Rule Ordering in Verb Cluster Formation: On the Extraposition Paradox and the Placement of the Infinitival Particle *te/zu*

Martin Salzmann[*]

Abstract

This paper addresses two puzzles in the domain of verb cluster formation and proposes a solution in terms of rule ordering. The first puzzle is the so-called extraposition paradox where extraposition can target a VP that is part of a verb cluster only if the VP is topicalized but not when the VP remains clause-final. I propose that verb cluster formation takes place at PF under adjacency and thus after extraposition and topicalization. Extraposition and topicalization can therefore bleed cluster formation, leading to a crash of the derivation if the VP remains in-situ. The second puzzle involves the placement of the infinitival marker *te/zu* in Dutch and German. I will show that the cross-linguistic differences in placement follow from the fact that the rule that associates the particle with the verb takes place at different points of the PF-derivation in the two languages. While it is an early operation in Dutch and is still sensitive to hierarchical structure, it is a late process in German and is therefore subject to linear order and adjacency. Both operations interact with other PF rules, and I will demonstrate that it is possible to determine a strict and non-contradictory (and predominantly intrinsic) ordering of the rules which as a side-effect provides evidence for the articulation of the PF-component. Finally, I will show that the *zu*-placement facts do not provide decisive evidence in favor of either a right-branching or a left-branching VP-structure; rather, the advantages and disadvantages of the two views turn out to largely balance each other out.

1. Introduction

While the PF-branch of grammar was for a long time kind of the syntacticians waste basket that hosted syntactic phenomena (e.g. stylistic rules) that could

---

not so easily be accommodated within the standard assumptions of syntactic theory (even though many of them seemed rather syntactic, e.g. locative inversion), the last two decades have seen a shift towards serious investigations into the structure of the PF-component. On the one hand, this was triggered by developments within syntactic theory that attempted to circumscribe more narrowly the operations that syntax proper is supposed to perform. On the other hand, the introduction of Distributed Morphology in Halle and Marantz (1993) radically changed the view on the morphology-syntax interface and opened new prospects for post-syntactic operations. An important point in that development was the proposal by Embick and Noyer (2001) who provided convincing evidence for a division of the PF-component into at least two subcomponents. The division was empirically motivated on the basis of post-syntactic movement operations. Since the PF-component gradually transforms hierarchical syntactic structure into a linear structure that can be interpreted at the interface, movement operations that apply early in the PF-branch will tend to be sensitive to hierarchical structure while later movement operations will be sensitive to linear properties of the structure such as adjacency. Operations of the first type are termed Lowering, which basically amounts to downward head-movement; operations of the second-type are called Local Dislocation. Both operations can be responsible for the placement of clitics and affixes that surface in a position different from the phrase marker whose terminal nodes they realize. A very recent contribution to the research into the PF-component is Arregi and Nevins (2012) who provide further evidence for a highly articulated post-syntactic component and pervasive interaction of post-syntactic operations instantiating classic feeding, bleeding and opaque relationships.

The goal of this paper is a modest attempt to contribute to this discussion by examining two phenomena from West-Germanic syntax that lend themselves to a post-syntactic treatment, viz. cluster formation and the placement of the infinitival particle. As I will argue, a full account requires the postulation of a number of post-syntactic operations that need to apply in a certain order. This rule interaction is then used as a diagnostic: If it can be demonstrated that a larger number of such rules can be ordered in a non-contradictory way, we have made progress towards an understanding of the structure of the PF-component.

The paper is organized as follows: In section 2, I will discuss an extraposition paradox and argue that verb cluster formation takes place post-syntactically. Based on these findings, in section 3, I will investigate the placement of the

infinitival particle in German and Dutch and propose solutions to account for the cross-linguistic variation as well as for the intricate placement pattern in German and its dialects. Section 4 concludes.

2. The Extraposition Paradox

2.1. The Problem

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

(1) 321 *Standard German*
dass man darüber [reden$_3$ können$_2$ sollte$_1$]
that one about.it talk.INF can.INF should.3SG
'that one should be able to talk about it'

Such sequences are usually referred to as verb clusters (for a detailed overview, cf. Wurmbrand 2005); furthermore, there is a long tradition (starting with Evers 1975) that analyzes verb clusters as complex syntactic heads, e.g. as follows:

(2)

---
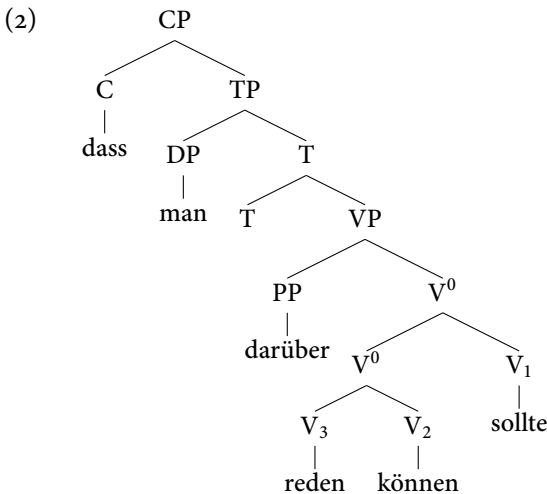
[1]Numbers on verbs indicate the embedding relations, i.e. 1 stands for the highest, i.e. the embedding verb, 2 for the immediately embedded verb etc.

The major empirical evidence for a complex head comes from the following observation: Extraposition, an operation which targets VP in German, cannot target VP_3 or VP_2 in a V-final structure involving a verb cluster (cf. van Riemsdijk 1998: 640ff., Haider 2003: 92ff., Bayer et al. 2005: 91):[2]

(3)   *Standard German*
       dass man [$_{VP_1}$ [$_{VP_1}$ [$_{VP_2}$ [$_{VP_2}$ [$_{VP_3}$ [$_{VP_3}$ t$_{darüber}$ reden$_3$] ***darüber**]
       that one   talk.INF about.it
       können$_2$] ***darüber**] sollte$_1$] **darüber**]
       can.INF   about.it   should about.it
       'that one should be able to talk about it'

This restriction follows directly under a complex head analysis since XP-movement cannot target segments of V, but only the maximal projection. The real challenge obtains under topicalization. Suddenly, extraposition *can* target VP_3:

(4)   *Standard German*
       [$_{VP_3}$ [$_{VP_3}$ t$_{darüber}$ reden$_3$] **darüber**] sollte$_1$ man schon [$_{VP_1}$ [$_{VP_2}$
                           talk.INF about.it   should.3SG one   indeed
       t$_{VP_3}$ können$_2$] t$_{sollte}$]
               can.INF
       'that one should be able to talk about it'

In other words, (4) seems to be derived from an ungrammatical structure. There are two possibilities to resolve this tension: Either cluster formation is taken to be optional or excorporation is allowed.

In the first case, cluster formation (in the sense of forming a complex head) would not take place in the V-final structure so that extraposition can target VP_3, thus providing a base for (4). However, once cluster formation is taken to be optional, it is no longer clear how the ungrammatical versions of (3) can be

---

[2]I adopt a left-branching VP-structure in what follows; the exact relationship between syntax and linear order is discussed in more detail in section 3.4.1; the consequences that arise once a right-branching VP-structure is adopted are discussed in section 3.7
For the purpose of this paper, I will label all verbal elements involved in verb clusters as lexical verbs, i.e. including modals, auxiliaries and others even though there may be reasons to classify some of them as functional elements (containing more structure), cf. Wurmbrand (2004a); as far as I can tell, nothing in what follows hinges on this.

derived. We are thus faced with a paradox: cluster formation would have to be optional and obligatory at the same time. The alternative, excorporation, does not fare much better: To derive (4), both V_1 and and V_2 would have to excorporate before topicalization takes place. While excorporation of V_1 may in principle be plausible as it moves to C, this is not the case with V_2, as it is generally assumed that the verbs stay in their base position in V-final structures.

