construction (precluding non-specific/non-referential interpretations) while reconstruction is fine in VPR. This generalization can be captured straightforwardly by the remnant extraposition analysis of the 3rd Construction, recall (26)/(28), because remnant movement is well-known to induce scope freezing effects (Barss 1986: 517–542, for a recent proposal, see Sauerland and Elbourne 2002):

- (102) reconstruction of α to its trace β is blocked if α does not c-command β at S-structure.

The following pair illustrates scope freezing with remnant topicalization (slightly adapted from Haider 2003: 101):

- (103) a. dass sie [_{VP1} **kein Fleisch** [_{VP2} *t_{kein Fleisch}* zu essen₂] *wagte*₁]
 that she no meat to eat.INF dared
 ‘that she didn’t dare to eat meat’ $\neg\exists > \text{dare}; \text{dare} > \neg\exists$
- b. [_{VP2} *t_{kein Fleisch}* zu essen₃] *wagte*₁ sie [_{VP1} **kein Fleisch** *t_{VP2}* *t_{wagte}*].
 to eat.INF dared she no meat
 ‘that she dared to eat no meat’ $\neg\exists > \text{dare}; * \text{dare} > \neg\exists$
Standard German

If the derivation of the 3rd Construction indeed involves remnant movement, obligatory wide-scope of the displaced XP follows straightforwardly. In VPR, however, where the lexical VP remains in its base-position, no remnant movement is involved so that movement can reconstruct, thus allowing for wide and narrow scope of the displaced XP.

9.2.4 Counterexamples

However, the facts are not as clear-cut as described above. The literature contains counter-examples suggesting that narrow scope is possible after all in the 3rd Construction: first, Geilfuß-Wolfgang (1991: 42f.) notes that existential/non-specific interpretations are sometimes (marginally) available; Wöllstein-Leisten (2001: 126f.) finds an existential interpretation in the 3rd Construction unproblematic, as does ter Beek (2008: 191ff.), who provides the following example:

- (104) omdat Jan **een huis** besloot₁ te kopen₂
 because John a house decide.PST.3SG to buy.INF
 ‘because John decided to buy a house’ *Standard Dutch*

Apparently, the type of matrix verb plays an important role. According to ter Beek (2008: 194, fn. 10): existential interpretations tend to be blocked with downward entailing matrix verbs like *vergetten* ‘forget’, *verlernen* ‘unlearn’, *vermeiden* ‘avoid’, *versäumen* ‘neglect’, *weigern* ‘refuse’ und *verbieten* ‘forbid’, while they are more readily available with verbs like *versuchen* ‘try’ etc. (although there are conflicting judgments).

Similarly, ter Beek (2008: 195ff.) provides well-formed examples with displaced idiom chunks in the 3rd Construction, some of which also seem well-formed in German:

- (105) a. omdat Jan **de zak** beweert₁ te krijgen₂
 because John the bag claim.3SG to get.INF
 ‘because John claims to get sacked’ *Standard Dutch*
- b. omdat Jan Marie **een loer** besloot₁ te draaien₂
 because John Mary a lurk decide.PST.3SG to turn.INF
 ‘because John decided to play a nasty trick on Mary’ *StD*
- c. weil Hans der Maria **einen Streich** versuchte₁ zu spielen₂
 because John the.DAT Mary a trick try.PST.3SG to play.INF
 ‘because John tried to play a trick on Mary’ *Standard German*

Interestingly, Geilfuß-Wolfgang (1991: 52) notes that the type of matrix verb is important; the examples deteriorate with verbs like *vergetten* ‘forget’.

Counter-examples are also found in the scopal interaction between the matrix verb and a QP: reconstruction does seem to be possible in the 3rd Construction in some instances ((106-a) is from Sternefeld 2006: 654):

- (106) a. Ratzinger, der **keine Kompromisse** *bereit ist*₁ einzugehen₂
 Ratzinger, who no compromises willing be.3SG to make.INF
 ‘Ratzinger, who is not willing to make any compromises’
 $\neg > \text{willing} > \exists$; *Standard German*
- b. dass du **keinen Schlips** *brauchst*₁ anzuziehen₂
 that you no tie need.2SG to wear.INF
 ‘that you need not wear a tie’
 $\neg > \text{need} > \exists$; *Standard German*
- c. dass er **kein Fleisch** *versuchte*₁ / *wagte*₁ zu essen₂
 that he no meat try.PST.3SG/dare.PST.3SG to eat.INF
 ‘That he tried/dared not to eat any meat’
 $\text{try} > \neg\exists$; *StG*

Again, the choice of matrix verb seems crucial: reconstruction with verbs like *vergessen* ‘forget’ or *verbieten* ‘forbid’ does not seem acceptable.

