Journal of Comparative Germanic Linguistics manuscript No.
(will be inserted by the editor)

On the limits of variation in Continental West-Germanic verb clusters

Evidence from VP-stranding, extraposition and displaced morphology for the existence of clusters with 213 order

Martin Salzmann

Received: date / Accepted: date

Abstract Recent work on verb clusters within Continental West-Germanic has argued in favor of restrictive theories of cluster orders 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., stranding of VP3 under topicalization of the governing VP2, short relative clause extraposition and displaced *zu*, I will show that verb clusters and Verb Projection Raising differ from the 3rd Construction with respect to the structural position of the dependent VP: While the dependent VP is contained within the projection of the governing verb in the former, it occurs outside of the projection of the governing verb in the latter. Applying the diagnostics to the Swiss German 213 construction delivers a clear result: The construction patterns with verb clusters rather than the 3rd Construction. I conclude from this that theories of verb clusters and unmarked word order more generally must be able to generate all six logically possible orders, including the 213 order.

Keywords Verb clusters · Verb Projection Raising · 3rd Construction · West-Germanic · Swiss German · Displaced *zu* · Adjunction · Extraposition · VP-Topicalization · Remnant Movement · Post-syntactic Morphology · Local Dislocation · Haplology · Word Order · Variation

Martin Salzmann
Department of Linguistics, University of Leipzig, Beethovenstraße 15, 04107 Leipzig, Germany
E-mail: martin.salzmann@uni-leipzig.de

1 Introduction: verb cluster orders

One prominent feature of Continental 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 examples (under verb second, where the finite verb moves to C, only the non-finite verbs occur together):[1]

(1) a. dass es jeder t_{es} gelesen$_4$ haben$_3$ zu müssen$_2$ glaubt$_1$
*that it everyone read.*PTCP *have.*INF *to must.*INF *believe.3*SG
'that everyone believes he must have read it' *Standard German*

b. dass das Buch keiner $t_{dasB.}$ zu lesen$_3$ versprochen$_2$ hat$_1$
*that the book no.one to read.*INF *promise.*PTCP *have.3*SG
'that no one promised to read the book' *Standard German*

Such sequences of verbs are referred to as verb clusters (for a detailed overview, cf. Wurmbrand 2017). Verb clusters involve the combination of a lexical verb with one or several functional verbs (auxiliaries, modals) and/or with one or several lexical verbs that embed non-finite complements.

There are three important properties of verb clusters that clearly set them apart from a mere sequence of verbs as it can be found in English: First, verb clusters display restructuring effects, second, the dependents of the lexical verb are often found outside the verb cluster, and third, they show massive cross-linguistic/dialectal and even intra-speaker variation with respect to the possible orders of the verbal elements. Since this paper will focus on word order variation, I will be rather brief concerning the first two aspects.

As for restructuring effects: Despite their sometimes bi- or multi-clausal semantics (cf. the translations of (1)), verb clusters behave like a monoclausal unit for syntactic operations, i.e., they show so-called clause-union or 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 (Bech 1983), verbs taking a *zu*-infinitive form different classes (Reis and Sternefeld 2004): Some are obligatorily restructuring (e.g., *scheinen* 'seem' in Standard German), some are optionally restructuring (e.g., *versuchen* 'try' in Standard German), while others do not allow 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] In (1),

[1] The glosses follow the Leipzig glossing rules, available at https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf. The numerical indices on the verbs indicate the hierarchical relations, i.e. 1 stands for the hierarchically highest verb of the cluster, 2 for the immediately dependent verb, etc.

[2] These 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

a pronominal/a DP-argument of the hierarchically lowest verb precedes the matrix subject, showing that it has left the projection of its predicate.

While restructuring effects can be found outside of West-Germanic (cf., e.g., clitic climbing in Romance), the transparency of the lexical VP in verb cluster languages is more general in that in many varieties its dependents can occur outside of the cluster without reordering with constituents of the embedding VP and thus without inducing the information structural markedness characteristic of scrambling. This is only visible in those varieties that have orders other than 321 and 312. The following pair from Swiss German shows that an object of the lexical verb can occur both inside and outside the cluster:[3]

(2) a. dass de Hans en Arie muess$_1$ singe$_2$
*that the John an aria must.3*SG *sing.*INF
b. dass de Hans muess$_1$ en Arie singe$_2$
*that the John must.3*SG *an aria sing.*INF
'that John must sing an aria' *Swiss German*

Importantly, both orders are acceptable under a wide-focus interpretation; unlike in scrambling that leads to reordering, the object in (2a) thus need not be given. 'Clusters' like (2b) containing non-verbal material are referred to as instances of Verb Projection Raising (VPR). VPR is only possible in certain orders (viz. 1X2X3, 1X32, and 2X31, where X indicates the position of non-verbal material). Furthermore, different varieties/dialects differ in the extent to which they allow VPR (cf., e.g., Barbiers et al. 2008, Dros-Hendriks 2018, chapter 5 on Dutch varieties). Since VPR obligatorily induces a wide-focus interpretation (Lötscher 1978, 6), non-VPR orders like (2a), which allow more information structural interpretations, will frequently be chosen. Consequently, the verbal elements will often occur together, hence the term 'cluster'.

I now turn to variation in the cluster orders. In 2-verb clusters, the two possible orders are instantiated both across West-Germanic as well as within a single variety (e.g. in Standard Dutch). In 3-verb clusters, there are six logically possible orders. Focusing on clusters with functional verbs (modals/auxiliaries) as V1 and V2 for the moment, the existence of the orders 123, 132, 321, 312 and 231 is undisputed (while largely absent in German varieties, 231 occurs frequently in West-Flemish and Afrikaans, cf. Wurmbrand 2004c, Barbiers 2005, Biberauer 2013). The following examples illustrate the various orders:[4]

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 since they also work if the transparent XP contains more structure, as is often the case in Verb Projection Raising, i.e. if the verbal complex contains non-verbal material, cf. below. In such configurations, some clause-union phenomena such as the lack of an independent tense or negation domain may be missing, while scrambling or pronoun fronting is still possible.

[3] Swiss G. examples without explicit references are from Zurich German, the author's native language. To my knowledge, the facts in this paper hold for all Swiss German dialects.

[4] Defenders of the Final-over-final-constraint, cf. Biberauer et al. (2014, 203–205), however, have treated 231 orders as illusory. In brief, they argue that either the verbs are not part of the same extended projection so that they are irrelevant for FOFC (West Flemish 231) or involve a complex verb consisting of V2 and V3 (Afrikaans 231).

(3) a. dat Hans het boek heeft$_1$ willen$_2$ lezen$_3$
*that John the book have.3*SG *want.*INF *read.*INF
'that John wanted to read the book' 123 *Standard Dutch*

b. dass Hans das Buch hat$_1$ lesen$_3$ wollen$_2$
*that John the book have.3*SG *read.*INF *want.*INF
'that John wanted to read the book' 132 *Standard German*

c. dass Peter das Buch gelesen$_3$ haben$_2$ muss$_1$
*'that Peter the book read.*PTCP *have.*INF *must.3*SG
'that Peter must have read the book' 321 *Standard German*

d. dass Hans das Buch lesen$_3$ hat$_1$ wollen$_2$
*that John the book read.*INF *have.3*SG *want.*INF
'that John wanted to read the book' 312 *Colloquial German*

e. dat hy die medisyne kon$_2$ drink$_3$ het$_1$
*that he the medicine could.*INF *drink.*INF *have.*PRS
'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 a subset of them. Furthermore, the availability of a certain cluster order often depends on the cluster type, i.e. the type of restructuring 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. What seems to characterize all West-Germanic verb cluster languages (and crucially sets them apart from restructuring languages like Italian), though, is that the word order pattern in the cluster is not completely rigid, i.e. there is at least one cluster order that deviates from the underlying order of the language. Concretely, while these languages/varieties are generally taken to be OV- and thus head-final languages, all of them allow for head-initial verb clusters to some extent. Standard German, for instance, mainly features strictly descending verb cluster, i.e. clusters where the governed verb precedes the governing verb. But in Aux-Mod-Inf clusters, the 132 order is mandatory.

Crucially, the 213 order has been claimed to be absent in the major cluster types involving functional verbs, see, e.g., Zwart (1996), Seiler (2004), Wurmbrand (2004c), Barbiers (2005) and Abels (2016). 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 Brandner et al. 2016). Similarly, the 213 order is (nearly) unattested in earlier stages of verb cluster languages; Sapp (2011) reports the absence of examples with 213 order in his Middle High German and Early New High German corpora. Haeberli and Pintzuk (2012, 225) report two Mod-Aux-Ptcp clusters with 213 order in their Old English corpus but generally take the order to be nonexistent.[5]

[5] There are a few claims to the contrary: Cooper (1995, 154) reports the 213 order 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

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

(4) a. wil er en ghöört$_2$ hät$_1$ choo$_3$
*because he him hear.*PTCP *have.3*SG *come.*INF
'because he heard him come'

b. wo ich scho Stars gseh$_2$ han$_1$ stah$_3$
*where I already stars see.*PTCP *have.1*SG *stand.*INF
'... where I have seen stars stand.' http://schnalletagebuech.blogspot.fr/2013/09/volleyball-hollywood.html, accessed March 21, 2016

c. Wo s aagfange$_2$ hät$_1$ rägne$_3$, simer i d Beiz
*when it start.*PTCP *have.3*SG *rain.*INF *are.*1PL *in the pub*
'When it started to rain, we went to the pub.'

d. wo i ufghört$_2$ ha$_1$ schaffe$_3$
*when I stop.*PTCP *have.1*SG *work.*INF
'after I stopped working.' www.babywelten.ch/community/forum/thema.html?TopicID=22 6&Page=2, accessed March 21, 2016

e. das er mer ghulffe$_2$ hät$_1$ abwäsche$_3$
*that he me.*DAT *help.*PTCP *have.3*SG *do.the.dishes.*INF
'that he helped me do the dishes'

f. bes er mer glehrt$_2$ het$_1$ skifahre$_3$
*until he me teach.*PTCP *have.3*SG *ski.*INF
'until he taught me how to ski'
http://www.heiraten.ch/forum/board/personenkreise/braut-braeutigamtalk/braeute-2013-12_250_3004.html, accessed October 19, 2017

Meran; Schwalm (2013, 65f., 69f., 81f., 86f.) reports 213 orders in Hessian Mod-Aux-Ptcp, Mod-Mod-Inf and Aux-Mod-Inf clusters; Schallert (2012, 285) reports 213 orders in Vorarlberg German Aux-Mod-Inf clusters; Heilmann (1999, 10) reports 213 orders in Swabian Aux-Mod-Inf clusters; finally, den Dikken (1994, 82f.) reports 213 orders 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/an unmarked order in any of these cases.

[6] V1 is normally the perfective auxiliary; examples with V1 being a modal are attested but less frequent (cf. Schallert 2012, 267 for a similar observation about Vorarlberg German):

(i) jetzt weiss i won i afange$_2$ muess$_1$ sueche$_3$
*now know.1*SG *I where I start.*INF *must.1*SG *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

Next to the benefactive meaning 'teach' as in (4f) below, the verb *leere* also has a more frequent intransitive use with the meaning 'learn', which also allows for the 213 order:

(ii) d Nacht, wo de Kobi glehrt$_2$ hät$_1$ flüüge$_3$
*the night, when the Jacob learn.*PTCP *have.3*SG *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 by about as many as the 123 order). Note also that 213 is an unmarked order in this cluster type; it alternates with the equally unmarked 123 order. According to Lötscher (1978, 3, fn. 2), the cluster type also allows the 231 order, but this clearly constitutes a marked order. 231 orders do not seem to occur elsewhere in Swiss German, except in clusters with motion verbs, see Salzmann (2013). The 213 order with these six verbs thus needs to be distinguished from the residual instances of 213 listed in fn. 5 that can be found in other cluster types, where they constitute apparently possible but marked orders. The cluster type in (4) is thus the only one where 213 represents the default order/an unmarked order.

The special status of the V2s occurring in the 213 order is not a peculiarity of Swiss German but has been noted in other varieties as well: Zwart (1995) reports 213 orders with perception verbs in the dialect of Stellingwerf; according to Schmid (2005, 64–72), the 213 order is found with inchoatives and control verbs (without *te*) in Afrikaans; Heilmann (1999, 62) reports 213 orders in Swabian with 'help' and 'begin', Zwart (2007, 80) reports 213 orders with 'learn' in Luxemburgish;[7] Wild (1991, 36) reports 213 orders with 'go' in the Fulda dialects of Hungary; finally, Louden (2011, 169, 175–179) reports 213 orders with perception verbs and benefactives in Pennsylvania German (as well as with V2 being causative '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 be shown to represent an obligatory restructuring configuration: While arguments of V3 can occur both outside the cluster as well as within VP3 (due to the possibility of VPR), cf. (5a), pronoun fronting is obligatory, cf. (5b):[8]

(5) a. dass er {✓d Chuchi} aagfange$_2$ hät$_1$ {✓d Chuchi}
*that he the kitchen start.*PTCP *have.3*SG *the kitchen*
butze$_3$
*clean.*INF
'that he started to clean the kitchen'

[7] Zwart (2007) discusses data from Samatimeric and Altai Plautdietsch with 1 ... 23 order in main clauses where there is insufficient data to determine whether these are 231 or 213 clusters underlyingly. The fact that the clusters feature a participle as V2 suggests that they are based on a 213 order where the participle form on V2 predominates; see also section 6.1 on the possibility of IPP in 213 orders.

[8] Pronoun fronting in Swiss German requires some care: Like other dialects, Zurich 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). The clitic versions seem rather unacceptable in non-finite complements, even if the matrix verb is non-restructuring (arguably since there is no proper host). The weak pronouns, however, can remain inside non-restructuring infinitival complements and will therefore be used in what follows.

b. dass er {✓si} aagfange$_2$ hät$_1$ {*si} butze$_3$
*that he it start.*PTCP *have.3*SG *it clean.*INF
'that he started to clean it' *Swiss German*

This is unsurprising given that verbs taking bare infinitival complements are usually regarded as restructuring verbs (recall that they are obligatorily restructuring in Standard German).[9]

I will argue in this paper that the existence of these 213 orders has important ramifications for the theory of verb clusters and theories of word order more generally as they imply that such theories must be able 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 compare the Swiss German 213-construction with another construction featuring a 213 order, viz., the 3rd Construction. In section three, I will introduce new diagnostics to distinguish between verb clusters proper/Verb Projection Raising and the 3rd Construction, viz., stranding of VP3 under topicalization of VP2, short relative clause extraposition and displaced morphology. I will then apply these diagnostics to the Swiss German 213 construction and show that it behaves like a proper verb cluster rather than the 3rd Construction. In section 4, I discuss the implications for the theory of verb clusters. Section five concludes. The appendix in section six briefly discusses diagnostics from the previous literature to distinguish verb clusters/VPR from the 3rd Construction that are inconclusive in my view.

[9] For some speakers, the phasal verbs and the benefactives can also be used with a *z*-infinitive, which entails optional restructuring. This can be seen in the fact that the pronoun can remain within VP3:

(i) dass er {si} aagfange$_2$ hät$_1$ {si} z butze$_3$
*that he it start.*PTCP *have.3*SG *it to clean.*INF
'that he started to clean it' *Swiss German*

Once *a z* is present, they thus behave like their Standard German counterparts. Note that in the standard language, the phasal verbs require a *zu*-infinitive while the benefactives occur both with and without *zu*. With *zu*, they are optionally restructuring; without *zu*, they are obligatorily restructuring (like bare infinitives more generally): This is shown by the fact that they cannot be extraposed, cf. (iia vs. b) and that pronoun fronting is obligatory, cf. (iic):

(ii) a. dass er ihr das Buch (zu) lesen half
*that he her the book to read.*INF *help.*PST.*3*SG
b. dass er ihr half das Buch *(zu) lesen
*that he her help.*PST.*3sg the book to read.*INF
'that he helped her read the book'
c. es ??(zu) lesen half er ihr nicht
*it to read.*INF *help.*PST.*3*SG *he her not*
'He didn't help her read the book.'

This crucially shows that the presence/absence of *z(u)* is not just a superficial matter of PF-realization but indicates a fundamental syntactic difference, viz., the presence or absence of syntactic structure. I am grateful to a reviewer for requesting clarification of this point.