One can thus conclude that neither excorporation nor cluster formation is sufficient to account for the pattern in (3) and (4). Interestingly, apart from Wurmbrand (2007), the issue has not been addressed in much detail. Haider (2003) assumes that the topicalization structure and the V-final structure are not transformationally related; rather, both the topicalized VP and the (partial) clause-final complex head are independently base-generated as such. Unfortunately, he does not spell-out how the two are related to each other: if the VP undergoes long-distance movement or occurs with a 4-verb-cluster, there will be at least part of a complex head clause-finally:

(5)   *Standard German*

[$_{VP_4}$ [$_{VP_4}$ t$_{darüber}$ reden$_4$] darüber] sollte$_1$ man schon [$_{VP}$ können$_3$
                        talk.INF about.it   should one indeed   can.INF

wollen$_2$ t$_{sollte}$].
want.INF

'One should want to be able to talk about it.'

The interpretation of the topicalized constituent as a complement of only a part of that complex head (i.e. V_3 *können* 'can') is certainly non-trivial (like other interpretive aspects such as adverbial modification, see Wurmbrand 2007 for critical discussion). A tentative solution is sketched in Haider (2010: 307).[3] Bader and Schmid (2009: 202, fn.11), who assume that verb clusters are base-generated complex heads, admit that for cases like (4), it may be necessary that V_1 actually selects a VP and not a V^0 as they assume elsewhere. But once this possibility is granted, the cluster property in (3) cannot be derived anymore.

---

[3]Note that such examples cannot be reanalyzed as cases of left-dislocation with deletion of the fronted proform (that would anaphorically refer to the fronted VP) because topicalization and left-dislocation do not always pattern the same. As pointed out in Haider (2003: 95f.), once lexical specifications of the verb (oblique cases, *wh*-complements) are involved, left-dislocation becomes impossible while topicalization does not.

Haegeman and van Riemsdijk (1986: 451), who assume that verb clusters result from reanalysis, simply mention that verb-clusters permit extraction and thus are not "lexical" in the sense of being impenetrable; the cluster paradox thus seems to remain unaccounted for under their approach as well. It is therefore fair to conclude that approaches adopting a complex syntactic head (either base-generated or derived in syntax) cannot resolve the contradiction between (3) and (4) in a straightforward way.[4]

2.2. The Solution: Timing

I would like to propose a solution to the extraposition paradox that makes crucial use of timing. I adopt the standard assumptions that extraposition, V-to-C-movement and topicalization take place in syntax. Where I differ from much of the literature is that I assume that verb cluster formation in the sense of forming a complex head takes place at PF under adjacency and thus arguably represents an instance of what Embick and Noyer (2001) have referred to as Local Dislocation. Importantly, what is inverted are not syntactic sisters (e.g. V_1 and VP_2 as e.g. in Haegeman and van Riemsdijk 1986, Wurmbrand 2004*b*). Since the empirical evidence for this position is discussed in much detail in Salzmann (to appear), I will not reproduce the arguments here but will simply take this as given.[5] Under these assumptions, the extraposition paradox can be accounted for as follows:

I will start with the cluster property of the V-final structure in (3). There are no complex heads in the syntax but stacked VPs instead. Extraposition can therefore in principle target either VP_3, VP_2 or VP_1, as indicated in the following structure:

---

[4]See Wurmbrand (2007) for an explanation of the paradox in (3) and (4) without the postulation of a complex head. She argues instead that extraposition is subject to prosodic restrictions which force extraposed constituents to appear at the edges of prosodic constituents. In a V-final structure, this is only the case if the extraposee occurs adjoined to VP_1 while under topicalization, VP_3 constitutes a separate prosodic domain and thus constitutes a legitimate attachment site.

[5]Adjacency constraints on verb cluster formation have, of course, been proposed before, cf. e.g. van Riemsdijk (1998: 639-645) where cluster formation involves syntactic head-movement. However, given that head-movement is normally not subject to such a constraint, an adjacency condition seems stipulative. Placing cluster formation as an adjacency-sensitive operation in the post-syntactic component seems more in line with current conceptions about the architecture of grammar.

(6)   *Standard German*
dass man [$_{VP_1}$ [$_{VP_1}$ [$_{VP_2}$ [$_{VP_2}$ [$_{VP_3}$ [$_{VP_3}$ t$_{darüber}$ reden$_3$] **darüber**]
that one   talk.INF about.it
können$_2$] **darüber**] sollte$_1$] **darüber**]
can.INF   about.it   should about.it
'that one should be able to talk about it'

I will assume that cluster formation is obligatory if the verbal elements are adjacent at PF, at least in descending order, cf. fn. 17 for details. Under a left-branching VP-structure, cluster formation will be string-vacuous in descending orders (which amounts to re-bracketing); under a right-branching VP-structure as discussed in section 3.7 (and also in 312 clusters based on left-branching structures, cf. section 3.4.2), cluster formation will additionally involve reordering/inversion. Technically, I assume that clustering is enforced by a surface constraint. In (6), cluster formation will be blocked if extraposition targets VP_3 or VP_2 as the extraposed PP will destroy adjacency between the verbal elements. Put differently, extraposition to VP_3 or VP_2 *bleeds* cluster formation. In this case, the derivation crashes because the surface constraint requiring the formation of a complex head is violated. Only extraposition to VP_1 is an option. This derives (3).

In the topicalization structure in (4), extraposition targets VP_3. This is licit since cluster formation does not take place at the point where VP_3 is still in its base-position. Rather, topicalization of VP_3 destroys the context for cluster formation: there is only one verbal element in the prefield so that no cluster formation needs to take place (the two remaining verbal elements at the end of the clause, however, do undergo cluster formation). In this case, topicalization also bleeds cluster formation, but the result is grammatical because no surface/PF-constraint is violated.

This approach based on timing of operations has an additional advantage. It directly explains why verb-second movement never involves a complex head: the element fronted to C is always just a single verb (the finite verb), complex verbs are ruled out:

(7)   *Standard German*
        a.   *Man [reden$_3$ können$_2$ sollte$_1$] darüber schon.
               one talk.INF can.INF   should about.it indeed
               'One should be able to talk about it.'

b.   Man sollte$_1$ darüber schon [reden$_3$ können$_2$ t$_{sollte}$].
      one   should about.it indeed talk.INF can.INF
      'One should be able to talk about it.'

Under the present approach, this follows automatically since at the point where verb second movement takes place, the verbs have not yet formed a cluster (i.e. like topicalization, verb second movement bleeds cluster formation). Consequently, only V_1 will move. Approaches that assume the formation of a complex head in syntax and relate V-final and V-second structures via movement have to make extra assumptions to rule out movement of the complex head (such as e.g. the complexity constraint in Neeleman and Weerman 1993: 460ff.).

To summarize: in this section, I have discussed the following rules/operations: extraposition, topicalization, and verb-cluster formation. They apply in the following order:

(8)   extraposition > topicalization > verb-cluster formation

The ordering between the first two operations is intrinsic as it follows from cyclicity, e.g. as formulated in the *Strict Cycle Condition* (Chomsky 1973) or the Extension Condition (Chomsky 1995). The ordering between topicalization and cluster formation is also intrinsic as it results from the fact that the two operations take place in different components of the grammar that are sequentially ordered, viz. syntax vs. PF (which can thus also be regarded as two separate cycles). In two configurations, application of R_1 can bleed application of R_2: In (3), extraposition bleeds cluster formation, in (4) topicalization bleeds cluster formation.

3.   The Placement of the Infinitival Particle *te/zu*

In this section, I will discuss the placement of the infinitival particle in Dutch/German. As we will see, the placement is not fully straightforward, at least in German, where the particle does not always occur in the position expected on the basis of its morphosyntactic properties. The major focus of attention is to pin down the point where the placement occurs. As it is dependent on a number of operations involving the verbal complex, the exact ordering of the various operations will be crucial.

3.1.   (Standard) Dutch vs. (Standard) German

Dutch and German differ from each other with respect to the placement of the infinitival particle: While it occurs at the end of the verb cluster in German, it surfaces at the beginning of the cluster in Dutch. In the following example, the matrix verb 'think' takes a non-finite complement clause where the hierarchically highest verb appears with the particle *te/zu*.[6]

(9)   a.   321 *Standard German*
            Er dachte, das Buch [lesen$_3$ können$_2$ **zu** müssen$_1$].
            he thought the book read.INF can.INF   to must.INF
            'He thought he had to be able to read the book.'
       b.   123 *Standard Dutch*
            Hij dacht   het boek [**te** moeten$_1$ kunnen$_2$ lezen$_3$].
            he thought the book to must.INF can.INF   read.INF
            'He thought he had to be able to read the book.'