The restrictions in the 3rd Construction are reminiscent of a weak island effect, which could be accounted for if the fronting operation is A'-movement. However, quite apart from the controversies about whether scrambling involves A- or A'-movement (cf. e.g. Müller 1995), it is not obvious that the fronting operation can be assimilated to regular scrambling (see the references above). Given these uncertainties, the data in the 3rd Construction remain puzzling.⁴⁵

9.2.5 Summary

There is a certain asymmetry between VPR and the 3rd Construction with regard to the interpretation of DPs displaced from the lexical VP: in the 3rd Construction, there is a strong tendency for the DP to take wide-scope while in VPR both wide and narrow scope seem equally available. This asymmetry would fit perfectly with the results reached in this paper that the 3rd Construction involves remnant extraposition so that the lexical VP is in an adjunction/non-complementation relationship while it is a complement in VPR. However, the data in the 3rd Construction are partially conflicting with a number of counter-examples so that no firm conclusions can be drawn.

⁴⁵If what looks like the 3rd Construction in Swiss German can actually involve a complementation structure, see sections 4.5 and 5.1.1, we expect the possibility of scope reconstruction, as in VPR-structures. Since the facts are subtle, I will only point out that Cooper (1995: 197, 199, fn. 39) argues that scope reconstruction is possible in the 3rd Construction in Zurich German.

References

- Abels, Klaus (2011): 'Hierarchy-order relations in the Germanic verb cluster and in the noun phrase', *Groninger Arbeiten zur Germanistischen Linguistik* **53**(2), 1–28.
- Abels, Klaus (2013): 'On 2-3-1', *UCL Working Papers in Linguistics* **25**, 67–101.
- Abels, Klaus (2015): The fundamental left-right asymmetry in the Germanic verb cluster. Ms. University College London, available at <http://ling.auf.net/lingbuzz/002613>.
- Abels, Klaus and Ad Neeleman (2012): 'Linear Asymmetries and the LCA', *Syntax* **15**(1), 25–74.
- Bader, Markus and Tanja Schmid (2009): 'Verb clusters in VB<colloquial German', *The Journal of Comparative Germanic Linguistics* **12**(3), 175–228.
- Bader, Thomas (1995): Missing and misplaced z' in Bernese Swiss German. In: Z. Penner, ed., *Topics in Swiss German syntax*. Peter Lang, Bern, pp. 19–28.
- Barbiers, Sjef (2000): The Right Periphery in SOV Languages: English and Dutch. In: P. Svenonius, ed., *The Derivation of VO and OV*. John Benjamins, Amsterdam, pp. 181–218.
- Barbiers, Sjef (2005): Word order variation in three verb clusters and the division of labor between generative linguistics and sociolinguistics. In: L. Cornips and K. P. Corrigan, eds, *Syntax and Variation. Reconciling the Biological and the Social*. John Benjamins, Amsterdam, pp. 233–264.
- Barbiers, Sjef (2008): 'Werkwoordclusters en de grammatica van de rechterperiferie', *Nederlandse Taalkunde* **13**(2), 160–188.
- Barbiers, Sjef and Hans Bennis (2010): De plaats van het werkwoord in zuid en noord. In: J. De Caluwe and J. V. Keymeulen, eds, *Voor Magda. Artikelen voor Magda Devos bij haar afscheid van de Universiteit Gent*. Academia Press, Gent, pp. 25–42.
- Barss, Andrew (1986): Chains and anaphoric Dependencies. Doctoral dissertation, MIT.
- Bayer, Josef and Jaklin Kornfilt (1994): Against Scrambling as an instance of Move-alpha. In: N. Corver and H. v. Riemsdijk, eds, *Studies on Scrambling. Movement and Non-Movement Approaches to Free Word-Order Phenomena*. Mouton, Berlin, pp. 17–60.
- Bayer, Josef, Tanja Schmid and Markus Bader (2005): Clause union and clausal position. In: M. d. Dikken and C. M. Tortora, eds, *The function of function words and functional categories*. John Benjamins, Amsterdam, pp. 79–113.
- Besten, Hans den and Hans Broekhuis (1992): 'Verb Projection Raising in het Nederlands', *Spektator* **21**, 21–34.
- Biberauer, Theresa (2013): Germanic verb clusters again – with Afrikaans center stage. Paper presented at the Comparative Germanic Syntax Workshop 28, October 3-5, Leipzig. .
- Bobaljik, Jonathan and Susi Wurmbrand (2005): 'The Domain of Agreement', *Natural Language & Linguistic Theory* **23**(4), 809–865.
- Bölsing, Friedrich (2011): *Niederdeutsche Sprachlehre: Plattdeutsch im Kirchspiel Lindhorst, Schaumburg-Lippe*. Olms, Hildesheim.
- Broekhuis, Hans (1993): 'Verb Projection Raising', *Spektator* **22**, 28–47.
- Broekhuis, Hans, Hans den Besten, Kees Hoekstra and Jean Rutten (1995): 'Infinitival complementation in Dutch: On remnant extraposition', *The Linguistic Review* **12**(93-122).
- Cinque, Guglielmo (2005): 'Deriving Greenberg's Universal 20 and its exceptions', *Linguistic Inquiry* **36**(3), 315–332.