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

Before concluding that the 213 order exists in Continental West-Germanic verb clusters, an alternative explanation for 213 orders must be taken into account: Unmarked 213 orders also occur in the so-called 3rd Construction.

2.1 Another construction with an unmarked 213 order

The 3rd Construction is a coherent construction where the restructuring predicate takes a *zu*-infinitive (rather than a bare infinitive as in the Swiss German construction introduced above) and where part of the non-finite clause follows the governing verb. At first sight it looks like an instance of VPR since in both constructions, the verbs occur in ascending order, there can be non-verbal material within the dependent VP and there are restructuring effects (scrambling in (6) indicates coherence):[10]

(6) a. dass er em Hans wett$_1$ t$_{em\ Hans}$ d Uhr stäle$_2$
*that he the.*DAT *John want.3*SG *the watch steal.*INF
'that he wants to steal John's watch' VPR, *SwG*

b. dass er em Hans versuecht$_1$ t$_{em\ Hans}$ debii z hälffe$_2$
*that he the John try.3*SG *there.with to help.*INF
'that he tries to help John with it' 3rdC, *SwG*

However, the 3rd Construction differs from clusters with functional verbs in that it can occur in the 213 order and in that V2 appears as a participle. Crucially, it thus seems more similar to the Swiss German 213 construction, which in addition to coherence and non-verbal material within VP3 also features a participle on V2. V2 also happens to have more lexical content than functional verbs. The pair in (7) iuxtaposes the two constructions (again, scrambling in (7a/b) indicates coherence):[11]

(7) a. wo s mer s Gschier ghulffe$_2$ händ$_1$ **gründlich**
*when they me the dishes help.*PTCP *have.3*PL *thoroughly*
t$_{s\ Gschier}$ abwäsche$_3$
*wash.*INF
'when they helped me wash the dishes thoroughly' *Swiss G.*

b. dass er em Hans versuecht$_2$ hät$_1$ t$_{em\ Hans}$ **debii** z
*that he the.*DAT *John tried has there.with to*
hälffe$_3$
*help.*INF
'that he tried to help John with it' 3rdC *Swiss German*

[10] To provide minimal pairs, I will usually use Swiss German data for the 3rd Construction. Unless noted otherwise, the 3rd Construction in Standard German behaves the same. Swiss German *zu*-infinitives are actually structurally ambiguous, see 3.3.2 and 3.4.3.1 below.

[11] In the Swiss German 213 construction V2 sometimes appears in the infinitive, instantiating the IPP-effect. See section 6.1 for discussion.

Given these similarities, it may be possible to unify the Swiss German 213 construction with the 3rd Construction.[12] 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.

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

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), and 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 construction represents a verb cluster or is an instance of the 3rd Construction. While some (e.g., Haegeman and van Riemsdijk 1986 and Geilfuß-Wolfgang 1991) have argued that the two constructions should be distinguished, others have claimed that the basic structure is the same, see, e.g., Vanden 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 section 6.2 of the appendix. The IPP-effect, which is often regarded as a diagnostic for proper verb clusters, is similarly inconclusive, see section 6.1 of the appendix. In the next section, I will introduce new structural diagnostics showing that verb clusters/VPR should indeed be distinguished from the 3rd Construction.

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

In this section, I will provide several new diagnostics that systematically differentiate between verb clusters and VPR on the one hand and the 3rd Construction on the other: stranding of VP3 under topicalization of the governing VP2, short relative clause extraposition and displaced *zu*. These diagnostics

[12] Zwart (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 (and thus should arguably be grouped with the 3rd Construction). Cf. also Kroch and Santorini (1991, 321) for a similar idea in a very different framework.

will show that in verb clusters and VPR the dependent VP has a tight relationship with the governing verb, while in the 3rd Construction, the relationship is looser. I will argue in section 4.2 below that this reflects a fundamental structural asymmetry: The dependent VP occupies a low position in the former, viz., is contained in the projection of the governing verb, but in a high position in the latter, viz., is not contained in the projection of the governing verb.

3.1 Stranding of VP3 under topicalization of VP2

A first asymmetry emerges under topicalization of VP2: In the 3rd Construction, VP3 can be stranded, cf. (8a), while this is not possible with VPR and verb clusters, cf. (8b/c) (pronoun fronting ensures we are dealing with coherent constructions; I will remain agnostic about possible PRO-subjects/traces of raised subjects in non-finite VPs; cf. Wurmbrand 2001, ch. 4 for discussion):

(8) a. [$_{VP2}$ versuecht$_2$] hät$_1$ er si scho, [$_{VP3}$ siine Eltere t$_{si}$ vorzstele$_3$]
try.PTCP *have.3*SG *he her indeed* *his*.DAT *parents* *introduce*.TO.INF
'He indeed tried to introduce her to his parents.' 3rdC

b. *[$_{VP2}$ wele$_2$] hät$_1$ er si scho [$_{VP3}$ siine Eltere t$_{si}$ vorstele$_3$]
want.INF *have.3*SG *he her indeed* *his*.DAT *parents* *introduce*.INF
'He indeed wanted to introduce her to his parents.' VPR

c. *[$_{VP2}$ wele$_2$] hät$_1$ er si siine Eltere scho [$_{VP3}$ t$_{siine\ Eltere}$ t$_{si}$ vorstele$_3$].
want.INF *have.3sg he her his*.DAT *parents indeed* *introduce*.INF
'He indeed wanted to introduce her to his parents.' VC

That the ungrammaticality of (8b/c) cannot be due to a problem with topicalization of VP2 as such is shown by the fact that topicalization of VP2 together with VP3 is generally possible, cf. (9) (the indirect object could also occur in the middle field like the weak pronoun if it scrambles out of VP3 before topicalization of VP2):[13]

[13] Standard German patterns the same with respect to topicalization of VP2. Incoherent infinitives behave like the 3rd Construction w.r.t. this diagnostic, cf. Müller (2002, 44–46). It has been claimed that ascending clusters cannot be topicalized, cf. Haider (2003, 108) on Dutch. An informal inquiry among Dutch linguists very much suggests that this is not correct; in fact, most speakers I have consulted strongly prefer the ascending order under topicalization (whether accompanied by an object or not). The same holds for Swiss German. Of course, (remnant) VP-topicalization requires appropriate contextualization to sound natural (and will thus often appear very marked for many speakers when presented in isolation). In Swiss German stronger deviance seems to result if a VPR structure is topicalized. This may be related to the fact that it induces a wide focus interpretation, which is arguably incompatible with the contrastive topic interpretation of the examples in the text.

(9) [$_{VP3}$ Siine Eltere wele$_2$ t$_{si}$ vorstele$_3$] hät$_1$ er si scho.
his.DAT *parents want*.INF *introduce.to*.INF *have.3*SG *he her* *indeed*

Taken together, the facts suggest that V2 and V(P)3 form a closer unit in verb clusters/VPR than in the 3rd Construction.

3.2 Relative clause extraposition

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

(10) a. dass Maria zu einem Mann sagte, [den sie nicht kannte], dass er ihr gefalle
that Mary to a man say.PST.*3*SG *who she not* *know*.PST.*3*SG *that he her*.DAT *please*.SUBJ.*3*SG
'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 Umzug zu helfen
that she a.DAT *man promise*.PST.3SG *who she not* *know*.PST.*3*SG *he*.DAT *at the 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) claims that both orders are possible and that preferences for one or the other order depend on heaviness. I will not take a stand on the behavior of CP-argument clauses but will instead focus on coherent infinitives:

As the following Swiss German pair shows, we again find an asymmetry in the tightness of the relationship between governing verb and the dependent VP: While in the 3rd Construction, the extraposed RC can both precede and follow the infinitival complement, it has to follow it in the VPR-variant:[14]

(11) a. dass mich jede versuecht$_1$, {✓ wo debii isch}, t$_{mich}$ devo z überzüüge$_2$, {✓ wo debii isch}
*that me everyone try.3*SG *who present be.3*SG *of.it to convince*.INF *who present be.3*SG
'that everyone who is present tries to convince me of it' 3rdC

[14] Again, pronoun fronting ensures coherence. Note that the verb cluster variant of (11b) with the RC only preceding the V2 is ungrammatical as well; however, it is arguably unacceptable for independent reasons because the extraposed infinitive is not heavy enough; a variant of (11a) with only the nonfinite V2 following the extraposed RC is equally unacceptable. The acceptability of short extraposition increases with the heaviness of the nonfinite VP (e.g. if 'convince' takes a CP-complement itself) in (11a) but not in (11b).

b. dass mich jede wett$_1$, {✗ wo debii isch}, t$_{mich}$
*that me everyone want.3*SG *C present be.3*SG
devoo überzüüge$_2$, {✓ wo debii isch}
*of.it convince.*INF *C present is*
'that everyone who is present wants to convince me of it' VPR
Swiss German

Descriptively, the data show that the relationship between governing verb and dependent VP is tighter in VPR than in the 3rd Construction in that it can be interrupted by RC-extraposition only in the latter.

3.3 The phenomenon of displaced *zu*

In German varieties, the placement of non-finite morphology crucially depends on the order in the verb cluster. While the non-finite morphology occurs on the expected verb if the cluster order is strictly descending, it is displaced in clusters that deviate from (3)21 (cf. Salzmann 2019a, this issue for details). In the following triple with different cluster orders, a selector outside the verb cluster selects a *zu*-infinitive (*ohne* in (12a/b), *freiheit* in (12c)).[15]

(12) 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' *Std. German* 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' *Std. German* 132
c. d freiheit, selber de tag chöne$_1$ **z** bestimme$_2$.
*the freedom self the day can.*INF *to determine.*INF
'the freedom to determine one's schedule' *Swiss German*12
cf. http://badoo.com/de-ch/0279246484/, accessed March 11, 2013

In (12a), which involves a 321-order, the *zu*-infinitive appears on the hierarchically highest verb of the cluster, viz., V1, thus on the verb that is immediately dependent on the *zu*-selector, as we would expect given the standard assumptions about morphological selection. In (12b/c), however, which involve a 132 and 12 order, respectively, *zu* does not occur on V1 but rather on V2. It thus seems to be displaced. As a descriptive generalization, *zu* always attaches to the last verb of the complement of the *zu*-selector. In a configuration as in (12) with the *zu*-selector outside the cluster, displacement occurs once V1 is not cluster-final (i.e. in all orders except 321 and 231). Importantly, if *zu* occurs on V1 in (12b/c), the result is sharply ungrammatical (unlike in Dutch).

[15] Note that while V2 appears as a participle in (12a), it appears as an infinitive in (12b/c), instantiating the IPP-effect; see section 6.1 below and Salzmann (2019a, this issue).

3.3.1 Displacement as a diagnostic for verb clusters

Displacement can be used as yet another diagnostic to distinguish between verb clusters/VPR and the 3rd Construction. As in the previous phenomena, the dependent non-finite VP forms a tighter unit with the governing verb in the first two constructions: The examples in (12) have shown that in bona fide verb clusters, *zu* occurs on the last verb of the complement of the *zu*-selector. *Zu* is thus displaced if V1 is not the last verb of the selector's complement. Displacement is equally obligatory with Verb Projection Raising, cf. (13):

(13) ohni mi **(*z)** welle$_1$ t$_{mi}$ uf d bullesite **z** stelle$_2$
*without me to want.*INF *on the cops.side to put.*INF
'without wanting to side with the cops' 1X2 *Swiss German*
http://www.fcbforum.ch/forum/showthread.php?4328-usschritige-nachem-spiel-!/page4; accessed March 11, 2013

In the 3rd Construction, however, there is no displacement. Rather, the *z(u)* selected by the matrix selector (*es bringt/gliich sii* 'be of avail/would you mind' in (14)) appears on V1; since V1 also selects a *z(u)*-infinitive, there is another *z(u)* on V2 (pronoun fronting/scrambling ensures coherence):[16]

(14) a. Aber es bringt nichts ihn$_1$ ***(zu)** versuchen$_1$ t$_{ihn}$ zu
*but it bring.3*SG *nothing him to try.*INF *to*
manipulieren$_2$
*manipulate.*INF
'But there is no point to try to manipulate him' *Std./Coll. G.*
https://www.elitepartner.de/forum/frage/kontakt-abgebrochen-und-dann-doch-interesse-bin-etwas-verwirrt.15022/, accessed April 10, 2017

b. wärs der glich, mer nomou schäu e
*be.*SBJV.*3*SG *you.*DAT *equal me.*DAT *again quickly a*
pn z probiere$_1$ z t$_{pn}$ schicke$_2$?
*personal.message to try.*INF *to send.*INF
'Would you mind to try again quickly to send me a personal message?' *Swiss German*
www.heiraten.ch/forum/board/verschiedene-themen/meine-traumhochzeit/wir-haben-uns-getraut-unser-trau-m-wochenende-15_221_2.html, July 20, 2017

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

(15) a. *ohni gläube, [$_{CP}$ dass de Peter **z** chunnt]
*without believe.*INF *that the Peter to come.3*SG

b. ohne **z** gläube, [$_{CP}$ dass de Peter chunnt]
*without to believe.*INF *that the Peter come.3*SG
'without believing that Peter will come' *Swiss German*

[16] Some speakers of Standard German also seem to allow for displacement; in Swiss German, displacement with *zu*-infinitive selecting verbs is optional, cf. sections 3.3.2 and 4.2.3.2.

To summarize, the facts from *zu*-placement show again that the dependent verb phrase in verb clusters and VPR forms a tighter unit with the governing verb: *Zu* occurs at the end of the entire cluster in orders where V1 is not final. In the 3rd Construction (and with CP-complements), *zu* separates the VP of the governing verb from the dependent infinitive.

3.3.2 Another diagnostic for verbclusterhood: missing z

In the so-called missing-*z* construction in Swiss German there are 2 *z(u)*-selectors in ascending order ('seem' and 'try' in (16)). Interestingly, unlike in the 3rd Construction in (14), there is only one *z*, namely on V3, while V2 appears in the bare infinitive, cf. Bader (1995, 22, 26):

(16) wüu dr Hans sine Fründe schiint$_1$ probiere$_2$ **z** häuffe$_3$
*because the John his.*DAT *friends seem.3*SG *try.*INF *to help.*INF
'because John seems to try to help his friends' *Bernese German*

Importantly, missing-*z* is optional, i.e. speakers generally accept versions with one (cf. (16)) or two *z*s (cf. (14b)). Missing *z* also seems to be available for some speakers of (spoken) Standard German (including one of the anonymous reviewers). One can find examples like (17) on the internet:

(17) ... ohne sie$_1$ versuchen$_1$ t$_{sie}$ zu therapieren$_2$
*without them try.*INF *to treat.*INF
'without trying to treat them'
https://erzaehlmirnix.wordpress.com/2015/10/07/die-realitaet-akzeptieren/, accessed April 10, 2017

The variant with two *zu*s is clearly more frequent, though, and many speakers of Standard German (including one of the anonymous reviewers) strongly reject examples like (17). Perhaps there is an underlying dialectal difference that affects the standard language of these speakers and is therefore responsible for the variation.

The missing-*z* construction shows that Swiss German *zu*-infinitives in ascending order can optionally behave like verb clusters in that *z* occurs at the end of the entire infinitival construction, thus revealing a tight relationship between governing verb and dependent VP. Crucially, the absence of *z* implies a structural difference (cf. fn. 9): The construction is obligatorily restructuring; weak pronouns have to leave the lexical VP, (18b), while they can stay there in the presence of a *z* on V2, which entails optional restructuring, cf. (18a):

(18) a. wil er {✓si} schiint$_1$ z probiere$_2$ {✓si} z küsse$_3$
b. wil er {✓si} schiint$_1$ probiere$_2$ {✗si} z küssse$_3$
*because he her.*ACC *seem.3*SG *try.*INF *her.*ACC *to kiss.*INF
'because he seems to try to help him' *Swiss German*

3.4 Applying the diagnostics to Swiss German 213 orders

This section has shown so far that there is a systematic asymmetry between verb clusters/VPR and the 3rd Construction in that governing verb and dependent verb form a tighter unit in the former. I will now apply these diagnostics to the Swiss German 213 construction.