Based on these data, the most straightforward analysis seems to be to treat the particle as a prefix, as e.g. proposed in Haider (1993). The particle simply surfaces on the hierarchically highest verb, viz. V_1. The surface difference would be the result of independent differences in verb cluster formation; for instance, one might argue (as in the classical analysis by Evers 1975) that cluster formation involves adjunction to the left in German but adjunction to the right in Dutch, leading to reversed, i.e. ascending surface structure in Dutch while the descending order of the base is retained in German. Further evidence for a prefix analysis comes from the following examples where there are two verbs (*versprechen* 'promise' and *versuchen* 'try') that select a *te-/zu*-infinitive. Here, the particle appears twice, on each verb that is dependent on a predicate selecting a *zu*-infinitive:

(10)   a.   321 *Standard German*
            dass er [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ das Buch **zu** lesen$_3$]   **zu** versuchen$_2$]
            that he   the book to read.INF to try.INF
            versprach$_1$]
            promised
            'that he promised to try to read the book'

---

[6]Lexical verbs selecting a non-finite complement usually take a so-called *te/zu*-infinitive where the infinitival verb is accompanied by the particle while modals and a few other verbs like perception verbs require a so-called bare infinitive, i.e. an infinitival verb without particle.

b.   123 *Standard Dutch*
dat hij [$_{VP_1}$ beloofde$_1$ [$_{VP_2}$ **te** proberen$_2$ [$_{VP_3}$ het boek **te**
that he   promised   to try.INF   the book to
lezen$_3$]]]
read.INF
'that he promised to try to read the book'

However, as we will see in the following subsection, a prefix analysis turns out to be too simple for certain configurations in German.

3.2.   Against Prefix Status: Misplaced *zu* in Standard German

In most cases – like those discussed in the previous subsection –, the infinitival particle indeed surfaces before the verb where one expects it to surface, viz. on the hierarchically highest verb of the selected *zu*-infinitive complement. However, there are two instances in the standard language where *zu* seems to occur in the wrong position, that is, to be misplaced. Both involve verb clusters that are partially ascending, i.e. involve 132 or 312 order. The placement of *zu* in non-finite 132 orders was first discussed in Bech (1983). In the following example, the conjunction *ohne* 'without' selects a *zu*-infinitive. The infinitival complement contains a verb cluster with 132 order. Consequently, one would expect the hierarchically highest verb of the cluster, viz. V_1, to be preceded by *zu*. However, this structure is ungrammatical. Instead, *zu* surfaces before the last element of the cluster, viz. V_2. In the following triple from Standard German, the first example involves *ohne* with a finite complement. The second one is the non-finite version of it with *zu* in the position expected on the basis of its morphosyntactic properties. The third example shows misplaced *zu* before V_2 (data from Haider 2011: 227):

(11)   a.   132 V_1 = finite *Standard German*
ohne   dass er es mich [hat$_1$ prüfen$_3$ lassen$_2$]
without that he it me   has   verify.INF let.INF
lit.: 'without that he let me verify it'
b.   132 V_1 = non-finite *Standard German*
*ohne   es mich [**zu** haben$_1$ prüfen$_3$ lassen$_2$]
without it me   to   have.INF verify.INF let.INF
'without having let me verify it'

   c.   132 V_1 = non-finite *Standard German*
       ?ohne   es mich [haben$_1$ prüfen$_3$ **zu** lassen$_2$]
       without it me   have.INF verify.INF to let.INF
       'without having let me verify it'

The second case of misplaced *zu* involves what Vogel (2009: 308) referred to as the 'scandal construction'. It involves a non-finite verb cluster with 312-order where *zu* again appears before V_2 instead of V_1. In the following example, the matrix verb 'regret' selects a complement with a *zu*-infinitive:[7]

(12)   a.   312 *Standard German*
       ?Er bedauert, es nicht [verhindert$_3$ haben$_1$ **zu** können$_2$].
       He regrets   it not   prevent.PRT have.INF to can.INF
       'He regrets not having been able to prevent it.'
   b.   312 *Standard German*
       *Er bedauert, es nicht [verhindert$_3$ **zu** haben$_1$ können$_2$].
       He regrets   it not   prevent.PRT to have.INF can.INF
       'He regrets not having been able to prevent it.'

Clearly, if *zu* is treated as a prefix, the two cases where it is misplaced cannot be derived. There is no agreement in the literature as to the grammatical status of misplaced *zu*. It is generally acknowledged that the constructions are somewhat marked, but beyond this one can find diametrically opposed views. While Vogel (2009) considers the constructions to be the result of rules of grammar, Bech (1983) treats them as not fully grammatical compromises: they cannot be fully grammatical because they are in conflict with independent principles of German grammar: *zu* should be placed on the hierarchically highest verb; furthermore, the particle has to occur before the last element of the cluster. However, in Aux-Mod-Inf clusters ('has want INF'), the structurally highest verb has to undergo inversion; as a consequence, either *zu* is not prefinal or the hierarchially highest verb fails to bear *zu*. According to Bech, there is no way to resolve this conflict without violating some constraint of German grammar.

---

[7]Another peculiarity of this construction is the participial morphology on V_3. Since V_2 is a modal verb, one would expect V_3 to occur in the bare infinitive (which is, in fact, a grammatical alternative). What seems to have happened is the following: the participial morphology required by V_1 is not realized on V_2 (it actually never is in these clusters; instead, the so-called *Infinitives Pro Participio* occurs); instead, it is displaced to V_3. Similarly, the non-finite *zu* expected to occur on V_1 is displaced to V_2; we are thus dealing with two instances of misplaced morphology.

This view may explain the markedness of (11-c) and (12-a), but it does not explain why (11-b) and (12-b) are completely unacceptable. Haider (2011) goes a step further and regards the constructions as grammatical illusions, i.e. the constructions are essentially ungrammatical but appear to be acceptable. I will not take a definitive stand on this issue with respect to Standard German because it cannot be decided so easily on theoretical or empirical grounds. But as we will see in the next section, once non-standard varieties are taken into account, there is reason to believe that the possibility of misplaced *zu* is part of German grammar even if it only surfaces very residually in the standard language.[8]

3.3.   Misplaced *z* in Alemannic Varieties of German

While misplaced *zu* in Standard German seems to be a somewhat marked phenomenon of unclear grammatical status, the empirical situation is different in Alemannic varieties of German: Even though infinitives are less common than in the standard language (and prepositional, finite or non-subordinate structures being used instead), misplaced *zu* (whose form is *z* in these dialects) is nevertheless more visible. This is related to the fact that ascending orders in verb clusters are much more common in Alemannic, and especially Swiss German varieties. As a consequence, the conflict between marking V_1 or the last verb of the verb cluster obtains much more frequently, especially in relatively simple 2-verb clusters. In all ascending orders the particle *z* systematically appears (misplaced) before the last verb of the verb cluster. Examples can be found on the internet but also in traditional grammatical descriptions (suggesting that the phenomenon is definitely not just an invention of the formal grammarian). [9] The last example of the following triple was tested in an informal survey with native speakers of various Swiss German dialects: [10]

---

[8]It should be stressed that the existence of misplaced and displaced morphology in West Germanic languages is beyond doubt given the observations in Höhle (2006) and den Dikken and Hoekstra (1997).

[9]The phenomenon is also discussed in Hodler (1969: 560), Bader (1995: 22), and Cooper (1995: 188f.), who provides a number of examples recorded from Swiss German radio.

[10]Because of the general preference for alternative constructions, the acceptability of misplaced *z* examples will invariably be degraded; furthermore, unlike in finite clusters, non-finite ascending clusters with only two verbs are usually judged much more acceptable than more complex ones. Why this should be the case is a question I have to leave for further research.