- Cinque, Guglielmo (2006): Restructuring and functional structure. In: G. Cinque, ed., *Restructuring and functional heads: The cartography of syntactic structures volume 4*. Oxford University Press, New York, pp. 11–63.
- Comrie, Bernard and Uli Frauenfelder (1992): ‘The verbal complex in Gurinerdeutsch’, *Linguistics* **30**(6), 1031.
- Cooper, Kathrin E. (1995): *Topics in Zurich German syntax*. Germanistisch Instituut, Groningen.
- Culicover, Peter (2013): *Grammar and complexity: Language at the intersection of competence and performance*. Oxford University Press, Oxford.
- de Vries, Mark (2002): *The Syntax of Relativization*. LOT, Utrecht.
- 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: W. Abraham, S. D. Epstein, H. Thrainsson and C. J.-W. Zwart, eds, *Minimal Ideas. Syntactic Studies in the Minimalist Framework*. Benjamins, Amsterdam, pp. 67–96.
- Den Dikken, Marcel and Eric Hoekstra (1997): ‘Parasitic participles’, *Linguistics* **35**(6), 1057–1090.
- Embick, David and Rolf Noyer (2001): ‘Movement Operations after Syntax’, *Linguistic Inquiry* **32**(4), 555–595.
- Evers, Arnold (1975): The transformational cycle of Dutch and German. Doctoral dissertation, University of Utrecht.
- Fanselow, Gisbert (2001): ‘Features, Q-Roles, and Free Constituent Order’, *Linguistic Inquiry* **32**(3), 405–437.
- Fleischer, Jürg and Oliver Schallert (2011): *Historische Syntax des Deutschen: eine Einführung*. Narr, Tübingen.
- Geilfuß-Wolfgang, Jochen (1991): Scrambling und Pseudoscambling. In: J. Geilfuß-Wolfgang, ed., *Verb- und Verbphrasensyntax. Arbeitspapier 11 des Sonderforschungsbereichs 340 Sprachtheoretische Grundlagen für die Computerlinguistik*. University of Tübingen, Tübingen, pp. 19–57.
- Glaser, Elvira (in preparation): Dialektsyntax des Schweizerdeutschen. Syntaktischer Atlas der Deutschen Schweiz (SADS). .
- Grewendorf, Günther and Joachim Sabel (1994): ‘Long Scrambling and Incorporation’, *Linguistic Inquiry* **25**(2), 263–308.
- Grimm, Jacob (1837): *Deutsche Grammatik. Vierter Theil.*. Dieterichsche Buchhandlung, Göttingen.
- Haegeman, Liliane (1992): *Theory and Description in Generative Syntax. A case study in West Flemish*. Cambridge University Press, Cambridge.
- Haegeman, Liliane (1998): ‘V-Positions and the Middle Field in West Flemish’, *Syntax* **1**(3), 259–299.
- Haegeman, Liliane and Henk van Riemsdijk (1986): ‘Verb Projection Raising, scope, and the Typology of Rules Affecting Verbs’, *Linguistic Inquiry* **17**(3), 417–466.
- Haider, Hubert (1993): *Deutsche Syntax – generativ: Vorstudien zur Theorie einer projektiven Gram-*