3.4.1 Stranding of VP3

We saw in section 3.1 above that stranding of VP3 under topicalization of VP2 is only possible in the 3rd Construction but not with verb clusters/VPR. I repeat the example with the 3rd Construction from above:

(19) [$_{\text{VP2}}$ versuecht$_2$] hät$_1$ er si scho, [$_{\text{VP3}}$ siine Eltere t$_{si}$ vorzstele$_3$]
try.PTCP *have.3*SG *he her indeed* *his*.DAT *parents* *introduce*.TO.INF
'He indeed tried to introduce her to his parents.' 3rdC

(19) is crucially relevant for the structure of 213 orders because it is in all likelihood based on a 213 order. It could theoretically also be based on a 321 or a 123 order. However, two facts speak against this: First, these orders are somewhat marginal with V2 selecting a *zu*-infinitive in Swiss German. Second, as we will see in section 4.2.1 below, they do not have the right structure to allow for stranding of VP3 (VP3 would first have to be extraposed). Crucially, stranding VP3 under topicalization of VP2 is not possible in the Swiss German 213 construction (again, pronoun fronting ensures coherence; a VR-variant of (20) with the object higher up in the middle field es equally ungrammatical):

(20) *[$_{\text{VP2}}$ ghöört$_2$] hät$_1$ er si scho [$_{\text{VP3}}$ t$_{si}$ es Lied singe$_3$]
hear.PTCP *have.3*SG *he her indeed* *a song sing*.INF
'He should have indeed heard her sing a song.' *Swiss German*

Importantly, (20) could in principle be derived from a 123, 213, 231 or 321 order (these are the possible orders with this cluster type; since all orders are compatible with a participle on V2, one cannot be sure). This implies that none of these orders allows for stranding of VP3; thus, the 213 order patterns like the other orders. As with the VR/VPR data in (9) above, the ungrammaticality of (20) is not due to topicalization of VP2 as such. Topicalizing VP2 together with VP3 is unproblematic (some speakers may prefer the IPP-form on V2):

(21) [$_{\text{VP2}}$ t$_{si}$ es Lied ghöört$_2$ singe$_3$] hät$_1$ er si scho t$_{V1}$ t$_{VP2}$
a song hear.PTCP *sing*.INF *have.3*SG *he her indeed*
'He indeed heard her sing a song.' *Swiss German*

I thus conclude that the Swiss German 213 construction patterns like regular verb clusters/VPR w.r.t. stranding of VP3 under topicalization of VP2.[17]

3.4.2 Relative clause extraposition

Recall that short relative clause extraposition was only possible in the 3rd Construction, while in verb clusters/VPR the lexical VP had to occur immediately after the governing verb. The 3rd Construction patterns as expected, short RC extraposition is possible in the 3rd Construction when it occurs in the 213 order:

(22) dass mich jede versuecht$_2$ hät$_1$, {✓ wo debii gsii
*that me everyone try.*PTCP *have.3*SG *C present be.*PTCP
isch}, t$_{mich}$ devoo z überzüüge$_3$, {✓ wo debii gsii isch}
*be.3*SG *of.it to convince.*INF *C present be.*PTCP *be.3*SG
'that everyone who was present tried to convince me of it' 3rdC

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 (as in 3.2 above, a verb cluster variant of (23) with only V3 following the relative clause is ungrammatical as well):

(23) dass si s eme Maa ghulffe$_2$ hät$_1$, {✗ wo si guet
*that she it a.*DAT *man help.*PTCP *have.3*SG *C she well*
kännt}, t$_s$ in Ornig bringe$_3$, {✓ wo si guet kännt}
*know.3*SG *in order bring.*INF *C she well know.3*SG
'that she helped a man who she knows well to bring it in order' 213

Again, the facts show that the Swiss German 213-construction behaves like a proper verb cluster/VPR rather than the 3rd Construction: VP3 forms a tight unit with the governing verb(s).[18]

[17] Recall from fn. 9 that the phasal and the benefactive verbs can (for some speakers) optionally occur with *z* as well. In that case they pattern with the 3rd Construction in allowing stranding of VP3:

(i) [$_{VP2}$ Ghulffe$_2$] hät$_1$ er mer scho [$_{VP3}$ miini Wonig *(z) butze$_3$]
*help.*PTCP *have.3*SG *he me.*DAT *indeed my apartment to clean.*INF
'He indeed helped me clean my apartment.' *Swiss German*

[18] Again, for those speakers who can use the phasal and the benefactive verbs with a *z*, the construction seems to behave like its Standard German equivalent and thus like the 3rd Construction in that short extraposition is possible:

(i) dass en jede aagfange$_2$ hät$_1$, wo debii gsii isch, t$_{en}$ demit
*that him everyone begin.*PTCP *have.3*SG *who present be.*PTCP *be.3*SG *it.with*
*(z) konfrontiere$_3$
*to confront.*INF
'that everyone who was present began to confront him with it' *Swiss German*

This shows again that the presence/absence of *z* is indicative of a difference in syntactic structure.

3.4.3 Displaced zu

Recall that *zu*-displacement and missing-*z* are characteristic of verb clusters/VPR but not of the 3rd Construction, where we find two *zu*s. I will first apply displaced *zu* to 213 orders with V2-selecting a *zu*-infinitive before addressing the Swiss German 213 construction.

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

In the Standard German 3rd Construction in the 213 order, the result is as expected: There is no displacement (scrambling ensures coherence):

(24) ohne Hans versucht$_2$ ***(zu)** haben$_1$ t$_{Hans}$ zu helfen$_3$
*without John.*DAT *try.*PTCP *to have.*INF *to help.*INF
'without having tried to help John' 3rC *Standard German*

Things are different in Swiss German: Recall from section 3.3.2 that complements of *z(u)*-selecting verbs can either occur in the missing-*z* construction with just one *z* or in the 3rd Construction variant with 2 *z*s. Crucially, if a *zu*-selecting verb occurs in the 213 order, embedded under another *zu*-selector, both variants are possible as well (the absence of *z* on V3 is ungrammatical). Thus, 213 orders with *z(u)*-selecting verbs can optionally behave like proper verb clusters/VPR in that they allow for the missing-*z* construction.

(25) a. ohni s versuecht$_2$ **z** ha$_1$ **z** läse$_3$ 3rdC
b. ohni s versuecht$_2$ ha$_1$ **z** läse$_3$ missing-*z*
c. *ohni s versuecht$_2$ **z** ha$_1$ läse$_3$
*without it try.*PTCP *to have.*INF *read.*INF
'without having tried to read it' *Swiss German*

3.4.3.2 z-placement in 213 clusters with V2 selecting a bare infinitive

Turning to the Swiss German 213 construction, before tackling three-verb clusters in 213 order, it is instructive to first look at 2-verb clusters with V1 belonging to the class of predicates that occur in the 213 order. Here, the result is very clear: There is displacement of *z*, showing that the dependent VP forms the tight unit with the governing verb that we have identified as characteristic of verb clusters (since some speakers optionally allow a *z*-infinitive with 'help', 'learn', 'stop' and 'begin, (26c–f) may be instances of missing *z*):

(26) a. soooo schön, di wieder mal ghöre$_1$ **z** singe$_2$
*so nice you again once hear.*INF *to sing.*INF
'so nice to hear you sing again' www.facebook.com/video/video.php?v=10200666450684322, accessed March 28, 2016
b. für di isch es gnue gsi mi gseh$_1$ **z** schiine$_2$
*for you be.3*SG *it enough be.*PTCP *me see.*INF *to shine.*INF
'for you it was sufficient to see me shine' http://www.songtexte.com/songtext/sandee/wind-womi-treit-73f5ea71.html, accessed April 6, 2017

c. ... mithälfe, e Teil vo de neue Bronceglogge hälfe$_1$ **z'** finanziere$_2$.
contribute.INF *a part of the new bronce bell help*.INF *to finance*.INF
'to contribute to help finance part of the new bronce bell'
https://www.google.ch/#psj=1&q=%22h%C3%A4lfe+z+%22&start=20, accessed March 28, 2016

d. Es esch cool gsi bi der lehre$_1$ **z** fahre$_2$.
*it be.3*SG *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

e. ich liieb es mit wildfrämde lüt afange$_1$ **z'** rede$_2$!
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, March 28, 2016

f. Wie bringt mer en Schlagzüger dezue, ufhöre$_1$ **z** spiele$_2$
*how bring.3*SG *one a drummer it.to stop*.INF *to play*.INF
'How can you make a drummer stop playing?'
http://www.vjmn.ch/files/Lagerbericht-Dienstag.pdf, accessed April 6, 2017

In the 213 order, the result is the same: There is displacement:

(27) a. *ohni en ghört$_2$ **z** ha$_1$ singe$_3$ non-cluster
b. ohni en ghört$_2$ ha$_1$ **z** singe$_3$ cluster
without him heard have.INF *to sing*.INF
'without having heard him sing' *Swiss German*

These facts have been verified with native speakers. An example from the internet is given in (28) (note again that since some speakers can use *aafange* 'begin' with a *z*-infinitive, this could also be an instance of missing *z*):

(28) Wieder en grund meh zum glücklich drüber sii, niä agfange$_2$ ha$_1$ **z** rauche$_3$!
again a reason more to happy about.it be.INF *never begin*.PTCP *have*.INF *to smoke*.INF
'Another reason to be happy to have never started smoking!'
www.facebook.com/Radio24/posts/10151574652070814, March 28, 2016

Thus, the Swiss German 213 construction crucially behaves like a proper verb cluster/VPR with respect to *z(u)*-displacement. While the six verbs in Swiss German that occur in the 213 order select a bare infinitive, they require a *zu*-infinitive in the standard language when they occur in the 213 order (including 'help' and 'learn', where *zu* is optional if the lexical VP is intraposed, cf. fn. 9). They are thus expected to pattern like the 3rd Construction in (24) above and indeed, in the Standard German equivalent of (28), *zu* before V1 is obligatory.

4 Implications for the theory of verb clusters

I will now investigate the implications of the empirical findings of the previous section for the theory of verb clusters. I will first provide a brief overview of the major theories of verb clusters and their predictions with respect to the possible orders and the structures they assign to them. Thereafter, I go through the three diagnostics introduced in the previous section and discuss to what extent the various theories can account for the empirical facts, especially for the cluster-nature of the Swiss German 213 orders. After discussing an alternative to capture the cluster-signature of the Swiss German 213 construction, I will, finally, briefly address questions pertaining to the rarity of 213 orders.

4.1 Approaches to word order in verb clusters

In this subsection, 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 which orders can be generated, the discussion below will be partly based on my interpretation of the mechanisms.

Note that I wish to remain neutral with respect to the question whether verb clusters are a genuine theoretical entity that requires a designated clustering mechanism (as, e.g., in Haider 2003) or whether they are just a pretheoretical notion, a descriptive term for a phenomenon whose properties follow from independently available general principles (as, e.g., in Wurmbrand 2004a, Koopman and Szabolcsi 2000, and Abels 2016). I will thus focus on the predictions the various approaches make w.r.t. the ordering possibilities and the constituency and leave a comparison of their other properties for another occasion (but see, e.g., Salzmann 2011, 2013 and Wurmbrand 2017 for some discussion). Even though I will frequently speak of 'theories of verb clusters', I will do so only for convenience and always intend all approaches that make predictions about the possible orders in verb clusters.

Importantly, this discussion rests on the assumption that verb clusters are the same phenomenon across Continental West-Germanic and that all varieties make use of essentially the same cluster-forming/-ordering mechanism. This is not a necessity as it is in principle conceivable that different varieties have different verb cluster grammars. However, all Continental West-Germanic varieties with verb clusters share certain important properties: They always show restructuring effects, the various orders do not affect the semantics and only certain 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/the same ordering principles 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 result from different ways of using a particular clustering/ordering mechanism (cf. also the discussion in section 4.4 below). 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.

4.1.1 Restrictive theories

A number of recently proposed theories are designed in a way that they cannot generate the 213 order while allowing the five other orders.[19]

Barbiers (2005) derives the five orders with feature-driven VP-movement from a right-branching base: 123 is thus base-generated, (29a). 132 involves movement of VP3 to SpecVP2, (29b). 321 involves VP3 to SpecVP2 movement followed by movement of VP2 (containing VP3) to SpecVP1, (29c). 312 involves successive-cyclic movement of VP3; it first moves to SpecVP2 and then moves on to SpecVP1, (29d). 231, finally, involves movement of VP2 containing VP3 to SpecVP1, (29e):

(29) a. 123: $[_{\text{VP1}}$ V1 $[_{\text{VP2}}$ V2 $[_{\text{VP3}}$ V3]]]
b. 132: $[_{\text{VP1}}$ V1 $[_{\text{VP2}}$ $[_{\text{VP3}}$ V3] V2 t_{VP3}]]
c. 321: $[_{\text{VP1}}$ $[_{\text{VP2}}$ $[_{\text{VP3}}$ V3] V2 $t_{\text{VP3}}]$ V1 t_{VP2}]
d. 312: $[_{\text{VP1}}$ $[_{\text{VP3}}$ V3] V1 $[_{\text{VP2}}$ t_{VP3} V2 t_{VP3}]]
e. 231: $[_{\text{VP1}}$ $[_{\text{VP2}}$ V2 $[_{\text{VP3}}$ V3]] V1 t_{VP2}]

213 is ruled out 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:

(30) $[_{\text{VP1}}$ $[_{\text{VP2}}$ $[_{\text{VP3}}$ V3$]_1$ V2 $[_{\text{VP3}}$ V3$]_1]_2$ V1 $[_{\text{VP2}}$ $[_{\text{VP3}}$ V3$]_1$ V2 $[_{\text{VP3}}$ V3$]_1]_2]$

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; the constraint is formulated such that the 312 order is ruled in).

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. It was first proposed in Wurmbrand (2004c), 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 (2016), who embeds his approach in a more general theory of neutral word order (building upon Cinque 2005

[19] I will not discuss the approach by Barbiers et al. (2018) and Dros-Hendriks (2018, ch. 3), 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 involve Verb Projection Raising. While this is a very radical and thus interesting proposal, I will not discuss it here because its focus is orthogonal to the question pursued in this paper, viz., the existence/non-existence of clusters with 213 order.

and Abels and Neeleman 2012). Under this approach, we obtain the following representations for the five different word orders:

(31) 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.[20]

4.1.2 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), neither the devices used by GEN to generate the six orders nor the structural properties of the resulting orders are specified so that I cannot provide any representations. In Bader and Schmid (2009), complex heads can be directly generated in different orders (by means of functional composition):

(32) 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.[21]

[20] These restrictive theories also make predictions about larger clusters, of course. Abels (2016), 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 do not discuss clusters consisting of four or more verbs in this paper because the structures become too complex to apply the diagnostics that I am using 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 in fact the most natural order (more so than 1234):

(i) dass er s Buech sött$_1$ aagfange$_3$ ha$_2$ läse$_4$
*that he the book should.3*SG *begin.*PTCP *have.*INF *read.*INF
'that he should have started to read the book' *Swiss German*

The unexpected orders 3214, 3241 and 3124 are also 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 Abels (2016) may constitute further counter-examples.

[21] Similar base-generation accounts are proposed by Bayer and Kornfilt (1994) and Sternefeld (2006, 619ff.) and can be found quite generally in HPSG, cf. Hinrichs and Nakazawa (1994). Unfortunately, none of them discuss 231 and 213 orders.