(13)   a.   1...23 *Zurich German*, cf. Weber (1987: 244,fn.1)
           Er schiint$_1$ nüüt   [wele$_2$ z wüsse$_3$] dervoo.
           He seems   nothing want.INF to know.INF about.it
           'He does not seem to be interested in it.'

       b.   12 *Swiss German*[11]
           Ich liebe d   freiheit, selber de tag [chöne$_1$ z bestimme$_2$].
           I   love the freedom self   the day can.INF to determine.INF
           'I love the freedom to determine my schedule.'

       c.   132 *Swiss German*
           er behauptet, s   Buech bis am   Mëëntig [müse$_1$ gläse$_3$
           he claims   the book till on.the Monday must.INF read.PRT
           z ha$_2$]
           to have.INF
           'he claims having to have read the book until Monday'

Additionally, a fact that to my knowledge has gone unnoticed so far, misplaced
z is not limited to constructions where the verbal elements are adjacent, as in
the examples discussed up to now (such constructions are also referred to as
instances of Verb Raising, VR). It also occurs with so-called Verb Projection
Raising (VPR), where the verbal elements are separated by non-verbal material:

(14)   1X2 *Swiss German*[12]
       ohni   mi [welle$_1$ [uf d   bullesite z stelle$_2$]], im   gegeteil,
       without me want.INF on the cops.side to put.INF   on.the contrary
       aber ...
       but
       'without wanting to side with the cops, on the contrary, but ...'

Misplaced *z* is thus arguably less marked than in the standard language and I
will regard it as a phenomenon that is part of the grammar of these varieties.
For the purpose of this paper, I will assume that misplaced *zu* in Standard
German is a grammatical phenomenon as well.

---

[11]http://badoo.com/de-ch/0279246484/, found on March 11, 2013
[12]http://www.fcbforum.ch/forum/showthread.php?4328-usschritige-nachem-spiel-!/page4;
found on March 11, 2013

Taking the misplaced *zu* facts into account, the descriptive generalization for the placement of *zu/z* is thus as follows: *zu/z* occurs before the last verbal element of a non-finite complement required to be marked with *zu/z*.[13]

3.4.   Proposal

I will pursue two goals in this section: First, I will derive the cross-linguistic variation between German and Dutch in a systematic way. Second, I will provide a coherent account of *zu-/z*-placement in German and its varieties.

3.4.1.   *Dutch vs. German*

The goal of this subsection is to show that the operation that associates the particle with the verb is a similar operation in both languages. The cross-linguistic differences result from the fact that the operation takes place at different stages of the PF-component so that different concepts of headedness play a role: while it is structural headedness in the case of Dutch, it is peripherality within a constituent in the case of German (cf. Embick and Noyer 2001: 562 for these notions of headedness).

3.4.1.1.   *German*

I will make the following assumptions for German: In line with recent work on the role of linear order in syntax, I assume that linearization takes place post-syntactically. More concretely (and more generally), complements of V, both verbal and non-verbal, are linearized to the left of the selecting head.[14] If no further operations apply, the left-branching structure reaches the surface. In verb cluster constructions, PF-operations may alter the syntactic structure, sometimes leading to an ascending order. Second, *zu* occupies a functional head above the VP which for reasons of simplicity I will label F.[15]

As a first step, I will tackle the placement of *zu* in a simple example like (9-a), repeated here for convenience:

---

[13]For two rare exceptions, cf. Schallert (2012: 252).

[14]I deviate from Kayne (1994) in that the order of head and complement is established by means of linearization parameters, as e.g. in Richards (2008).

[15]I refrain from glossing it as I/T as in older work because the evidence for a separate IP/TP in German is rather scarce; furthermore, *zu* is also present in structures that are arguably smaller than IP/TP, cf. section 3.4.2 below.

(15) 321 *Standard German*
Er dachte, das Buch [lesen$_3$ können$_2$ **zu** müssen$_1$].
he thought the book read.INF can.INF to must.INF
'He thought he had to be able to read the book.'

The starting point will be a left-branching VP-structure with *zu* occupying a functional head above the highest verbal projection; for expository purposes, I will assume that all material that occurs outside the verb cluster has scrambled to Specv; in the left-branching VR-cases, this is optional (the surface string is usually ambiguous between a scrambled and a non-scrambled structure unless there is additional material like adverbials); in the ascending structures to be dicussed below it is obligatory to evade VP-inversion:

(16) 321 *Standard German*
Er dachte, [$_{vP}$ das Buch [$_{FP}$ [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ t$_{das\ Buch}$ lesen$_3$]
he thought the book read.INF
können$_2$] müssen$_1$] **zu**]].
can.INF must.INF to
'He thought he had to be able to read the book.'

In a next step, *zu* is associated with the final verb of the verb cluster *müssen* 'must'. At this point, there are two possibilities: Either *zu* undergoes Lowering in the sense of Embick and Noyer (2001), i.e. it undergoes downward head-movement. Alternatively, the reordering takes place at some later stage and involves linear reordering as in Local Dislocation. Before we can decide between these two options, we need to look at ascending structures as in the examples with misplaced *zu* in (11-c), (13-b) and (14), repeated here for convenience:

(17) a. 132 V$_1$ = non-finite *Standard German*
?ohne es mich [haben$_1$ prüfen$_3$ **zu** lassen$_2$]
without it me have.INF verify.INF to let.INF
'without having let me verify it'

b. 12 *Swiss German*
Ich liebe d freiheit, selber de tag [chöne$_1$ **z** bestimme$_2$].
I love the freedom self the day can.INF to determine.INF
'I love the freedom to determine my schedule.'

    c.   1X2 *Swiss German*

          ohni   mi [welle$_1$ [uf d   bullesite **z** stelle$_2$]], im

          without me want.INF on the cops.side to put.INF   on.the

          gegeteil, aber …

          contrary but

          'without wanting to side with the cops, on the contrary, but …'

The input for the reordering operations will look as follows:

(18)   a.   132, ex. (11-c) *Standard German*

          ohne   [$_{vP}$ es mich [$_{FP}$ [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ t$_{es}$ t$_{mich}$ prüfen$_3$]

          without   it me   verify.INF

          lassen$_2$] haben$_1$] zu]]

          let.INF   have.INF to

    b.   12, ex. (13-b) *Swiss German*

          … [$_{vP}$ de Taag [$_{FP}$ [$_{VP_1}$ [$_{VP_2}$ t$_{de\ Taag}$ bestimme$_2$]   chöne$_1$] z]]

              the day   determine.INF can.INF to

    c.   12, ex. (14) *Swiss German*

          ohni   [$_{vP}$ mi [$_{FP}$ [$_{VP_1}$ [$_{VP_2}$ t$_{mi}$ uf d   Bullesiite stelle$_2$]

          without   me   on the cops.side put.INF

          wele$_1$]   z]]

          want.INF to

I will assume that ascending structures come about by means of VP-inversion at PF (as e.g. in Haegeman and van Riemsdijk 1986, Williams 2004), i.e. V$_1$ inverts with VP$_2$. If all lexial material is scrambled out of the lexical VP, ascending Verb Raising structures obtain (cf. Broekhuis 1993). If the non-verbal material does not move, it is affected by VP-inversion so that Verb Projection Raising structures obtain. The three examples above are thus transformed into the following structures:

(19)   a.   132, ex. (11-c) *Standard German*

          ohne   [$_{vP}$ es mich [$_{FP}$ [$_{VP_1}$ haben$_1$ [$_{VP_2}$ [$_{VP_3}$ t$_{es}$ t$_{mich}$

          without   it me   have.INF

          prüfen$_3$] lassen$_2$]] zu]]

          verify.INF let.INF   to

b. 12, ex. (13-b) *Swiss German*

... [$_{vP}$ de Taag [$_{FP}$ [$_{VP_1}$ chöne$_1$ [$_{VP_2}$ t$_{de\ Taag}$ bestimme$_2$]] z]]
      the day   can.INF   determine.INF to

c. 12, ex. (14) *Swiss German*

ohni   [$_{vP}$ mi [$_{FP}$ [$_{VP_1}$ wele$_1$ [$_{VP_2}$ t$_{mi}$ uf d Bullesiite
without   me   want.INF   on the cops.side

stelle$_2$]] z]]
put.INF to

The following tree diagrams illustrate the inversion operation for (11-c):