- matik*. Narr, Tübingen.
- Haider, Hubert (2003): V-clustering and clause union: Causes and effects. In: P. Seuren and G. Kempen, eds, *Verb constructions in German and Dutch*. John Benjamins, Amsterdam, pp. 91–126.
- Haider, Hubert (2010): *The syntax of German*. Cambridge University Press, Cambridge.
- Haider, Hubert (2011): ‘Grammatische Illusionen – Lokal wohlgeformt – global deviant’, *Zeitschrift für Sprachwissenschaft* **30**(2), 223–257.
- Halle, Morris and Alec Marantz (1993): Distributed Morphology and the pieces of inflection. In: K. Hale and S. J. Keyser, eds, *The view from building 20. Essays in linguistics in honour of Sylvain Bromberger*. MIT Press, Cambridge, pp. 111–176.
- Hendriks, Lotte, Sjef Barbiers and Hans Bennis (2015): Mapping the linguistic system. Paper presented at the Edisyn conference on June 11, 2015, Zurich.
- Hinterhölzl, Roland (2006): *Scrambling, remnant movement, and restructuring in West Germanic*. Oxford University Press, Oxford; New York.
- Hinterhölzl, Roland (2009): ‘The IPP-Effect, Phrasal Affixes and Repair Strategies in the Syntax-Morphology Interface’, *Linguistische Berichte* **2009**(218), 191–215.
- Hodler, Werner (1969): *Berndeutsche Syntax*. Francke Verlag, Bern.
- Höhle, Tilman N. (2006): Observing non-finite verbs: some 3v phenomena in German-Dutch. In: P. Brandt and E. Fuß, eds, *Form, Structure, and Grammar. A Festschrift Presented to Günther Grewendorf on Occasion of His 60th Birthday*. Akademie Verlag, Berlin, pp. 55–77.
- Kayne, Richard S. (1994): *The antisymmetry of syntax*. MIT Press, Cambridge, Mass.
- Koopman, Hilda and Anna Szabolcsi (2000): *Verbal complexes*. MIT Press, Cambridge, Mass.
- Kroch, Anthony and Beatrice Santorini (1991): The derived constituent structure of the West-Germanic verb-raising construction. In: R. Freidin, ed., *Principles and parameters in comparative grammar*. MIT Press, Cambridge, pp. 269–338.
- Lötscher, Andreas (1978): ‘Zur Verbstellung im Zürichdeutschen und in anderen Varianten des Deutschen’, *Zeitschrift für Dialektologie und Linguistik* **45**, 1–29.
- Louden, Mark L. (2011): Synchrony and diachrony of verb clusters in Pennsylvania Dutch. In: M. T. Putnam, ed., *Studies on German-language islands*. John Benjamins, Amsterdam, pp. 165–185.
- Merkes, Peter Wilhelm (1895): Der neuhochdeutsche Infinitiv als Teil einer umschriebenen Zeitform: Historisch-grammatische Betrachtungen. Phd dissertation, University of Göttingen.
- Meurers, Walt Detmar (2000): Lexical generalizations in the syntax of German non-finite constructions. Phd dissertation, Universität Tübingen.
- Müller, Gereon (1995): *A-bar syntax: a study in movement types*. Mouton de Gruyter, Berlin.
- Neeleman, Ad (1990): The Third Construction as a classical case of head-movement. Ms. Utrecht University.
- Reis, Marga (1979): Ansätze zu einer realistischen Grammatik. In: K. Grubmüller, E. Hellgardt, H. Jellissen and M. Reis, eds, *Befund und Bedeutung. Zum Verhältnis von Empirie und Interpretation in Sprach- und Literaturwissenschaft*. Niemeyer, Tübingen, pp. 1–21.
- Reis, Marga and Wolfgang Sternefeld (2004): ‘Review article: Susanne Wurmbrand: Infinitives. Restructuring and clause structure’, *Linguistics: An Interdisciplinary Journal of the Language Sciences* **42**(2), 469–508.