The theory proposed by Haider (2003) can probably also generate all six orders. In his system, the 321 order is base-generated as a complex head, while orders that deviate from it involve excorporation of a verb of the complex head and reprojecting movement (projection of an additional VP-shell). The starting point is a base-generated complex head in 321 order, (33a). A 132 order is then derived by reprojection of V1, (33b). To derive the 123 order, both V2 and V1 have to reproject, (33c). 231 orders are an intermediate step of a 123 derivation, i.e., they only involve reprojection of V2, (33d). 213 involves excorporation of V1 followed by excorporation of V2, (33e). 312 involves excorporation of V1 followed by excorporation of V3, (33f):[22]

(33) a. $[_{VP}\ [_{V^\circ}\ [_{V^\circ}\ [_{V^\circ}$ V3] V2] V1]] →
b. $[_{VP}$ V1 $[_{VP}\ [_{V^\circ}\ [_{V^\circ}\ [_{V^\circ}$ V3] V2] t_{V1}]]]
c. $[_{VP}$ V1 $[_{VP}$ V2 $[_{VP}\ [_{V^\circ}\ [_{V^\circ}\ [_{V^\circ}$ V3] t_{V2}] t_{V1}]]]]
d. $[_{VP}$ V2 $[_{VP}\ [_{V^\circ}\ [_{V^\circ}\ [_{V^\circ}$ V3] t_{V2}] V1]]]
e. $[_{VP}$ V2 $[_{VP}$ V1 $[_{VP}\ [_{V^\circ}\ [_{V^\circ}\ [_{V^\circ}$ V3] t_{V2}] t_{V1}]]]]
f. $[_{VP}$ V3 $[_{VP}$ V1 $[_{VP}\ [_{V^\circ}\ [_{V^\circ}\ [_{V^\circ}$ t_{V3}] V2] t_{V1}]]]]

In Salzmann (2013), verb-cluster formation takes place at PF by means of Local Dislocation (Embick and Noyer 2001). Importantly, cluster formation may be accompanied by inversion, but it may also be string-vacuous. The order 123 is the base order that does not involve any cluster-formation operations. 132 involves cluster-formation (and inversion) between V3 and V2. 321 is based on 132 and involves cluster formation (and inversion) of [32] with V1. 213 can be generated quite easily as it involves cluster-formation (and inversion) between V2 and V1. 231 and 312 are somewhat different in that they involve string-vacuous cluster-formation: 231 orders first involve string-vacuous cluster-formation between V2 and V3 followed by cluster formation (and inversion) between [23] and V1. 312, finally, involves string-vacuous cluster-formation between V1 and V2 followed by cluster formation (and inversion) between [12] and V3. The resulting representations are given in (34):[23]

(34) 123: 1 2 3 132: 1[32] 321: [[32]1] 213: [21]3 231: [[23]1] 312: [3[12]]

[22] It is not made explicit whether the reprojection movements have to apply in a specific order and whether this allows for reordering. According to the description in Haider (2003, 105), reprojection starts with V1; subsequent reprojection of V2 then involves tucking in and derives the 123 order. If excorporation has to start with V1 and subsequent reprojection must tuck in, the orders 213 and 312 cannot be derived this way. Haider (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 (in German) or to the right (in Dutch), again starting from a base-generated complex head in 321 order. This option is arguably envisaged for the 312 order (as well as the 4132 and 4312 order) as this would correctly derive the compactness of this order (i.e., it does not tolerate any non-verbal interveners). Depending on the precise assumptions about reprojection, Haider's approach may thus eventually turn out to be restrictive rather than powerful.

[23] This mechanism can also generate most of the logically possible orders in 4-verb clusters, except for 3142 and 2413. Unlike the restrictive theories discussed in 4.1.1 above it can thus 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 the 213 order (recall fn. 20).

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 (arguably followed by reanalysis of the complex [V1V2] with V3). 213, finally, requires reanalysis of V2 and V1, followed by reanalysis and inversion of the complex [V2V1] with V3.[24]

Similarly powerful is the remnant movement approach of Koopman and Szabolcsi (2000). The 213 order can be generated if a constituent containing V3 is moved to a position between V1 and V2 (SpecFP1 in (35)) and the remnant VP2 is subsequently moved to a position above V1:

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

In fact the actual derivations in this work are much more complex; none of the orders is base-generated, even the 123 order involves massive XP-movement, cf. Koopman and Szabolcsi (2000, 62–72).[25]

The two types of approaches thus mainly differ with respect to the availability 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/theories of word order. Given that we have seen that the construction indeed behaves like a proper verb cluster, it clearly provides an argument for more powerful theories (an alternative account of 213 orders under restrictive theories will be discussed in 4.3.2 below). In the next subsection, I will provide an explicit syntactic analysis of the three diagnostics introduced above and discuss to what extent the different approaches to cluster orders can account for the properties of verb clusters in these configurations.

[24] One 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; furthermore, if there were, the 312 order probably couldn't be derived either. Thus, depending on the precise assumptions, the approach can arguably either derive all six logically possible orders or just four.

[25] Traditional 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-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].
Note that since the 213 order allows for VPR, cf. (5a), deriving 213 by head-movement from a left-branching base makes incorrect predictions in this respect.

The second issue that will play an important role is the more general question of what it takes to be a verb cluster, i.e., what verb clusters/VPR have in common and what sets them apart from the 3rd Construction. If we look at the representations provided in this subsection, most approaches assume that in all the possible orders of verb clusters all verbal elements/projections are contained within the projection of the highest verb of the cluster. The only exceptions are the reprojection approach by Haider (2003) and the remnant movement approach by Koopman and Szabolcsi (2000). In the former, there is, however, always containment within a higher VP, but this need not be VP1 (as in the 213, 231 and 312 orders). In the latter, there are certain orders, including 12(3), where V2 (and V3) are not contained within VP1 or any other verbal projection (cf., e.g., Koopman and Szabolcsi 2000, 158). We will see that containment within the projection of the hierarchically highest V of the cluster is indeed characteristic of verb clusters but not of the 3rd Construction, where the dependent VP is not contained within the projection of the governing verb (in 2-verb clusters)/VP3 is not contained within VP1 (in the 213 order). We will see that the data in fact motivate an even more narrow characterization of verb clusters: In verb clusters, the dependent verb/VP is contained within the projection of the governing verb. The 312 order will be crucial in this respect and provide an argument against those theories that involve syntactic movement of VP3 to a position outside of VP2.

4.2 Analysis: Verb clusters/VPR vs. 3rd Construction

The previous section has shown that there are systematic asymmetries between verb clusters/VPR and the 3rd Construction that descriptively can be captured in terms of tightness: The relationship between the governing verb and the dependent VP is very close in the former constructions such that they cannot be separated by topicalization, relative clause extraposition or *zu*. In this section, I will show that the pretheoretic notion of tightness translates into a systematic structural difference: In verb clusters, the dependent VP occurs in a structurally low position, more precisely, in a position within the projection of the structurally highest verb of the cluster; in fact, the topicalization/stranding data will very much suggest that it must be contained within the projection of the governing verb. In the 3rd Construction, however, it occurs in a higher position, a position that is crucially not contained in the projection of the governing verb. In what follows, I will show that the three diagnostics I have introduced are sensitive to this difference.

4.2.1 Stranding of VP3

I will start with the stranding data. I will first account for the difference between verb clusters/VPR and the 3rd Construction before discussing the implications for the analysis of the 213 order.

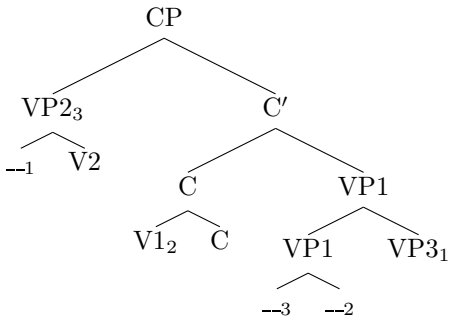
Recall that stranding of VP3 under topicalization of VP2 is only possible in the 3rd Construction but not in verb clusters/VPR. I repeat the the examples with the 3rd Construction/VPR from above:

(36) a. [$_{VP2}$ versuecht$_2$] hät$_1$ er si scho, [$_{VP3}$ siine Eltere
try.PTCP *have.3*SG *he her indeed* *his*.DAT *parents*
t$_{si}$ vorzstele$_3$]
introduce.TO.INF
'He indeed tried to introduce her to his parents.' 3rdC

b. *[$_{VP2}$ wele$_2$] hät$_1$ er si scho [$_{VP3}$ siine Eltere t$_{si}$
want.INF *have.3*SG *he her indeed* *his*.DAT *parents*
vorstele$_3$]
introduce.INF
'He indeed wanted to introduce her to his parents.' VPR

This asymmetry follows straightforwardly if the 3rd Construction involves extraposition of VP3. Topicalization of VP2 then instantiates remnant movement. In verb clusters/VPR, however, VP3 remains in-situ, i.e. within the projection of VP2 so that it cannot be stranded. The derivation of (36a) (based on a 213 order) is illustrated in (37):[26]

(37)



[26] I assume for concreteness' sake that the middle field only consists of a VP as proposed in Haider (2010), but nothing would change if there were a TP-projection.
I represent extraposition by means of rightward movement/adjunction mainly for ease of exposition (as in the classical remnant extraposition analysis of the 3rd Construction, cf. Broekhuis et al. 1995), but alternative implementations of extraposition would work as well: In anti-symmetric approaches such as Hinterhölzl (2006), there would be leftward movement of VP3 followed by remnant movement of VP2. If extraposition involves coordination + ellipsis as proposed by de Vries (2002, chapter 7), the 3rd Construction would arguably involve VP2-coordination with VP3 occurring only in the right conjunct [$_{\&P}$ [$_{VP2}$ V2] & [$_{VP2}$ ~~V2~~ [VP3]]]. However, since VP3 would arguably be a complement in both verb clusters/VPR and the 3rd Construction, it is not obvious how to restrict extraposition to the 3rd Construction. If extraposition is analyzed as a PF-operation, it would have to take place from the topicalized VP2 and prosodic conditions would then have to make sure that it ends up clause-finally. See also fn. 28 below. I have no new insights to offer as to what triggers extraposition.
Recall from section 3.4.3.1 that *z*-infinitives in Swiss German can instantiate two structures, one corresponding to the 3rd Construction and one behaving like a proper verb cluster; obviously, for the stranding data and the extraposition data below, the 3rd Construction variant is used since the other structure would not lead to a grammatical result.

That extraposition has taken place in (36b) is suggested by the fact that VP3 is clause-final and follows the modal particle.[27] An alternative to form a remnant would be scrambling of VP3 in which case VP3 would precede the particle. Extraposition of VP3 could also target VP2; in that case, it could either be stranded as in (37) (VP-adjuncts can be stranded by VP-movement) or be pied-piped under VP2-topicalization. Note that (36a) cannot be directly based on a 321 or 123 order since VP3 is contained within VP2 in both cases. Thus, VP3 would have to be extraposed first to evade pied-piping by VP2.

Crucially, with verb clusters/VPR there is no possibility for VP3 to leave VP2 and thus create a remnant: Extraposition and scrambling of bare infinitival VPs is generally taken to be impossible in German, cf. Müller (1995, 154). Consequently, when VP2 is topicalized, VP3 has to be pied-piped.

Importantly, the stranding data also provide more information about the structure of verb clusters. The fact that VP3 cannot be stranded follows if VP3 is contained within VP2 (whether as a complement or as a specifier). The 312 order is crucial in this respect because in many approaches (Barbiers 2005, Wurmbrand 2004a, Abels 2016, Haider 2003, Koopman and Szabolcsi 2000) V(P)3 undergoes movement to a position outside of VP2. Consequently, in those varieties that allow for the 312 order, stranding of VP3 should be grammatical. To the best of my knowledge, this is not the case. Stranding of VP3 is generally unavailable, irrespective of whether the cluster order in a given variety is 123, 132, 312, 231 or 321. Consequently, the stranding data argue against those analyses where V(P)3 is not contained within VP2. This includes the V(P)3-movement analyses of the 312 order just mentioned but also XP-movement analyses like those in Koopman and Szabolcsi (2000, 72) more generally where even in orders like 123 VP3 is not contained within VP2. Of course, there may be independent reasons why remnant movement of VP2 fails in these cases, but that requires an additional assumption that is not necessary in those approaches (e.g. Bader and Schmid 2009, Haegeman and van Riemsdijk 1986, Salzmann 2013), where VP3 is contained within VP2. I now turn to the Swiss German 213 construction.

As shown above, the Swiss German 213 construction behaves like a proper verb cluster because stranding of VP3 is impossible: (repeated from above):

(38) *[$_{\text{VP2}}$ ghöört$_2$] hät$_1$ er si scho [$_{\text{VP3}}$ t$_{si}$ es Lied singe$_3$]
hear.PTCP *have.3*SG *he her indeed* *a song sing*.INF
'He has indeed heard her sing a song.' *Swiss German*

Since (38) can be based on a 123, 213, 231 or a 321 order, we can infer that VP3 is contained within VP2 in all these orders. This is compatible with most powerful theories (Bader and Schmid 2009, Haider 2003, Haegeman and van Riemsdijk 1986, and Salzmann 2013) but proves problematic for the remnant

[27] Extraposition of the lexical VP is clearly visible in more complex clusters like (i), where it follows V2 in the right sentence bracket (the base order would be either 1324 or 3214):

(i) versuecht$_3$ sött$_1$ er si scho haa$_2$, siine Eltere vorzstele$_4$
try.PTCP *should.3*SG *he her indeed have*.INF *his* *parents introduce.to*.INF

movement based approach by Koopman and Szabolcsi (2000) because VP3 is not contained within VP2 as shown in the simplified representation in (39):

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

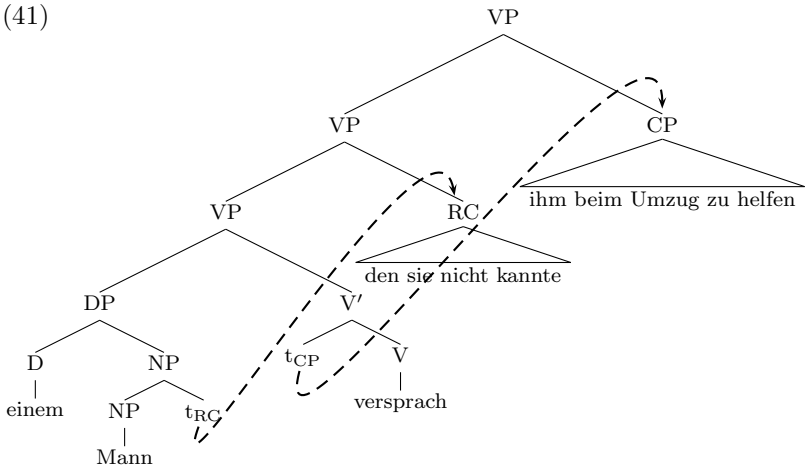
The implications of the stranding test are summarized in (40):

(40) a. stranding under topicalization → VP3 not contained in VP2
b. no stranding under topicalization → VP3 contained within VP2

4.2.2 Extraposition

Recall that when both a complement CP and a relative clause are extraposed, the RC can precede the complement CP. This crucially presupposes an attachment site between the main clause and the complement clause. Consequently, short RC extraposition is only possible if the complement clause is extraposed as well; the structure of (10b) above is thus as in (41) (the non-finite complement is taken to be incoherent and is therefore represented as a CP):

(41)



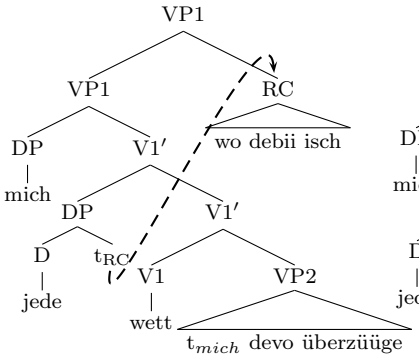
Since verb clusters/VPR and the 3rd Construction differ from each other in that the former involves a low position of the dependent VP (contained within the projection of the governing verb), while in the latter the dependent VP is in a higher position (not contained in the projection of the governing verb), we also expect an asymmetry with respect to RC-extraposition: Suppose RC-extraposition involves adjunction to the local VP; if extraposition of the dependent VP involves adjunction to the governing VP as well, the RC should be able to precede the non-finite complement clause in the 3rd Construction, but not in verb clusters/VPR where the dependent VP remains within the projection of the governing verb. As the data in 3.2 above have shown, this prediction is borne out. In the 3rd Construction, we find both short and long RC-extraposition, while in verb clusters/VPR only long RC-extraposition is possible. I repeat the relevant Swiss German examples for convenience:

(42) a. dass mich jede versuecht$_1$, {✓ wo debii isch}, t$_{mich}$
*that me everyone try.3*SG *who present be.3*SG
devo z überzüüge$_2$, {✓ wo debii isch}
*of.it to convince.*INF *who present be.3*SG
'that everyone who is present tries to convince me of it' 3rdC

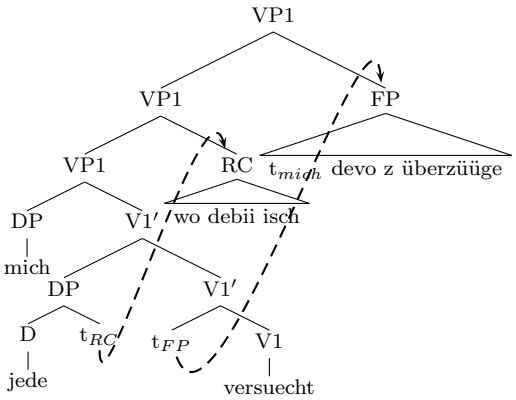
b. dass mich jede wett$_1$, {✗ wo debii isch}, t$_{mich}$ devo
*that me everyone want.3*SG *C present be.3*SG *of.it*
überzüüge$_2$, {✓ wo debii isch}
*convince.*INF *C present is*
'that everyone who is present wants to convince me of it' VPR

The structures of the two examples are given in (43) (I will henceforth represent bare infinitives as VPs and coherent *zu*-infinitives as FPs, cf. 4.2.3 below):

(43) a. VPR:



b. 3rd Construction:



Since the dependent VP is low, contained within VP1 in VPR, the RC will invariably follow it, irrespective of where it is adjoined (as long as extraposition obeys the extension condition). In the 3rd Construction, however, both the RC and FP are adjoined to VP1. Since nothing regulates their order, both options are possible.[28]

[28] The contrast between VPR and the 3rd Construction follows straightforwardly under a syntactic approach to extraposition since only the 3rd Construction provides an adjunction site. Under a PF-approach, the asymmetry would have to follow from prosodic asymmetries. I leave it to further research whether such asymmetries can be found.
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 (although that is unlikely in German as they generally reconstruct, cf. Sternefeld (2006, 774–779).