(20)   Structure of (11-c) before VP-inversion:
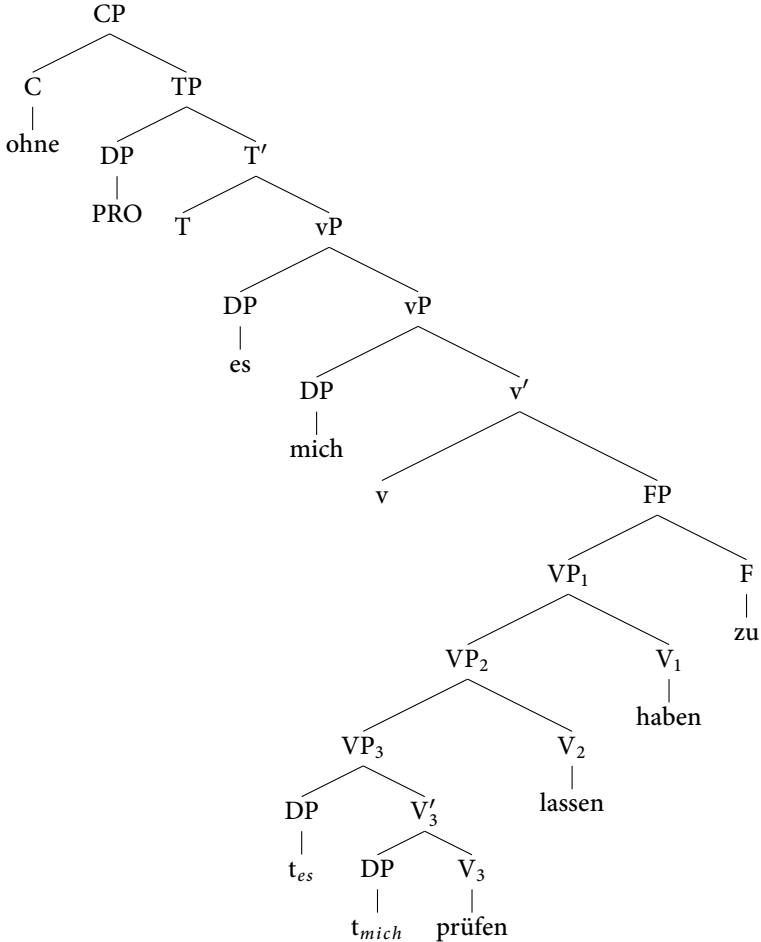
(21)   Structure of (11-c) after VP-inversion:

```
                        CP
                    /        \
                  C            TP
                  |          /    \
                ohne      DP        T'
                          |       /    \
                         PRO    T        vP
                                       /    \
                                    DP        vP
                                    |       /    \
                                    es    DP        v'
                                          |       /    \
                                        mich     v        FP
                                                        /      \
                                                     VP₁         F
                                                    /    \       |
                                                  V₁      VP₂     zu
                                                  |      /    \
                                                haben  VP₃      V₂
                                                      /   \      |
                                                    DP     V'₃  lassen
                                                    |     /   \
                                                   t_es  DP     V₃
                                                         |      |
                                                       t_mich prüfen
```
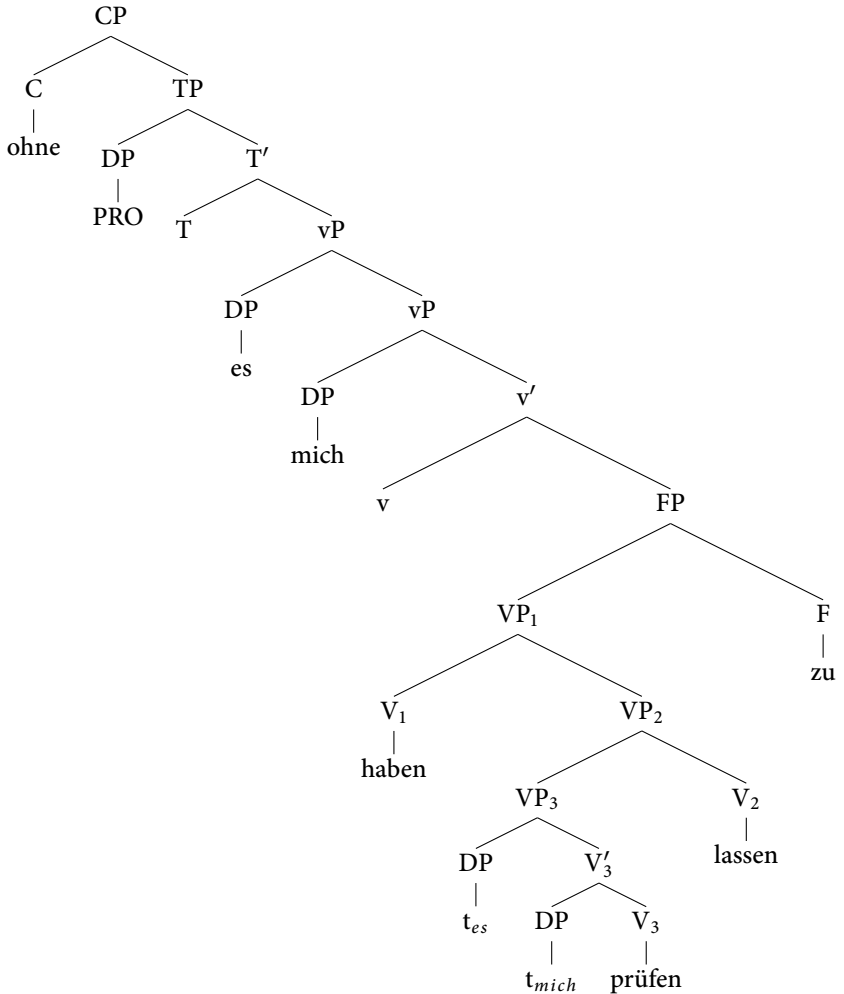
In the next step, *zu*-placement applies. Since *zu* does not end up on the head of VP$_1$, viz. *haben* 'have', but on the rightmost element of the VP, viz. V$_2$ *lassen* 'let', it must be an operation that is sensitive to linear order and adjacency, viz. a late PF-operation like Local Dislocation that affixes a head onto another one and inverts the two.

The operation can be sketched as follows for the three examples under discussion (I retain the VP-brackets for purposes of illustration, but it should

be noted that there is no [full] hierarchical structure anymore at this point of the derivation; furthermore, the trace of *zu* is only present for purposes of illustration):

(22) a. $[_{FP} [_{VP_1} V_1 [_{VP_2} [_{VP_3} V_3] V_2]]$ zu]
          $\Rightarrow [_{FP} [_{VP_1} V_1 [_{VP_2} [_{VP_3} V_3]$ **zu**+$V_2]] t_{zu}]$
     b. $[_{FP} [_{VP_1} V_1 [_{VP_2} V_2]]$ z]
          $\Rightarrow [_{FP} [_{VP_1} V_1 [_{VP_2}$ **z**+$V_2]] t_{zu}]$
     c. $[_{FP} [_{VP_1} V_1 [_{VP_2}$ PP $V_2]]$ z]
          $\Rightarrow [_{FP} [_{VP_1} V_1 [_{VP_2}$ PP **z**+$V_2]] t_{zu}]$

Zu-cliticization will proceed in the same fashion in (9-a). Since both constructions involve (at least partially) descending orders, there will also be string-vacuous cluster formation. The relative ordering between cluster formation and *zu*-cliticization is discussed in section 3.4.2.