- Sauerland, Uli and Paul Elbourne (2002): 'Total Reconstruction, PF Movement, and Derivational Order', *Linguistic Inquiry* **33**(2), 283–319.
- Schallert, Oliver (2012): Untersuchungen zu Ersatzinfinitivkonstruktionen in den Vorarlberger und Liechtensteiner Dialekten. Doctoral dissertation, University of Marburg.
- Schallert, Oliver (2014): *Zur Syntax der Ersatzinfinitivkonstruktion: Typologie und Variation*. Stauffenburg, Tübingen.
- Schmid, Tanja (2005): *Infinitival syntax. Infinitivus pro participio as a repair strategy*. Benjamins, Amsterdam.
- Schmid, Tanja and Ralf Vogel (2004): 'Dialectal Variation in German 3-Verb Clusters: A Surface-Oriented Optimality Theoretic Account', *The Journal of Comparative Germanic Linguistics* **7**(3), 235–274.
- Schwalm, Johanna (2013): Verbalsyntax in den Mundarten Hessens: eine Untersuchung anhand der SyHD-Materialien. Master's thesis, Marburg University.
- Seiler, Guido (2004): On Three Types of Dialect Variation and Their Implications for Linguistic Theory: Evidence from Verb Clusters in Swiss German Dialects. In: B. Kortmann, ed., *Dialectology meets Typology. Dialect Grammar from a Cross-Linguistic Perspective*. Mouton de Gruyter, Berlin, Germany, pp. 367–399.
- Sternefeld, Wolfgang (1990): Scrambling and minimality. In: G. Grewendorf and W. Sternefeld, eds, *Scrambling and barriers*. John Benjamins, Amsterdam, pp. 239–257.
- Sternefeld, Wolfgang (2006): *Syntax: eine morphologisch motivierte generative Beschreibung des Deutschen*. Stauffenburg Verlag, Tübingen.
- ter Beek, Janneke (2008): *Restructuring and Infinitival Complements in Dutch*. LOT, Utrecht.
- Vogel, Ralf (2009): 'Skandal im Verbkomplex. Betrachtungen zur scheinbar inkorrekten Morphologie in infiniten Verbkomplexen des Deutschen', *Zeitschrift für Sprachwissenschaft* **28**(2), 307–346.
- von Stechow, Arnim (1990): Status government and coherence in German. In: G. Grewendorf and W. Sternefeld, eds, *Scrambling and barriers*. John Benjamins, Amsterdam, pp. 143–198.
- Weber, Albert (1987): *Zürichdeutsche Grammatik: ein Wegweiser zur guten Mundart*. Schweizer Spiegel Verlag, Zürich.
- Weise, Oskar (1900): *Syntax der altenburger Mundart*. Breitkopf u. Härtel, Leipzig.
- Wöllstein-Leisten, Angelika (2001): *Die Syntax der dritten Konstruktion: eine repräsentationelle Analyse zur Monosententialität von "zu"-Infinitiven im Deutschen*. Stauffenburg-Verl., Tübingen.
- Wurmbrand, Susanne (2001): *Infinitives: Restructuring and clause structure*. Mouton de Gruyter, Berlin.
- Wurmbrand, Susanne (2004): West Germanic verb clusters: The empirical domain. In: K. É. Kiss and H. v. Riemsdijk, eds, *Verb clusters: A study of Hungarian, German, and Dutch*. John Benjamins, Amsterdam, pp. 43–85.
- Wurmbrand, Susanne (2005): Verb clusters, verb raising, and restructuring. In: M. Everaert and H. v. Riemsdijk, eds, *The Blackwell companion to syntax. Volume 5*. Blackwell, Oxford, pp. 227–341.
- Wurmbrand, Susanne (2007): 'How Complex Are Complex Predicates?', *Syntax* **10**(3), 243–288.

- Wurmbrand, Susanne (2012): 'Parasitic participles in Germanic: Evidence for the theory of verb clusters', *Taal en Tongval: Tijdschrift voor Taalvariatie* **64**(1), 129–156.
- Wurmbrand, Susanne (to appear): Verb clusters, verb raising, and restructuring. In: M. Everaert and H. v. Riemsdijk, eds, *The Blackwell companion to syntax. 2nd edition*. Wiley-Blackwell, Oxford.
- Wurmbrand, Susi (2004): 'Two types of restructuring—Lexical vs. functional', *Lingua* **114**(8), 991–1014.
- Wyngaerd, Guido vanden (1989): Verb Projection Raising and the status of infinitival complements. In: D. Jaspers, Y. Putseys, W. Klooster and P. Seuren, eds, *Sentential Complementation and the Lexicon. Studies in Honour of Wim de Geest*. Linguistic Models, Foris, Dordrecht, pp. 423–438.
- Zwart, Jan-Wouter (1993): Dutch syntax. A Minimalist approach. Phd dissertation, University of Groningen.
- Zwart, Jan-Wouter (1995): A Note on Verb Clusters in the Stellingwerf Dialect. In: M. den Dikken and K. Hengeveld, eds, *Linguistics in the Netherlands 1995*. Linguistics in the Netherlands, John Benjamins, Amsterdam, pp. 215–226.
- Zwart, Jan-Wouter (1996): Verb clusters in continental West Germanic dialects. In: J. R. Black and V. Motapanyane, eds, *Microparametric syntax and dialect variation*. John Benjamins, Amsterdam, pp. 229–258.
- Zwart, Jan-Wouter (2007): 'Some notes on the origin and distribution of the IPP-effect', *Groninger Arbeiten zur Germanistischen Linguistik* **45**, 77–99.