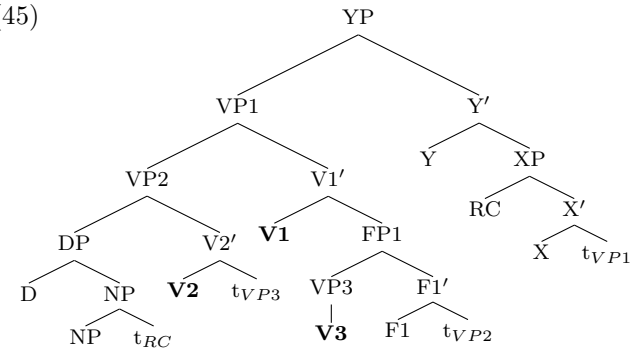
Note that the contrast can be derived most straightforwardly if the dependent VP occupies the complement position of the governing verb in proper verb clusters. A higher position within VP1 is conceivable as well, but then RC-extraposition must crucially involve adjunction to a maximal projection (or, under remnant movement-based approaches, movement to a position of appropriate height) to ensure that the RC ends up higher than the dependent VP (see also the discussion in 4.3.2 below). As discussed in the previous subsection, containment within the governing VP is assumed in most approaches to cluster orders, at least for ascending 12(3) clusters; the only exception is again the approach by Koopman and Szabolcsi (2000); consequently, special provisions are needed under that approach to rule out short extraposition in verb clusters (for instance, VP2 could move above VP1 followed by extraposition of an RC from VP1 to a spec above VP2; subsequent movement of (the XP containing) VP1 to an even higher spec would derive the undesired order).

I now turn to the extraposition facts in the Swiss German 213 construction. Recall from above that short RC extraposition is not possible; I repeat the relevant example from above:

(44) dass si s eme Maa ghulffe$_2$ hät$_1$, {✗ wo si guet
*that she it a.*DAT *man help.*PTCP *have.3*SG *C she well*
känt}, t$_s$ in Ornig bringe$_3$, {✓ wo si guet känt}
*know.3*SG *in order bring.*INF *C she well know.3*SG
'that she helped a man who she knows well to bring it in order' 213

Under the remnant movement approach by Koopman and Szabolcsi (2000) (44) would have roughly the structure in (45) (again strongly simplified):

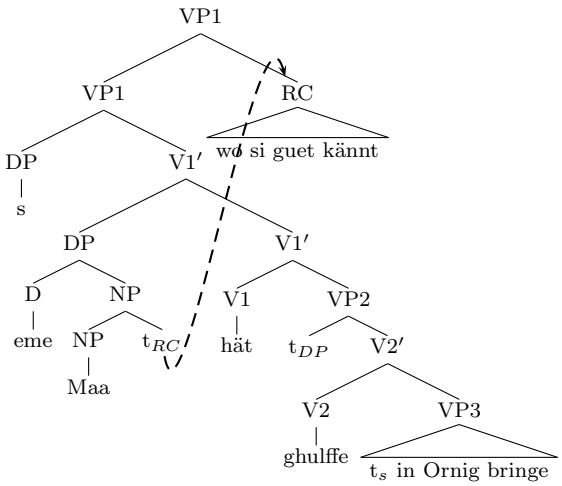
(45)



As just mentioned, special provisions are necessary under such an approach to ensure that short RC-extraposition is blocked. Concretely, one has to rule out the possibility that the RC moves to a specifier between FP1 and V1 in (45). There may be ways of ensuring that, but it requires extra assumptions which, as I will show presently, are not necessary under the other powerful theories:

Under the post-syntactic approach by Salzmann (2013), the structure of (44) looks as follows (since this is a syntactic representation, cluster formation and inversion between V1 and V2 has not yet taken place):

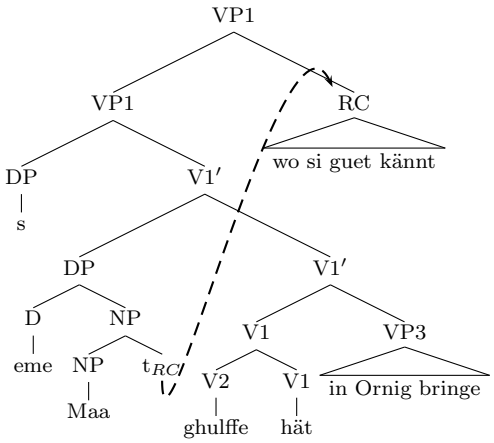
(46)



Since VP3 remains in the complement position of the governing verb, RC-extraposition will necessarily target a higher position. Under the reanalysis approach by Haegeman and van Riemsdijk (1986) and the approach by Haider (2003) the structure would arguably look quite similar.

In approaches adopting base-generated clusters like Bader and Schmid (2009), finally, (44) looks as in (47). Since VP3 is a complement of V1, the extraposed RC will invariably occur at the end of the cluster.

(47)



Thus, as with the stranding diagnostic, most powerful theories can account for the behavior of the Swiss German 213 construction. The approach by Koopman and Szabolcsi (2000) can do so only with extra assumptions. The implications of the short RC-extraposition diagnostic are summarized in (48):

(48) a. short RC extraposition → dependent VP outside of the projection of the governing verb
b. obligatory long RC extraposition → dependent VP contained within the projection of the governing verb

4.2.3 zu-placement

In this subsection, I will first explain why *zu*-displacement distinguishes between verb clusters/VPR and the 3rd Construction. Then, I show why the missing-*z* construction also bears the hallmarks of verb clusters. Finally, I address *zu*-placement in 213 orders.

4.2.3.1 Zu-displacement

Recall that displacement also systematically distinguishes between verb clusters/VPR and the 3rd Construction in that it is found only in the former. I repeat the relevant contrast from above:

(49) a. ohni mi **(*z)** welle$_1$ t$_{mi}$ uf d bullesite **z** stelle$_2$
*without me to want.*INF *on the cops.side to put.*INF
'without wanting to side with the cops' 1X2 *Swiss German*
http://www.fcbforum.ch/forum/showthread.php?4328-usschritige-nachem-spiel-!/page4; accessed March 11, 2013

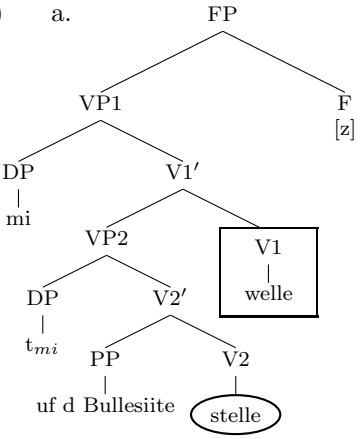
b. wärs der glich, mer nomou schäu e
*be.*SBJV.*3*SG *you.*DAT *equal me.*DAT *again quickly a*
pn z probiere$_1$ z t$_{pn}$ schicke$_2$?
*personal.message to try.*INF *to send.*INF
'Would you mind to try again quickly to send me a personal message?' *Swiss German*
www.heiraten.ch/forum/board/verschiedene-themen/meine-traumhochzeit/wir-haben-uns-getraut-unser-trau-m-wochenende-15_221_2.html, July 20, 2017

Before showing that this asymmetry also results from the nature of the structural relationship between governing verb and dependent VP, I will briefly address derivational and representational approaches to *zu*-placement. Both types of approaches share the idea that *zu*-placement depends on the surface order in the cluster; as we will see below, both approaches are compatible with different approaches to cluster orders although not all combinations will work; see Salzmann (2019a, this issue) for detailed discussion of displacement.

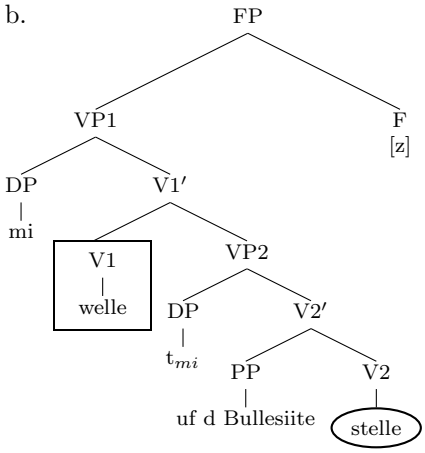
In a derivational approach as in Salzmann (2019a, this issue), *zu* is inserted into an independent syntactic head F above VP and and is associated with its host post-syntactically by means of Local Dislocation (cf. Embick and Noyer 2001). This functional head takes its VP-complement to the left, in accordance with the head-final nature of the German VP. This captures 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 V1 is not final in the cluster, *zu* will appear to be displaced, but this is just a side-effect. As for the structure of verb clusters, I will assume that the cluster order is mainly a question of linearization between head and complement (cf., e.g. Abels 2016) and that coherent verbal projections are VPs (cf., e.g., Wurmbrand 2007).

Turning to *zu*-placement in (49a), there are two possible ways of linearizing a two-verb cluster, either VP2-V1 or V1-VP2:[29]

(50) a.



b.



In Swiss German, Mod-Inf clusters can generally be linearized both ways. At vocabulary insertion, the hierarchical structures are converted into a linear string. *Zu* is inserted into F and thus occurs at the end of the cluster. Importantly, *zu* needs to attach to the left of a verbal host. By Local Dislocation it is affixed onto and inverted with the linearly adjacent verb (for morphological aspects, cf. Salzmann 2019a, this issue). Depending on how V1 and VP2 are ordered, this will target a different verb. In the descending 21 order (50a), *z* will target V1, leading to a 'well-behaved' case of *zu*-placement, schematically illustrated in (51a). In the VPR-example at hand, VP2 is ordered after V1 as in (50b). In that case, *z* attaches to V2, leading to displacement, cf. (51b).

(51) a. DP PP V_2 V_1 z ⇒ DP PP V_2 **z**+V_1
LD

b. DP V_1 PP V_2 z ⇒ DP V_1 PP **z**+V_2
LD

In the Standard German example in (12a), both VPs are ordered to the left of the governing verb; we thus obtain a 321 order and *zu* attaches to V1, cf. (52a). In examples like (12b), the linearization is mixed, leading to a 132 order so that *zu* ends up on V2, (52b).

(52) a. V3 V2 V1 zu ⇒ V3 V2 **zu**+V1
LD

b. V1 V3 V2 zu ⇒ V1 V3 **zu**+V2
LD

[29] I assume that verb clusters differ from VPR in that all non-verbal material is scrambled out of the lexical VP. For a base-generation alternative to scrambling, see e.g. Salzmann (2011). In the tree diagrams in the text, scrambled material is located in a specifier of V1, but scrambling to a specifier of F would be equally possible.

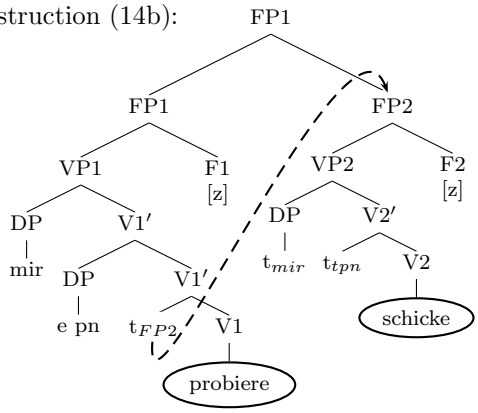
In representational approaches, illustrated on the basis of Vogel (2009), *zu* is a feature that is assigned by the governing verb to the entire infinitival complement; its realization is determined by realizational rules like (53):

(53) zu is realized on the right-most verb within the XP bearing [zu]

Given that the realizational rule applies to the surface order, both 'well-behaved' *zu* in descending orders as well as displacement in orders where V1 is not cluster-final can be derived straightforwardly: The feature is realized on the right-most verbal terminal of the relevant phrase.

Turning to the contrast between verb clusters/VPR and the 3rd Construction, what is crucial is that the dependent VP is contained within the projection of the governing verb in the former. If the dependent VP is ordered after the governing verb, the governed verb will end up adjacent to *zu*/will be the target for the realizational rule so that displacement results. The absence of displacement in the 3rd Construction follows under the by now familiar assumption that the dependent VP is not contained in the projection of the governing verb. Rather, it involves remnant extraposition; in (54), the dependent VP together with its functional structure is right-adjoined to FP1:

(54) 3rd Construction (14b):



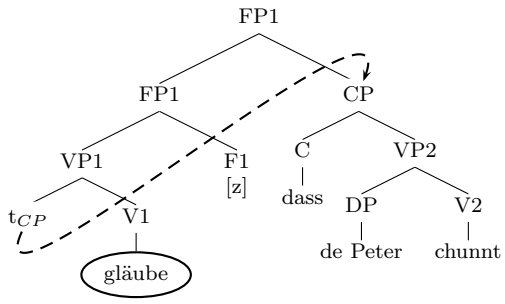
After linearization, both *z*s are adjacent to different verbs and Local Dislocation derives the correct result:[30]

(55) DP $\mathbf{z}_1$+V_1 $\mathbf{z}_2$+V_2
LD LD

Zu-placement with CP-arguments ((15)) involves extraposition (of CP to FP1):

[30] If extraposition occurs at PF, it will automatically apply before *zu*-placement because it is sensitive to hierarchical (prosodic) structure (as it moves a constituent), while *zu*-placement applies to linear structure. The ordering of the operations would thus be intrinsic.

(56)



Zu is linearized after the matrix verb and can thus felicitously undergo Local Dislocation; the finite CP does not interfere:[31]

(57) **z**+V_1 C DP V_2

Extraposition also leads to the correct result under the representational approach to *zu*-placement: since both the matrix and the dependent/adjoined VP bear a *zu*-feature, we obtain two *zu*s.