Zu thus shows the behavior of a clitic. It must be stressed, however, that *zu* should be classified as an affix, viz. a phrasal affix (the terminology is somewhat confusing in this area): *zu* has selectional restrictions: it can only occur before verbs in the bare infinitive. This was guaranteed in all the examples discussed so far. There are cases, however, where the last verb of the cluster is a participle, as e.g. in Aux-Part clusters that show ascending order in Western dialects of Switzerland. Interestingly, while the ascending order is unproblematic in finite contexts, it is unacceptable in non-finite contexts – irrespective of the position of *zu*. Instead, only the descending order is acceptable in non-finite contexts (data from Raffaela Baechler, p.c.):

(23) a. 12/21; *Swiss German, Western dialects*
          das er s   Buech hät$_1$ gläse$_2$/ gläse$_2$ hät$_1$
          that he the book   has read.PRT read.PRT has
          'that he read the book'
     b. 12; *Swiss German, Western dialects*
          *ohni   s   Buech ha$_1$ z gläse$_2$
          without the book   have.INF to read.PRT
          'without having read the book'
     c. 12; *Swiss German, Western dialects*
          *ohni   s   Buech z ha$_1$ gläse$_2$
          without the book   to have.INF read.PRT
          'without having read the book'

    d.   21; *Swiss German, Western dialects*
          ohni   s   Buech gläse$_2$ **z**   ha$_1$
          without the book   read.PRT to have.INF
          'without having read the book'

The ungrammaticality of example (23-c) is unsurprising since *zu* does not occur
before the last element of the verb cluster. What is unexpected, though, is the
ungrammaticality of example (23-b). It makes perfect sense, though, if *zu* is
treated as an affix: since V$_2$ already bears a prefix, there is no space for another
affix. Furthermore, the selectional restrictions of *zu* would be violated. As a
consequence, only the descending order is grammatical in non-finite clusters.

3.4.1.2.   Dutch

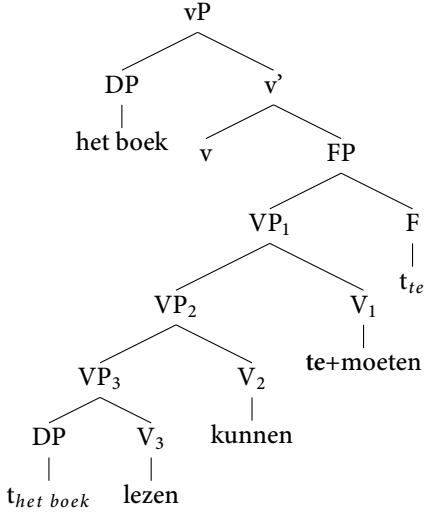I now turn to the derivation of the Dutch example in (9-b), repeated here for
convenience:

(24)   123 *Standard Dutch*
        Hij dacht   het boek [**te** moeten$_1$ kunnen$_2$ lezen$_3$].
        he thought the book to must.INF can.INF   read.INF
        'He thought he had to be able to read the book.'

The starting point will be the same as in German: The VP is left-branching with
verb clusters starting out as stacked VPs. The particle *te* occupies a functional
head above the highest VP and non-verbal material has scrambled from the
lexical VP to Specv:

(25)   [$_{vP}$ het boek [$_{FP}$ [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ t$_{het\ boek}$ lezen$_3$]   kunnen$_2$] moeten$_1$]
             the book   read.INF can.INF   must.INF
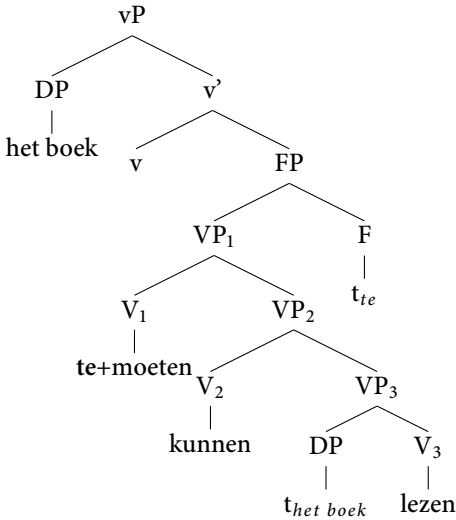   te]]
        to

Then, inversion and *te*-placement come into play. Since in Dutch the particle
is inverted together with V$_1$, *te*-placement has to precede inversion. Since
inversion involves sister nodes and thus hierarchical structure, *te*-lowering will
also apply at a point where the hierarchical structure is still available; therefore,
te undergoes downward head-movement and targets the head of the highest
VP, viz. V$_1$ *moeten*. This is illustrated in the following tree diagram:

(26)   *te*-Lowering in (9-b):

```
                        vP
                   ┌────┴────┐
                  DP         v'
                   |     ┌───┴───┐
                het boek v       FP
                             ┌───┴───┐
                           VP₁       F
                        ┌───┴───┐    |
                      VP₂       V₁   t_te
                   ┌───┴───┐    |
                 VP₃       V₂  te+moeten
                ┌──┴──┐    |
               DP     V₃  kunnen
                |     |
            t_het boek  lezen
```

Thereafter, VP-inversion takes place: V_1 inverts with VP_2 and V_2 inverts with VP_3, resulting in an ascending structure:[16]

(27)   VP-inversion in (9-b):

```
                        vP
                   ┌────┴────┐
                  DP         v'
                   |     ┌───┴───┐
                het boek v       FP
                             ┌───┴───┐
                           VP₁       F
                        ┌───┴───┐    |
                      V₁       VP₂   t_te
                       |    ┌───┴───┐
                   te+moeten V₂     VP₃
                             |    ┌──┴──┐
                          kunnen DP     V₃
                                  |     |
                             t_het boek  lezen
```

---

[16]The example in (10-b) will be derived similarly, with *te*-inversion applying twice and with inversion of V_1 with FP_1 and V_2 with FP_2.

The crucial difference between *te*-lowering and *zu*-cliticization is thus the point at which they apply: While *te*-lowering applies at an early stage of the PF-derivation where there is still full hierarchical structure, *zu*-cliticization applies at a later stage where the hierarchical structure has already been converted into a linear one. In the terms of Embick and Noyer (2001), *te*-lowering corresponds to Lowering while *zu*-cliticization corresponds to Local Dislocation. The cross-linguistic variation thus results from the fact that two operations – whose function is a similar one, viz. associate a verb with a marker of non-finiteness – apply at different points in the PF-component.

One may ask at this point why Dutch *te* is treated as an independent element in syntax at all and not as a prefix. As far as I can tell, such an analysis would certainly work for Standard Dutch since *te* is always adjacent to the verb it is supposed to mark. However, an analysis in terms of an independent syntactic element that undergoes lowering is more interesting if the contrast with German is to be described in a systematic way: Instead of treating the two particles as two completely different morphological objects (prefix vs. clitic), it seems more attractive to derive the difference by having the placement rules apply at different points of the derivation. Thereby, a common core can be captured: Both elements need to be associated with a non-finite verb. Furthermore, a look at other Dutch varieties and Afrikaans lends some support to this approach: In both languages, *te* is not lowered onto V_1; instead, it seems to remain an independent element and is inverted like the VPs. This is shown by the following examples involving a 231 order in the verbal complex: *te* does not associate with V_1 but occurs before the entire verbal complex (which in West Flemish contains non-verbal material):

(28)   a.   *West Flemish*, cf. Haegeman (1998: 635)
            mee Valere **te** [[willen$_2$ [dienen boek kuopen$_3$]] een$_1$]
            with Valere to want.INF that   book buy.INF   have.INF
            'with Valere having wanted to buy that book'
       b.   *Afrikaans*, cf. Donaldson (1993)
            Die banke moes   oop   gewees het, om dit gister   **te** [[kan$_2$
            the bank   should open been   have to it   yesterday to can.INF
            betaal$_3$] het$_1$].
            buy.INF have.INF
            'The bank should have been open to have been able to buy it yesterday.'

If *te* is treated as an independent element in syntax, its varied distribution can be accounted for in a straightforward manner: It undergoes lowering in some (Standard Dutch) but not in all varieties (West Flemish). Furthermore, lowering may apply at different points of the PF-derivation, which accounts for the contrast between (Standard) German and Standard Dutch.

3.4.2.   *Other Configurations in German*

In this subsection, I will discuss the placement of *zu* in Standard German in more detail. I will first determine the relative ordering of *zu*-placement and cluster formation before analyzing the placement of *zu* in the so-called 3rd construction.