The structural implications of *zu*-(non-)displacment are given in (58):

(58) a. *zu*-displacement → dependent VP contained within the projection of the governing verb
b. no displacement → dependent VP not contained within the governing of the governing verb

4.2.3.2 Missing z

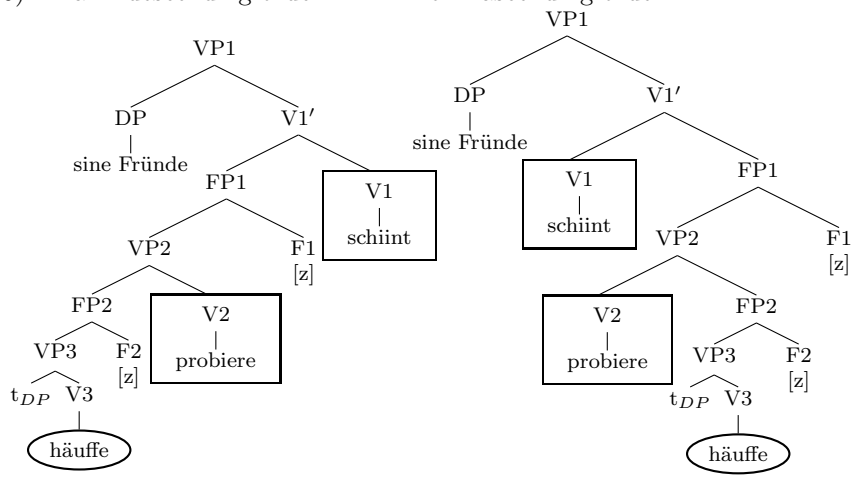
Recall that in Swiss German ascending orders involving two *zu*-infinitives, *zu* on the first dependent verb is optional. I repeat an example from above:

(59) wüu dr Hans sine Fründe schiint$_1$ **(z)** probiere$_2$ **z** häuffe$_3$
*because the John his.*DAT *friends seem.3*SG *to try.*INF *to help.*INF
'because John seems to try to help his friends' *Bernese German*

The version with two *z*s is simply an instance of the 3rd Construction with FP2 undergoing extraposition as in the derivation in (54) above. The missing-*z* variant can be derived by assuming that the dependent FPs need not be extraposed but can simply be ordered after the governing verb. Thus, next to the strictly descending order as in (60a) (which is less prominent than in Standard German), the strictly ascending order in (60b) is available as well:

[31] Note that the lack of displacement suggests that postverbal object CPs are not sisters of V at surface structure (pace Zwart 1993, Haider 2010). Whether finite (and non-finite non-restructuring) 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; see Bayer et al. (2005) for discussion.

(60) a. descending order: b. ascending order:



After linearization of (60a), there will be a *z* adjacent to each verb, thus deriving the unspectacular 'well-behaved' case. After linearization of (60b), however, both *z*s follow the verb cluster, cf. (61a). I propose that the two *z*s are reduced to one by morphological haplology: Given cyclicity, the *z* adjacent to V3 is attached first, cf. (61b). The second *z* is attached thereafter, cf. (61c) and finally deleted under identity with the *z* already attached to V3, cf. (61d):

(61) a. linearization: V1 V2 V3 z z
b. *z*-placement: V1 V2 **z**+V3 z
LD
c. *z*-placement: V1 V2 **z**+z+V3
LD
d. deletion under identity: V1 V2 ~~**z**~~+z+V3

Note that the haplology effect only obtains if FP2 remains low, viz., within the projection of the governing verb. This is the same structural relationship as in verb clusters. Therefore, missing-*z* can be considered another diagnostic for verbclusterhood (for discussion of the haplology effect in representational approaches, cf. Salzmann 2019a, this issue):

(62) missing *z* → dependent VP contained within the governing VP

4.2.3.3 zu-placement and the 213 order

Recall that the Swiss German 213 construction as well as Swiss German 213 orders involving two *zu*-infinitives behave like proper verb clusters/VPR because they instantiate *zu*-displacement/missing-*z* (examples repeated from above):

(63) a. ohni s versuecht$_2$ ha$_1$ **z** läse$_3$ missing-*z*
*without it try.*PTCP *have.*INF *to read.*INF
'without having tried to read it' *Swiss German*

b. Wieder en grund meh zum glücklich drüber sii, niä
*again a reason more to happy about.it be.*INF *never*
agfange$_2$ ha$_1$ **z** rauche$_3$!
*begin.*PTCP *have.*INF *to smoke.*INF
'Another reason to be happy to have never started smoking!'
www.facebook.com/Radio24/posts/10151574652070814, March 28, 2016

While all powerful theories are in principle compatible with *zu*-placement in 213 orders, different theories require different approaches to *zu*-placement. The derivational approach works for remnant movement based theories like Koopman and Szabolcsi (2000) (cf. also Hinterhölzl 2009): The constituent containing the three verbal elements (VP1 in (64)) moves into the specifier of the functional head into which *zu* is inserted, viz., F1 in (64) (as before, the derivation is strongly simplified; note that as this derivation shows, FP2 need not necessarily be contained within the immediately governing VP2, it is sufficient if it is contained within VP1). (64) is thus the derivation for (63b):

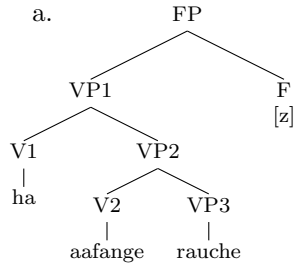
(64) $[_{\text{FP1}}\ [_{\text{VP1}}\ [_{\text{VP2}}\ \mathbf{V2}\ t_{VP3}]\ [_{\text{V1}'}\ \mathbf{V1}\ [_{\text{FP2}}\ \mathbf{VP3}\ [_{\text{F2}'}\ \text{F2}\ t_{VP2}\ ]]]]\ \text{F1}\ t_{VP1}]$

In the missing-*z* construction, the complement of V2 would be an FP3 hosting the lower *zu*, into whose specifier VP3 moves. The rest of the derivation then proceeds as in (64) so that we obtain the representation in (65) for (63a):

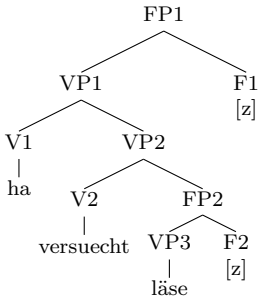
(65) $[_{\text{FP1}}\ [_{\text{VP1}}\ [_{\text{VP2}}\ \mathbf{V2}\ t_{FP3}]\ [_{\text{V1}'}\ \mathbf{V1}\ [_{\text{FP2}}\ [_{\text{FP3}}\ \mathbf{VP3}\ \text{F3}\ t_{VP3}\ ]\ [_{\text{F2}'}\ \text{F2}\ t_{VP2}\ ]]]]\ \text{F1}\ t_{VP1}]$

After linearization, the *zu*s in F1 and F3 will be adjacent (F2 is silent) and haplology derives the correct result.
In the post-syntactic approach by Salzmann (2013) the derivation of (63b) proceeds as follows: Given that verbal complements are ordered to the right of the verbal head, while the functional head F takes its VP-complement to the left, the input to the PF-derivation is as in (66a):

(66) a.



b.



After linearization of (66a), there is cluster formation between V2 and V1, leading to a [21]3 order, while *z* is affixed onto the adjacent V3:

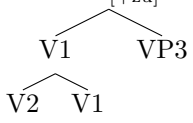
(67) $[\text{V}_2{+}\text{V}_1]$ $\boldsymbol{z}{+}\text{V}_3$ Local Dislocation
LD

In the missing-*z* construction in (63a), the input to the PF-derivation is as in (66b). After linearization, both *z*s will be adjacent to V3, and deletion under identity derives the correct result (Local Dislocation inverts V1 and V2).

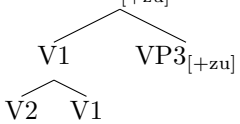
The derivational *zu*-placement mechanism can arguably also be combined with the reanalysis approach by Haegeman and van Riemsdijk (1986), provided that the late inserted morphology can interact with the reanalyzed structure; but given that this is where inversion takes place, this seems unproblematic.

The derivational approach to *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 as in (12) (although infixation of *zu* into the cluster may be an issue). However, if one of the verbs selects a *zu* as in the missing-*z*-construction (59), there simply is no space for that functional head inside the complex head. The representational approach, however, can be combined with 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): Taking the approach by Bader and Schmid (2009), it is conceivable that a head outside the cluster imposes a *zu*-feature on its VP1-complement that involves a 213 order as in (68a). The realizational rule in (53) will make sure that *zu* is realized on the last verb within VP1, irrespective of the cluster order and thus on V3 in (63b).

(68) a.



b.



In the missing-*z* construction, an outside head would impose a *zu*-feature on VP1; V2 would also impose a *zu*-feature on VP3 (the selectional requirements of V2 are inherited by V1), cf. (68b). The two *zu*s will thus both have to be realized on V3, leading to missing *z* as in (63a).

4.2.4 Intermediate summary

This subsection has provided two important results, one pertaining to the nature of verb clusters/VPR, one related to the 213 order. I have shown that verb clusters/VPR and the 3rd Construction differ from each other in a fundamental structural property: In 2-verb 'clusters', the dependent VP is contained within the projection of the governing verb in the case of verb clusters/VPR, while it is located outside the projection of the governing verb in the 3rd Construction. In 3-verb 'clusters', there is clear evidence that VP3 is contained within VP1 in verb clusters/VPR but not in the 3rd Construction. This characterization is compatible with most approaches to cluster orders, except for the remnant movement approach by Koopman and Szabolcsi (2000), where in several orders the dependent VPs are not contained within VP1. It is thus

the only one that does not provide a simple structural characterization of verb clusters. While this may only be a conceptual objection, the stranding data provide very clear evidence that VP3 is not only contained within VP1 but also within the projection of the governing V2. All structural descriptions of cluster orders where this is not the case thus make incorrect predictions for the stranding of VP3 under topicalization of the governing VP(2). Crucially, this not only affects the remnant movement approach but several other approaches (Barbiers 2005, Wurmbrand 2004a, Abels 2016, Haider 2003) where in the 312 order VP3 is not contained within VP2. In other words, unless remnant movement of VP2 can be ruled out by other plausible means under such approaches, the stranding data argue in favor of a more restrictive characterization of clusters such that the dependent VP must be contained within the projection of the governing verb, an assumption that is, e.g., made in Haegeman and van Riemsdijk (1986), Bader and Schmid (2009) and Salzmann (2013).[32]

Turning to the 213 order, we have seen that all powerful theories can provide appropriate descriptions for the three diagnostics. If we combine the two findings, we can conclude that the diagnostics introduced in this paper provide important insights into the nature of verb clusters that crucially limit the possible structural descriptions. Still, the results are compatible with several theories of verb clusters, which I take to be a positive result as it suggests that the diagnostics capture a rather general structural property. In the remaining subsections, I will briefly address an alternative account of 213 orders and their general rarity.

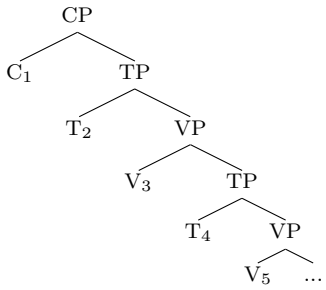
4.3 An alternative to derive 213 orders

An alternative perspective on 213 orders is proposed in Abels (2016). He capitalizes on the fact that 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'). Therefore, the Swiss German clusters might instantiate lexical restructuring rather than functional restructuring. At least under the perspective taken in Abels (2016) 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 (e.g. auxiliaries vs. adverbs vs. PPs). 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

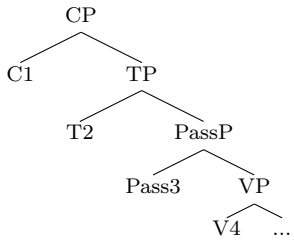
[32] In an approach based on functional composition like Bader and Schmid (2009), the lexical verb is indirectly contained in the projection of the governing verb in the orders 312 and 213 in that V1 inherits the selectional restrictions of V2.

thus two independent domains. Given the evidence that the two classes (lexical vs. functional restructuring verbs) indeed pattern differently in some respects (cf. Wurmbrand 2004b), 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 from Abels (2016, 196):[33]

(69) a.



b.



As Abels points out, while C1, T2 and Pass3 can straightforwardly be analyzed as modifiers of V4 in the representation of functional restructuring in (69b), C1, T2 and V3 cannot easily be treated as modifiers of V5 in lexical restructuring as in (69a). For this approach to be viable, two conditions need to be met: First, the classification of the six Swiss German verbs as lexical should be compatible with diagnostics that have been independently arrived at for the lexical status of restructuring verbs rather than just their compatibility with the 213 order. Second, there has to be an operation that generates the 213 order and derives the cluster behavior of these verbs, viz., an operation that is sufficiently different from classical (high) extraposition as in the 3rd Construction. I will discuss both points in turn.

4.3.1 On the lexical/functional nature of the six Swiss German verbs

Wurmbrand (2001, 2004b) discusses a number of semantic and syntactic criteria to separate lexical from functional restructuring verbs. However, when applied to the six Swiss German verbs, there is little evidence that they should be classified as lexical restructuring verbs. Their syntactic properties are clearly those of functional restructuring verbs: First, recall that pronoun fronting is obligatory (see (5b)), suggesting that the 213 construction involves obligatory restructuring. Second, these six verbs all display the IPP-effect (it is invisible in the case of 'see' because it does not have a separate form for the participle; furthermore, the IPP-effect seems optional for most of these verbs, see also

[33] Note that the lexical XP 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 then rather be located in the nature of the projection of the restructuring verb rather than in the size of its complement.
The difference between the two classes is captured in quite a different way in Wurmbrand (2004b, 992), where functional verbs embed a vP while lexical restructuring verbs only embed a VP.

section 6.1 below). Third, the verbs occurring in the 213 order all behave the same with respect to the verb cluster diagnostics discussed in section 3.[34]

As for the semantic properties, the six verbs do not form a homogenous group: Focusing on the thematic properties, the phasal verbs *aafange* 'begin' and *ufhöre* 'stop', which generally come in a functional restructuring (raising) and a lexical restructuring (control) version, do show the behavior of raising/functional verbs in the 213 order in that they are compatible with a weather-*it*. The other four verbs of the 213 class *gsee* 'see', *ghöre* 'hear', *hälffe* 'help' and *lehre* 'teach/learn' display the inverse pattern in that the subject has to be thematic (and they involve ECM or control). Similarly, unlike the phasal verbs, 'help' and 'teach' (and perhaps also the perception verbs, cf. fn. 39) take internal arguments, another hallmark of lexical restructuring verbs.

To summarize, while the six Swiss German verbs thus pattern the same with respect to the syntactic criteria, they do not with respect to the semantic criteria (thematic subjects and objects). Furthermore, while the syntactic properties are those of functional verbs, four verbs behave like lexical restructuring verbs and two like functional verbs w.r.t. the thematic criteria. This argues against the idea that the special behavior of the six verbs can be captured by referring to their 'lexical' nature. More generally, it casts doubts on the basic idea that functional and lexical restructuring verbs can be neatly separated into groups with homogenous syntactic and semantic behavior. As these six Swiss German verbs show, the properties do not always line up.[35]

[34] Note also that there is no perfect match between the possible cluster orders and the putative lexical status of a restructuring verb. While verbs like *probiere/versueche* 'try' that are undoubtedly lexical restructuring verbs can occur in 321 and 123 order next to the 213 order resulting from extraposition, the verbs that occur in the 213 construction additionally allow the orders 123 and, somewhat marginally, 231.

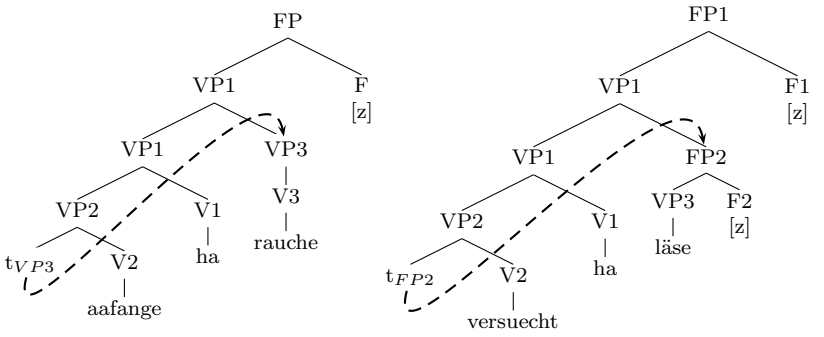
[35] For criticism of the lexical/functional divide, cf. Reis and Sternefeld (2004).
It is not completely surprising that some of the verbs occurring in the 213 order behave inconsistently w.r.t. the lexical/functional diagnostics. The special behavior of perception verbs has been observed before; because of their mixed behavior, they have been referred to as semi-functional/semi-lexical, cf., e.g., Wurmbrand (2001, 158, 165, 215–225). Wurmbrand (2001) argues that their behavior can be accounted for if they are treated as voice elements. Whether 'help' and 'teach/learn' can be subsumed under this class as well remains to be seen, though. Even if this should turn out to be possible, the class of verbs occurring in the 213 order would still involve elements that form distinct (more or less functional) classes according to the classification in Wurmbrand (2001, 206, 216).
There are two further potentially semi-functional verbs in Swiss German, viz., *laa* 'let' and *bliibe* 'stay'. Neither of them occurs in the 213 order, though. In the case of 'let' this may be related to the fact that it often occurs as a weak/clitic form, viz. *la*, which is not compatible with the prosodic requirements of the 213 order (secondary stress on V2); furthermore, since not all speakers have a separate participle form for 'let' but the V2 in the 213 order usually appears as a participle, the possibility of 213 is further reduced (but recall that the 213 order is found with 'let/cause' in Pennsylvania German, cf. Louden 2011, 175). I have no explanation why 'stay' does not occur in the 213 order (it only seems to occur in the 123, the 321 and perhaps the 132 order in Swiss German, cf. Schmid 2005, 57–62). I can only point out that in Standard German, where it generally occurs in the 321 order, it also neither patterns with modal verbs (in that the 132 order is marginal in Aux-durative-Inf clusters) nor with the semi-functional or the lexical restructuring verbs (which in addition occur in the 132 or the 213 order). Thus, for a descriptively adequate account of the word order possibilities in verb clusters, it seems necessary to be able to refer to single predicates.

Admittedly, the problem of capturing the class of verbs that allows for the 213 order is a challenge for every theory. Thus, the arguably more pressing question is whether there is – within more restrictive theories – an alternative syntactic mechanism to generate 213 orders with cluster properties. This is the topic of the next subsection.

4.3.2 Low extraposition to derive the 213 order

Such an alternative was proposed by one of the anonymous reviewers: The cluster-like behavior of the six verbs in question can be captured if they involve low extraposition of VP3, viz., to a position that is crucially below (a) the functional head F into which *zu* is inserted and (b) the extraposition site of relative clauses. In the case of *zu*-displacement, the structure could look as in (70a): Unlike in the 3rd Construction (cf. (54)), extraposition thus targets a position *below* FP. Such an analysis is in principle compatible with the restrictive approaches by Wurmbrand (2004a), Barbiers (2005) and Abels (2016), although given that these approaches rule out rightward movement, extraposition would have to be handled differently (e.g. by leftward movement followed by remnant movement). The analysis of missing *z* would be similar, instead of remaining a complement of V2 as in (66b) above, FP2 (containing the VP3) would be extraposed to a position between VP1 and F1, cf. (70b):

(70) a. Swiss G. 213: dispaced *zu* b. Swiss German 213 missing-*z*



In the case of RC-extraposition, low extraposition of VP3 has to target a position below the extraposition site of the RC. To derive the correct result, VP3 must crucially not target the same projection as the RC as this would wrongly predict optionality (especially if extraposition is adjunction). Obviously, this can only be ensured if there is at least one functional head above VP1 whose projection/specifier the RC can target (as e.g. in remnant movement-based approaches, cf. (45)).

What the low extraposition analysis crucially cannot account for is the asymmetry under VP-topicalization discussed in section 3.1 above: The lowest extraposition site for VP3 under the 213 order is right above VP1 as in (70a) above. VP3 is thus no longer contained within VP2 so that stranding

should be unproblematic, contrary to fact. Adjoining VP3 to VP2 would not help either since segments of VP can be stranded under VP-topicalization (cf. stranded VP-adverbs under English VP-topicalization: *read a book he did in the evening*).[36] There are also more general concerns one may raise against the low extraposition analysis: First, related to the implementation of extraposition, while nothing speaks against having both low and high extraposition, it may be technically tricky to ensure that extraposition has to be low with certain predicates (those in the 213 construction and the missing-*z* construction) but high with others (those in the 3rd Construction); furthermore, to implement the difference structurally, one needs more functional structure, i.e., VP and RC extraposition must crucially not target the same VP (cf. above). Second, since extraposition is widely available across West Germanic, one might expect 213 orders to be more wide-spread. Third, once low extraposition is an option, various cluster orders become ambiguous: Both the 123, the 231 and the 132 order can also be generated this way from a 321-base. Low extraposition is thus also an alternative to flexible linearization of sister nodes/VP-inversion.[37] Fourth, in the powerful approaches discussed above that can also account for the stranding asymmetry, verb clusters/VPR are characterized by a very natural and simple structural property: The dependent VP is contained within the projection of the governing verb. This is not the case if the restrictive theories are combined with low extraposition.

In conclusion, then, combining restrictive theories with low extraposition to derive 213 orders only represents a partial alternative.[38]

4.4 On the rarity of the 213 order

The last sections have established that the Swiss German 213 construction displays the properties of verb clusters. This raises the obvious question of why this order is so rare (if not inexistent) in other cluster types, both within Swiss German and elsewhere in Continental West Germanic. I believe that the rarity of the 213 order 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

[36] There are instances of displaced morphology where the morphology selected by V1 is displaced to V3 in a 213 order, cf. Heilmann (1999, 62, ex. 3d), Höhle (2006, 74, ex. 59ii) and Salzmann (2019a, this issue, section 2.4). This crucially cannot be accounted for under low extraposition/adjunction of VP3 to VP1 since VP3 would then end up above the functional head F into which the displaced morphology is inserted: F would be lower than VP1 so that F should attach to V2, contrary to fact.

[37] For an adjunction/extraposition analysis of VPR, see Vanden Wyngaerd (1989), Besten and Broekhuis (1992), and Haegeman (1992).

[38] When there is scrambling from VP3 in the 213 construction, the low extraposition approach predicts scope freezing; since these facts are generally very subtle, cf. section 6.2.3 below, I will not pursue this.

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 orders than in descending orders. 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 types does not come as a surprise.

The fact that the 213 order 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 the benefactives 'help/teach': 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.

Let me emphasize that these functional factors only provide motivation for the acceptability of the Swiss German 213 orders, but no explanation. The grammar sometimes grammaticalizes structures that may be suboptimal from a functional point of view. A good example is the 231 order in verb clusters. According to Culicover's metric, it should be just as marked as the 213 order. 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

[39] This is obvious for 'help/teach' but less so for the 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 where verb projection raising is otherwise readily available:

(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 VP as c-command should be sufficient for Case-assignment. But if it is 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'. Unfortunately, this cannot be the whole story because the Swiss German verbs are in principle compatible with an expletive (as in *I saw it snow*) so that the ECM-structure seems to be at least an option. 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.

Schmid 2005, 78) and even more so in Afrikaans (Biberauer 2013), 231 orders are frequent in Aux-Mod-Inf clusters, where Mod comprises modal, causative, benefactive, perception, durative and evidential verbs.[40]

As a final point, the facts discussed in this paper imply that a theory of verb clusters (and theories of word order more generally) 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 (classes of) 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. 5 as well as for examples found on the Internet like the following:

(71) und ine verzell was i alles chöne$_2$ han$_1$ mache$_3$
*and them tell.1*SG *what I all can.*INF *have.1*SG *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, accessed March 22, 2016

Similar restrictions will be needed for the 231 order. To my knowledge, within West-Germanic, 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).[41]

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 as more plausible to treat such gradient data in terms of markedness, i.e., orders with different

[40] Abels (2013) argues the 231 order is marked because of a mismatch between the prosodic and the syntactic structure that is not found in other cluster orders.

[41] See Dros-Hendriks (2018, chapter 3.7) 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.

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.[42]

5 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.

Against this background, it is remarkable that Swiss German (and possibly other varieties within Continental West-Germanic) has 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 V(P)3 as its complement. Obligatory weak pronoun fronting shows that the construction involves obligatory restructuring. I have argued that it is not an instance of the 3rd Construction, which is also a coherent construction and occurs in the 213 order. While previous work had not reached a consensus whether verb clusters/VPR and the 3rd Construction should be distinguished, I have shown, based on new diagnostics, that the relationship between the governing verb and the dependent VP is tighter in the case of verb clusters/VPR than in the 3rd Construction. This asymmetry in tightness can be related to the structural position of the dependent VP: While the dependent VP is in a low position, contained within the projection of the governing verb in verb clusters/VPR, it is in a high position in the 3rd Construction, outside of the projection of the governing verb. Several diagnostics converge on this result: displaced *zu*, missing *z*, short relative clause extraposition and stranding of VP3 under topicalization of the governing VP2. Applying these diagnostics to the Swiss German 213 construction delivers a clear result: It behaves as if VP3 is contained

[42] The theory proposed in Abels (2016, 186) only applies to neutral word orders. At first sight, this may seem to help solve the gradience issue (as only one order would have to be accounted for in a given cluster type). 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. Abels essentially follows Cinque (2005) in assuming that orders that only occur as marked orders but never constitute the neutral order in a single variety are derived by different means than the neutral orders (these alternative means are not dealt with in the paper). The question that now arises is whether this also holds for orders that are marked in a given variety but may be the neutral order in a different variety: If it does, this would imply that a marked 132 order in dialect A is derived differently than an unmarked 132 order in dialect B. This strikes me as rather implausible given that there is no evidence (to my knowledge) that marked 132 orders differ syntactically from unmarked 132 orders. If instead orders that are marked in some dialect but occur as unmarked orders in other dialects are derived by the same means, it is no longer clear why marked 213 orders (i.e. as in (71) and the residual occurrences in the major cluster types listed in fn. 5) should not be derived by regular means as well. And then the gradience problem returns.

within the projection of the governing verb 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 and theories of word order more generally must be powerful enough to generate all six logically possible orders.

Finally, while the data from relative clause extraposition and *zu*-placement could also be accounted for by the weaker assumption that all dependent VPs occur within the projection of the highest verb of the cluster, the stranding data crucially require VP3 to be *contained within the governing VP2*. This provides an argument against those approaches where in certain cluster orders, especially in the 312 order, VP3 is not contained within VP2. It also argues against a low-extraposition account of the Swiss German 213 order, which otherwise might be combined with restrictive approaches to cluster orders. Importantly, though, both the 213 order as such and its behavior with respect to displaced *zu*, short RC-extraposition and VP-stranding are compatible with several 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 verb clusters/VPR are characterized by a very general structural notion, viz., containment, that any theory must be able to capture.

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

In this appendix I will discuss two diagnostics that have been proposed 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.[43]

6.1 IPP-effect

The original motivation for the IPP-effect as a diagnostic for clusterhood 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):

(72) a. dat ik Ad {een ei} heb$_1$ *proberen*$_2$ {*een ei} te geven$_3$
*that I Ad an egg have.1*SG *try.*INF *an egg to give.*INF$_3$
b. dat ik Ad {een ei} *geprobeerd*$_2$ heb$_1$ {een ei} te geven$_3$
*that I Ad an egg try.*PTCP *have.1*SG *an egg to give.*INF
'that I tried to give Ad an egg' *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

[43] I 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.

3rd Construction. However, it is actually not quite clear what the IPP-effect diagnoses because it also occurs in VPR, where there is certainly no complex head, see the following examples from Swiss German (where the IPP-effect is generally optional, cf. Schmid 2005, 22f., Lötscher 1978, 3, fn. 2) and West Flemish, see Haegeman (1998, 275f.):

(73) a. dass i de Hans ha$_1$ ghöört$_2$/ghööre$_2$ en Arie singe$_3$
*that I the John have.1*SG *hear.*PTCP*/hear.*INF *an aria sing.*INF
'that I heard John sing an aria' *Swiss German*

b. dass i em Hans ha$_1$ ghulffe$_2$/hälffe$_2$ s Gschier
*that I the.*DAT *John have.1*SG *help.*PTCP*/help.*INF *the dishes*
abwäsche$_3$
*wash.*INF
'that I helped John do the dishes' *Swiss German*

(74) da Valére ee$_1$ willen$_2$ Marie dienen boek geven$_3$
*that Valere have.3*SG *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 (75) (IPP in 213 orders is also attested in Vorarlberg German, see Schallert 2014, 195f., in Swabian, cf. Heilmann 1999, 62, ex. 3d, and in earlier Pennsylvania Dutch, cf. Louden 2011, 178; as far as I can tell, though, at least in Swiss German the participle is clearly the more frequent option):

(75) ... wo de Alkohol [...] afange$_2$ hät$_1$ e Rolle spile$_3$
*when the alcohol start.*INF *have.3*SG *a role play.*INF
'when the alcohol started playing a role in my life'
http://hpgmuender.blogspot.fr; blog on September 29,2007; March 25, 2016

Whatever the IPP-effect indicates, given (75), there is no reason to conclude that the Swiss German 213 construction is an instance of the 3rd Construction.

6.2 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).

6.2.1 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):

(76) a. Was$_k$ hat$_1$ Heinrich __$_k$ für einem Kind vergessen$_2$ die
*what have.3*SG *Hendrik for a.*DAT *child forget.*PTCP *the*
Zebras zu zeigen$_3$?
*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$_k$ tänksch, dass de Hans hät$_1$ __$_k$ für Lüüt
*what believe.2*SG *that the Hans have.3*SG *for people*
wele$_2$ vo siine Idee überzüüge$_3$?
*want.*INF *of his ideas convince.*INF
'What kind of people do you think John wanted to convince of his ideas?' *Swiss German*, cf. Salzmann (2011, 462)

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.):

(77) a. Er hat$_1$ einem Kind versucht$_2$ das MÄRCHEN
*he have.3*SG *a.*DAT *child try.*PTCP *the fairy.tale*
vorzulesen$_3$.
*read.to.*INF
'He tried to read the fairy tale to a girl.' *Standard German*

b. Wenn er einem Kind hätte$_1$ das MÄRCHEN vorlesen$_3$
*if he a.*DAT *child had.*SBJV.3SG *the fairy.tale read.to.*INF
dürfen$_2$
*may.*INF
'if he had been allowed to read the fairy tale to a child'
Standard German

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

(78) a. ?Ich habe$_1$ ihm **was** versucht$_2$ nach Berlin zu schicken$_3$.
*I have.1*SG *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** hätte$_1$ nach Berlin schicken$_3$
*if I he.*DAT *something had.*SBJV.1SG *to Berlin send.*INF
können$_2$
*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. pseudoscrambling (Geilfuß-Wolfgang 1991) is involved or no movement whatsoever (cf. e.g. Bayer and Kornfilt 1994, Fanselow 2001, Salzmann 2011) is an open question; quite probably, the differences follow from independent factors (surface generalizations, freezing being restricted to topical XPs etc.).

6.2.2 Semantic differences between the 3rd Construction and VPR

Next to these similarities, a number of 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 should stress that all facts in the remainder of this section are very subtle and native speakers (including the anonymous reviewers) frequently disagree on the judgments. I will not take a stand here but merely cite from the literature.

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.):

(79) a. ??Peter hat$_1$ **einen Adventskalender** vergessen$_2$ zu
*Peter have.3*SG *a advent.calendar forget.*PTCP *to*
basteln$_3$.
*make.*INF
'Peter forgot to make an advent calendar.' *Standard German*

b. dass Peter **einen Adventskalender** hätte$_1$ für mich
*that Peter a advent.calendar have.*SBJV.*3*SG *for me*
basteln$_3$ sollen$_2$
*make.*INF *want.*INF
'that Peter should have made an advent calendar for me' *Std. G.*

Nor can the displaced XP be an *idiom chunk* in the 3rd Construction (Geilfuß-Wolfgang 1991, 52), while this is unproblematic in VPR (Salzmann 2011, 461):

(80) a. ?*Er hat$_1$ seinem Onkel **einen Bären** versucht$_2$
*he have.3*SG *his uncle a bear try.*PTCP
aufzubinden$_3$.
*tie.to.*INF
'He tried to pull his uncle's leg.' *Standard German*

b. dass er **känere Flüüg** hät$_1$ chöne$_2$ öppis z
*that he no.*DAT *fly have.3*SG *can.*INF *something to*
Leid tue$_3$
*suffering do.*INF
'that he could not harm anyone' *Swiss German*

As for 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):

(81) a. Er HAT$_1$ *mindestens ein Geschenk* versucht$_2$ *fast*
*he have.3*SG *at least one present try.*PTCP *almost*
jedem Gast t$_{mindestens\ ein\ Geschenk}$ zu überreichen$_3$
*every.*DAT *guest to hand.over.*INF
'He tried to hand over at least one present to almost every guest.'
$\exists \succ \forall$; *$\forall \succ \exists$; *Standard German*

b. DASS er *mindestens ein Geschenk* hat$_1$ *fast jedem*
*that he at least one present have.3*SG *almost every.*DAT
Gast t$_{mindestens\ ein\ Geschenk}$ überreichen$_3$ wollen$_2$
*guest hand.over.*INF *want.*INF
'that he wanted to hand over at least one present to almost every guest' $\exists \succ \forall$; $\forall \succ \exists$; *Standard German*

If a quantified XP interacts with a scopal matrix verb, wide-scope seems obligatory in the 3rd Construction but not in VPR, cf. Salzmann (2011, 454) for VPR and Bobaljik and Wurmbrand (2005, 810, 831) for the 3rd Construction:

(82) a. weil er [$_{VP1}$ **alle Fenster** *vergass*$_1$ [$_{VP2}$ t$_{alleFenster}$
*because he all windows forget.*PST.*3*SG
zu schliessen$_2$]
*to close.*INF
'because he forgot to close all the windows'
$\forall \succ$ forget; *forget $\succ \forall$; *Standard German*

b. dass er [$_{VP1}$ **2 Manager** *wett*$_1$ [$_{VP2}$ t$_{2\ Manager}$ vo siine
*that he 2 managers want.3*SG *of his*
Idee überzüüge$_2$]]
*ideas convince.*INF
'that he wants to convince two managers of is ideas'
$2 \succ$ want; want $\succ 2$; *Swiss German*

6.2.3 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 (54), because remnant movement is well-known to induce scope freezing effects (Barss 1986, 517–542; for a recent proposal, see Sauerland and Elbourne 2002):

(83) 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 (adapted from Haider 2003, 101; (84a) additionally has a split-scope reading, which (84b) also lacks):

(84) a. dass sie [$_{\text{VP1}}$ **nichts** [$_{\text{VP2}}$ t$_{nichts}$ zu essen$_2$] *wagte*$_1$]
that she nothing to eat.INF *dare*.PST.*3*SG
'that she didn't dare to eat anything/dared not to eat anything'
¬∃ ≻ dare; dare ≻ ¬∃

b. [$_{\text{VP2}}$ t$_{nichts}$ zu essen$_3$] *wagte*$_1$ sie [$_{\text{VP1}}$ **nichts** t$_{VP2}$ t$_{wagte}$].
to eat.INF *dared she nothing*
'there is nothing that she dared to eat' ¬∃ ≻ dare; *dare ≻¬∃
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.