3.4.2.1.   *The Ordering of Cluster Formation*

Based on the data in (3) I have been assuming that there is cluster formation in German if the verbal elements end up adjacent to each other in descending order. Importantly, this holds for all verbal elements in descending order in German and its dialects, i.e. also for (9-a), (10-a), for V_3 and V_2 in (11-c) and for V_3 and V_2 in (13-c). In all these constructions, extraposition cannot target the dependent VP, only the clause-final VP is a possible attachment site. [17] Cluster formation was argued to apply after syntax, viz. at some point of the PF-derivation. Since on my assumptions extraposition can block cluster formation, it must apply under adjacency and thus at a late stage of the PF-derivation. Since *zu*-cliticization was shown to apply at a late stage as well, questions arise with respect to the relative ordering of the two operations. The testing ground to decide this issue are examples like (10-a), repeated here, which contain clusters with two *zu*-infinitives:

---

[17] With verbs in ascending orders, things are somewhat more complicated: In ascending Swiss German structures, there can in principle be non-verbal material between the verbs (i.e. VPR is always a possibility); consequently, there cannot be a general surface constraint requiring string-vacuous cluster formation in that case. The same goes for partially ascending structures in German like (11-c) which also permit residual VPR (i.e. non-verbal material beween V_1 and V_3, cf. Bader and Schmid 2009: 224f.). In Dutch, however, ascending structures are almost completely impenetrable; most interveners can be considered incorporated X^0-elements. Consequently, a surface constraint requiring string-vacuous cluster formation would lead to the correct result and may in fact act as a trigger for the evacuation of the verb cluster.

(29)   321, (10-a) *Standard German*
dass er [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ das Buch **zu** lesen$_3$]   **zu** versuchen$_2$]
that he   the book to read.INF to try.INF
versprach$_1$]
promised
'that he promised to try to read the book'

The starting point, before the PF-operations apply, will be the following left-branching structure (again, I assume for the sake of concreteness that non-verbal material has scrambled to SpecV, even though this is optional in descending structures):

(30)   321, (10-a) *Standard German*
dass er [$_{vP}$ das Buch [$_{VP_1}$ [$_{FP_1}$ [$_{VP_2}$ [$_{FP_2}$ [$_{VP_3}$ t$_{das\ Buch}$ lesen$_3$]   **zu**]
that he   the book   read.INF to
versuchen$_2$] **zu**] versprach$_1$]]
try.INF   to   promised
'that he promised to try to read the book'

Since verb cluster formation is based on adjacency, it seems that *zu*-cliticization has to take place *before* cluster formation:

(31)   a.   *zu*-cliticization
dass er das Buch [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ **zu**+lesen$_3$]   **zu**+versuchen$_2$]
that he the book   to+read.INF to+try.INF
versprach$_1$]
promised
       b.   cluster formation
dass er das Buch [$_V$ [$_V$ **zu**+lesen$_3$]+ [$_V$ **zu**+versuchen$_2$]+ [$_V$
that he the book   to+read.INF   to+try.INF
versprach$_1$]]
promised

Unfortunately, this ordering seems to lead to a contradiction once the 'scandal construction' (12-a) is taken into account. I repeat the relevant example:

(32)   312, (12-a) *Standard German*
        Er bedauert, es nicht [verhindert$_3$ haben$_1$ **zu** können$_2$].
        He regrets   it not   prevent.PRT have.INF to can.INF
        'He regrets not having been able to prevent it.'

The starting point will be the following structure:

(33)   ... es nicht [$_{FP}$ [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ t$_{es}$ verhindert$_3$] können$_2$] haben$_1$] **zu**]
        it not   prevent.PRT have.INF can.INF to

Since *zu*-cliticization was assumed to precede cluster formation to derive the
correct result for (10-a) = (29), one would expect *zu* to surface on V$_1$, contrary to
fact. Rather, the facts suggest that reordering between V$_1$ and V$_2$ must precede
zu-cliticization. However, this reordering cannot be an instance of VP-inversion
because the two verbs are not sisters – VP-inversion would incorrectly place V$_1$
at the beginning of the cluster (string vacuous movement of VP$_3$ to some higher
position seems ill-motivated). Consequently, the reordering between V$_1$ and V$_2$
must be the result of a different process. The obvious choice is cluster formation
under adjacency. While the cases of cluster formation we have looked at so far,
cf. (3), (9-a) and (10-a) were all string-vacuous, this one involves reordering. In
fact, it is arguably the same kind of process as *zu*-cliticization. To derive the
correct result for the scandal construction, reordering cluster formation has to
take place *before zu*-cliticization:

(34)   a.   reordering/cluster formation (12-a)
              ... es nicht [[verhindert$_3$] haben$_1$+können$_2$] **zu**
                it not   prevent.PRT have.INF+can.INF to
         b.   *zu*-cliticization (12-a)
              ... es nicht [[verhindert$_3$] haben$_1$+**zu**+können$_2$]
                it not   prevent.PRT have.INF+to+can.INF

It seems that we have arrived at an impasse: To derive (10-a) = (29), we need to
assume that cluster formation follows *zu*-cliticization while the reverse ordering
is necessary to derive (12-a) = (32). There are two possibilities to resolve the
contradiction.

    First, since the two cases of cluster formation differ in that one is string-
vacuous while one involves reordering, one could classify them as two in-
dependent operations that can be ordered differently with respect to other

operations.[18] Concretely, one has to assume that reordering cluster formation can precede *zu*-cliticization, which in turn precedes string-vacuous cluster formation. The resulting order thus looks as follows:

(35)   reordering cluster formation > *zu*-cliticization > string-vacuous cluster
        formation

Cluster formation is an instance of Local Dislocation in both cases as both are sensitive to linear order. As pointed out above, *zu*-cliticization is also an instance of this type of PF-operation.

    Second, there is an alternative to resolve the contradiciton that avoids extrinsic ordering: If we assume that the bracketed structure in (29) (= (10-a)) and (32) (= (12-a)) is still available at the point where the PF-operations apply, the placement facts follow under cyclicity if the PF-derivation unfolds bottom-up: In (32), verb cluster reordering will first produce $[3[1+2]zu]$, then, *zu* is inverted with V_2, correctly deriving $[3[1+zu+2]]$. In (29), the lower *zu* (F_2) is first cliticized onto V_3: $[zu_2+V_3]$. Then, $[zu_2+V_3]$ undergoes string-vacuous cluster formation with V_2, resulting in $[zu_2+V_3+V_2]$. Then, the higher *zu* (F_1) cliticizes onto V_2, resulting in $[zu_2+V_3+zu_1+V_2]$. Finally, the entire complex undergoes string vacuous cluster formation with V_1: $[zu_2+V_3+zu_1+V_2+V_1]$. The second solution has the advantage that it avoids extrinsic ordering. Furthermore, the full derivation of (34) will additionally involve string-vacuous cluster formation between V_3 and the complex $[V_1+V_2]$ because 312 clusters behave like complex heads (nothing may intervene between V_3 and V_1). If the ordering is as in (35), cyclicity would be violated. A fully cyclic derivation that avoids rule ordering, however, can derive the correct result. Finally, the cyclic derivation treats both instances of cluster formation as the same operation. Given these advantages, I opt for the second solution. [19]

---

[18] This is a somewhat unsatisfactory move since it seems straightforward to treat string-vacuous cluster formation as a subcase of reordering cluster formation, the latter simply being more complex in involving an additional operation.

[19] An alternative to both solutions would be to derive the impenetrability of descending clusters without cluster formation as e.g. in Wurmbrand (2007) where extraposition is sensitive to prosodic principles. In that case, *zu*-cliticization can be ordered after cluster formation without having any detrimental consequences.