6.2.4 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 example (85):

(85) omdat Jan **een huis** besloot$_1$ te kopen$_2$
because John a house decide.PST.*3*SG *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 *vergessen* '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:

(86) a. omdat Jan **de zak** beweert$_1$ te krijgen$_2$
*because John the bag claim.3*SG *to get*.INF
'because John claims to get sacked' *Standard Dutch*

b. omdat Jan Marie **een loer** besloot$_1$ te draaien$_2$
because John Mary a lurk decide.PST.*3*SG *to turn*.INF
'because John decided to play a nasty trick on Mary' *StD*

c. weil Hans der Maria **einen Streich** versuchte$_1$ zu
*because John the.*DAT *Mary a trick try.*PST.*3*SG *to*
spielen$_2$
*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 *vergessen* '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 ((87a) is from Sternefeld 2006, 654):

(87) a. Ratzinger, der **keine Kompromisse** *bereit ist*$_1$
*Ratzinger, who no compromises willing be.3*SG
einzugehen$_2$
*to make.*INF
'Ratzinger, who is not willing to make any compromises'
$\neg \succ$ willing $\succ \exists$; *Standard German*

b. dass du **keinen Schlips** *brauchst*$_1$ anzuziehen$_2$
*that you no tie need.2*SG *to wear.*INF
'that you need not wear a tie' $\neg \succ$ need $\succ \exists$; *Standard German*

c. dass er **kein Fleisch** *versuchte*$_1$*/wagte*$_1$ zu essen$_2$
*that he no meat try.*PST.3SG*/dare.*PST.3SG *to eat.*INF
'That he tried/dared not to eat any meat' try $\succ \neg\exists$; *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 a challenging topic for further research.[44]

6.2.5 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 not to reconstruct, while in VPR the reconstructed and the non-reconstructed interpretation seem equally available. This asymmetry would fit perfectly with the results reached in this

[44] If what looks like the 3rd Construction in Swiss German can actually involve a verb cluster structure, recall sections 3.3.2 and 3.4.3.1, 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) indeed argues that scope reconstruction is possible in the 3rd Construction in Zurich German.

paper that the 3rd Construction involves remnant extraposition of the lexical VP (so that the scope facts can be subsumed under Barss' generalization), while it remains in a complement position 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 at this point.

Acknowledgements This research has been supported by the SNSF-grant PA00P1_136379/1 and the DFG-grant SA 2646/1-1. Earlier versions of this research were presented at the DGfS in Potsdam (March 2013), at the University of Tübingen (June 2013), at the University of Leipzig (June 2013), at the workshop on Swiss German syntax in Arezzo (September 2013), at the CRISSP seminar in Brussels (June 2014), at SinFonIJA in Graz (September 2014), at CGSW 29 in York (September 2014), at the workshop on dialect syntax: state of the art in Frankfurt (December 2014), at the workshop on verb clusters in Amsterdam (May 2015) and at the University of Göttingen (January 2017). I thank the audiences for helpful feedback, in particular Sjef Barbiers, Hans Bennis, Theresa Biberauer, Claudia Bucheli, Jeroen van Craenenbroeck, Jürg Fleischer, Doreen Georgi, Fabian Heck, Lotte Hendriks, Anke Himmelreich, Roland Hinterhölzl, Tony Kroch, Guido Mensching, Walt Detmar Meurers, Gereon Müller, Henk van Riemsdijk, Oliver Schallert, Guido Seiler, Markus Steinbach, Wolfgang Sternefeld, Tanja Temmerman, Jochen Trommer, Philipp Weißer, Guido Vanden Wyngaerd, Hedde Zeijlstra, Eva Zimmermann, and Jan-Wouter Zwart. I am very grateful to Martin Graf and Christoph Landolt from the Idiotikon for help with Swiss German data. Particular thanks go to Klaus Abels for very enriching discussions about the 213 order. The paper has been significantly improved by comments of three anonymous JCGL-reviewers. Finally, I would like to thank the editor in Chief, Susi Wurmbrand, for support and enthusiasm for my project and her numerous valuable suggestions on different versions of this paper. The usual disclaimers apply.

References

Abels, Klaus. 2013. On 2-3-1. *UCL Working Papers in Linguistics* 25: 67–101.

Abels, Klaus. 2016. The fundamental left–right asymmetry in the Germanic verb cluster. *The Journal of Comparative Germanic Linguistics* 19 (3): 179–220.

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 Colloquial German. *The Journal of Comparative Germanic Linguistics* 12 (3): 175–228.

Bader, Thomas. 1995. Missing and misplaced z' in Bernese Swiss German. In *Topics in Swiss German syntax*, ed. Zvi Penner, 19–28. Bern: Peter Lang.

Barbiers, Sjef. 2005. Word order variation in three verb clusters and the division of labor between generative linguistics and sociolinguistics. In *Syntax and variation. Reconciling the biological and the social*, eds. Leonie Cornips and Karen P. Corrigan, 233–264. Amsterdam: John Benjamins.

Barbiers, Sjef. 2008. Werkwoordclusters en de grammatica van de rechterperiferie. *Nederlandse Taalkunde* 13 (2): 160–188.

Barbiers, Sjef, Hans Bennis, and Lotte Dros-Hendriks. 2018. Merging verb cluster variation. *Linguistic Variation* 18 (1): 144–196.

Barbiers, Sjef, Johan van der Auwera, Hans Bennis, Eefje Boef, Gunther de Vogelaer, and Margreet van der Ham. 2008. *Syntactic atlas of the Dutch dialects, volume II.* Amsterdam: Amsterdam University Press.

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 *Studies on scrambling. Movement and non-movement approaches to free word-order phenomena*, eds. Norbert Corver and Henk van Riemsdijk, 17–60. Berlin: Mouton.

Bayer, Josef, Tanja Schmid, and Markus Bader. 2005. Clause union and clausal position. In *The function of function words and functional categories*, eds. Marcel den Dikken and Christina M. Tortora, 79–113. Amsterdam: John Benjamins.

Bech, Gunnar. 1983. *Studien über das deutsche Verbum infinitum*. Tübingen: Niemeyer.

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.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45 (2): 169–225.

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*. Hildesheim: Olms.

Brandner, Ellen, Martin Salzmann, and Gerhard Schaden. 2016. Zur Syntax und Semantik des doppelten Perfekt aus alemannischer Sicht. In *Syntaktische Variation – areallinguistische Perspektiven*, eds. Alexandra Lenz and Franz Patocka, 13–45. Vienna: Vienna University Press.

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 *Restructuring and functional heads: The cartography of syntactic structures volume 4*, ed. Guglielmo Cinque, 11–63. New York: Oxford University Press.

Cooper, Kathrin E. 1995. *Topics in Zurich German syntax*. Groningen: Germanistisch Instituut.

Culicover, Peter. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.

de Vries, Mark. 2002. *The syntax of relativization*. Utrecht: LOT.

den Dikken, Marcel. 1994. Minimalist Verb (Projection) Raising. *Groninger Arbeiten zur Germanistischen Linguistik* 37: 71–88.

den Dikken, Marcel. 1995. Verb (Projection) Raising, scope, and uniform phrase structure. *Proceedings of NELS* 25: 95–110.

den Dikken, Marcel. 1996. The minimal links of Verb (Projection) Raising. In

Minimal ideas. Syntactic studies in the minimalist framework, eds. Werner Abraham, Samuel David Epstein, Hoskuldur Thrainsson, and C. Jan-Wouter Zwart, 67–96. Amsterdam: Benjamins.

Dros-Hendriks, Lotte. 2018. *Not another book on verb raising.* Utrecht: LOT.

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.

Geilfuß-Wolfgang, Jochen. 1991. Scrambling und Pseudoscrambling. In *Verb- und Verbphrasensyntax. Arbeitspapier 11 des Sonderforschungsbereichs 340 Sprachtheoretische Grundlagen für die Computerlinguistik*, ed. Jochen Geilfuss-Wolfgang, 19–57. Tübingen: University of Tübingen.

Glaser, Elvira. in preparation. Dialektsyntax des Schweizerdeutschen. Syntaktischer Atlas der Deutschen Schweiz (SADS).

Haeberli, Eric, and Susan Pintzuk. 2012. Revisiting verb (projection) raising in Old English. In *Grammatical change*, eds. Dianne Jonas, John Whitman, and Andrew Garret, 219–238. Oxford: Oxford University Press.

Haegeman, Liliane. 1992. *Theory and description in generative syntax. A case study in West Flemish.* Cambridge: Cambridge University Press.

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. 2003. V-clustering and clause union: Causes and effects. In *Verb constructions in German and Dutch*, eds. Pieter Seuren and Gerard Kempen, 91–126. Amsterdam: John Benjamins.

Haider, Hubert. 2010. *The syntax of German.* Cambridge: Cambridge University Press.

Heilmann, Axel. 1999. Die VP im Schwäbischen. Doctoral dissertation, University of Stuttgart, Stuttgart.

Hinrichs, Erhard, and Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In *German in Head-Driven Phrase Structure Grammar*, eds. John Nerbonne, Klaus Netter, and Carl Pollard. Stanford: CSLI Publications.

Hinterhölzl, Roland. 2006. *Scrambling, remnant movement, and restructuring in West Germanic.* Oxford; New York: Oxford University Press.

Hinterhölzl, Roland. 2009. The IPP-effect, phrasal affixes and repair strategies in the syntax-morphology interface. *Linguistische Berichte* 2009 (218): 191–215.

Höhle, Tilman N. 2006. Observing non-finite verbs: some 3v phenomena in German-Dutch. In *Form, structure, and grammar*, eds. Patrick Brandt and Eric Fuß, 55–77. Berlin: Akademie Verlag.

Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes.* Cambridge,

Mass: MIT Press.
Kroch, Anthony, and Beatrice Santorini. 1991. The derived constituent structure of the West-Germanic verb-raising construction. In *Principles and parameters in comparative grammar*, ed. Robert Freidin, 269–338. Cambridge: MIT Press.
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 *Studies on German-language islands*, ed. Michael T. Putnam, 165–185. Amsterdam: John Benjamins.
Müller, Gereon. 1995. *A-bar syntax: a study in movement types*. Berlin: Mouton de Gruyter.
Müller, Stefan. 2002. *Complex predicates: verbal complexes, resultative constructions and particle verbs in German*. Stanford, CA: CSLI Publications.
Neeleman, Ad. 1990. The Third Construction as a classical case of head-movement. Ms. Utrecht University.
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.
Salzmann, Martin. 2011. Resolving the movement paradox in verb projection raising. In *Empirical issues in syntax and semantics 8*, eds. Olivier Bonami and Patricia Cabredo Hofherr, 453–485. Paris: CNRS.
Salzmann, Martin. 2013. New arguments for verb cluster formation at PF and a right-branching VP. Evidence from verb doubling and cluster penetrability. *Linguistic Variation* 13 (1): 81–132.
Salzmann, Martin. 2019a, this issue. Displaced morphology in German verb clusters. An argument for post-syntactic morphology. *Journal of Comparative Germanic Linguistics*.
Sapp, Christopher D. 2011. *The verbal complex in subordinate clauses from medieval to modern German*. Amsterdam; Philadelphia: John Benjamins.
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*. Tübingen: Stauffenburg.
Schmid, Tanja. 2005. *Infinitival syntax*. Amsterdam: Benjamins.
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 *Dialectology meets typology*, ed. Bernd Kortmann, 367–399. Berlin, Germany: Mouton de Gruyter.

Sternefeld, Wolfgang. 2006. *Syntax: eine morphologisch motivierte generative beschreibung des Deutschen*. Tübingen: Stauffenburg Verlag.

ter Beek, Janneke. 2008. *Restructuring and infinitival complements in Dutch*. Utrecht: LOT.

Vanden Wyngaerd, Guido. 1989. Verb projection raising and the status of infinitival complements. In *Sentential complementation and the lexicon. Studies in honour of Wim de Geest*, eds. Dany Jaspers, Yvan Putseys, Wim Klooster, and Pieter Seuren. *Linguistic models*, 423–438. Dordrecht: Foris.

Vogel, Ralf. 2009. Skandal im Verbkomplex. Betrachtungen zur scheinbar inkorrekten Morphologie in infiniten Verbkomplexen des Deutschen. *Zeitschrift für Sprachwissenschaft* 28 (2): 307–346.

Weber, Albert. 1987. *Zürichdeutsche Grammatik: ein Wegweiser zur guten Mundart*. Zürich: Schweizer Spiegel Verlag.

Wild, Katharina. 1991. Zur Satzgliedstellung in den "Fuldaer" deutschen Dialekten Südungarns. *Zeitschrift für Dialektologie und Linguistik* 58: 24–43.

Wöllstein-Leisten, Angelika. 2001. *Die syntax der dritten Konstruktion: eine repräsentationelle analyse zur Monosententialität von "zu"-Infinitiven im Deutschen*. Tübingen: Stauffenburg-Verlag.

Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.

Wurmbrand, Susi. 2004a. Syntactic vs. post-syntactic movement. In *Proceedings of the 2003 annual conference of the Canadian linguistic association*, eds. Sophie Burelle and Stanca Somesfalean, 284–295.

Wurmbrand, Susi. 2004b. Two types of restructuring–lexical vs. functional. *Lingua* 114 (8): 991–1014.

Wurmbrand, Susi. 2004c. West Germanic verb clusters: The empirical domain. In *Verb clusters: A study of Hungarian, German, and Dutch*, eds. Katalin É Kiss and Henk van Riemsdijk, 43–85. Amsterdam: John Benjamins.

Wurmbrand, Susi. 2007. How complex are complex predicates? *Syntax* 10 (3): 243–288.

Wurmbrand, Susi. 2017. Verb clusters, verb raising, and restructuring. In *The Blackwell companion to syntax. 2nd edition*, eds. Martin Everaert and Henk van Riemsdijk. Oxford: Wiley-Blackwell.

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 *Linguistics in the Netherlands 1995*, eds. Marcel den Dikken and Kees Hengeveld, 215–226. Amsterdam: John Benjamins.

Zwart, Jan-Wouter. 1996. Verb clusters in continental West Germanic dialects. In *Microparametric syntax and dialect variation*, eds. James R. Black and Virginia Motapanyane, 229–258. Amsterdam: John Benjamins.

Zwart, Jan-Wouter. 2007. Some notes on the origin and distribution of the IPP-effect. *Groninger Arbeiten zur Germanistischen Linguistik* 45: 77–99.