3.4.2.2.   *zu-Placement in the Third Construction/Extraposition*

I now turn to *zu*-infinitival complements with ascending order in (Standard) German. There are generally two types of *zu*-infinitives irrespective of whether they occur intraposed, i.e. in descending order, or whether they occur in postverbal (= extraposed) position, i.e. in ascending order. The two types differ with respect to their integration into the matrix clause: Some *zu*-complements are so small in size that they form a monoclausal unit together with the matrix clause for clause-bound processes like scrambling or weak pronoun fronting, i.e. they show restructuring/transparency effects. They are usually treated as VPs, as has been presupposed for (10-a) (= (29)), and are usually referred to as restructuring infinitives. Other complements do not show any of the transparency effects and are therefore usually treated as CPs; they are termed non-restructuring infinitives. Whether a complement is restructuring or non-restructuring usually depends on the selecting predicate. Some select only restructuring infinitives (e.g. *scheinen* 'seem' in Standard German), others select only non-restructuring infinitives (e.g. *bedauern* 'regret'), and a third class can select either type of complement (e.g. *versuchen* 'try'). This classification is also found when the *zu*-infinitive occurs in post-verbal or extraposed position. When a restructuring infinitive occurs in postverbal (= extraposed) position, the construction is referred to as the *3rd construction*. Both constructions are illustrated by the following examples (scrambling of the object pronoun in example (36-b) indicates transparency; the same operation would lead to ungrammaticality in example (36-a)):

(36)   a.   CP-complement, *Standard German*
         ohne   [$_{VP_1}$ zu bedauern$_1$ [$_{CP}$ mich zu mögen$_2$]]
         without   to regret.INF   me   to like.INF
         'without regretting to like me'
     b.   VP-complement/3rd construction, *Standard German*
         ohne   mich [$_{VP_1}$ zu versuchen$_1$ [$_{VP_2}$ t$_{mich}$ zu mögen$_2$]]
         without me   to try.INF   to like.INF
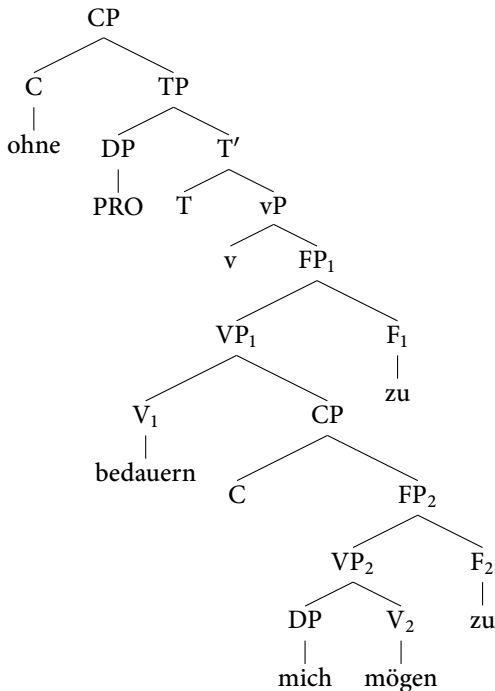   'without trying to like me'

An obvious question is how the ascending orders are derived. There are two possibilities: Either they are the result of PF-inversion – like VPR-structures – or they are derived by extraposition, i.e. movement to the right. In what follows,

I will use the placement of *zu* as a diagnostic to determine which of the two options is correct. The starting point will again be a descending order:
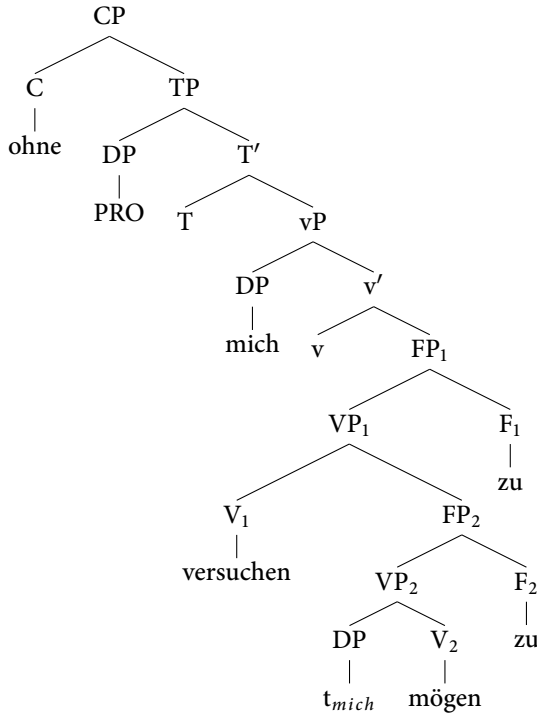
(37)   a.   CP-complement
             ohne   [$_{FP_1}$ [$_{VP_1}$ [$_{CP}$ [$_{FP_2}$ [$_{VP_2}$ mich mögen$_2$] zu]] bedauern$_1$]
             without   me   like.INF to   regret.INF
             zu]
             to
       b.   VP-complement/3rd construction
             ohne   [$_{vP}$ mich [$_{FP_1}$ [$_{VP_1}$ [$_{FP_2}$ [$_{VP_2}$ mögen$_2$] zu] versuchen$_1$]
             without me   like.INF to   try.INF
             zu]]
             to

Suppose ascending structures are derived by means of PF-inversion of V_1 with CP/FP_2; the result is indicated in the following tree diagrams:

(38)   a.   non-restructuring predicate, ex. (36-a)
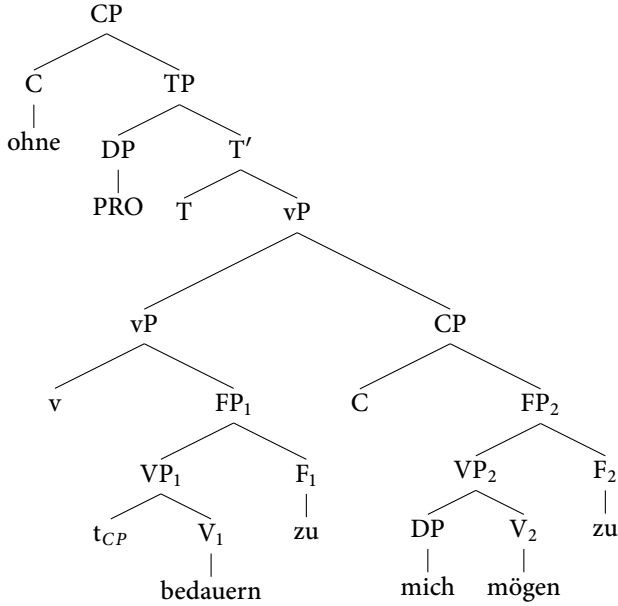
b.   restructuring predicate, ex. (36-b)

```
                         CP
                    ┌─────┴──────┐
                    C            TP
                    │        ┌────┴────┐
                  ohne      DP         T′
                            │      ┌────┴────┐
                           PRO     T         vP
                                        ┌────┴────┐
                                       DP         v′
                                        │     ┌────┴────┐
                                      mich     v       FP₁
                                           ┌─────────────┴───────┐
                                          VP₁                    F₁
                                     ┌─────┴──────┐              │
                                     V₁          FP₂            zu
                                     │       ┌────┴────┐
                                 versuchen  VP₂        F₂
                                        ┌────┴────┐     │
                                       DP        V₂    zu
                                        │         │
                                     t_mich    mögen
```

If *zu* then cliticizes onto the right-most verb, an ungrammatical result obtains: both *zu*s are cliticized onto the the last verb and V_1 fails to bear *zu* (or, alternatively, the two *zu*s are reduced to one by haplology and there is only one *zu*, which occurs on the right-most verb of the complement):

(39)   a.   CP-complement
            *ohne   bedauern₁ mich [**zu**+**zu**+mögen₂]
            without regret.INF me   to+to+like.INF
       b.   VP-complement/3rd construction
            *ohne   mich versuchen₁ [**zu**+**zu**+mögen₂]
            without me   try.INF   to+to+like.INF

Consequently, ascending structures must be derived by means of movement to the right (= extraposition) of CP/FP_2, as indicated in the following tree

structures (For the sake of concreteness, I am assuming adjunction to matrix vP, but a different/higher position would also be conceivable):[20]

(40) a. non-restructuring predicate, ex. (36-a)

```
                        CP
                   ┌────┴────┐
                   C         TP
                   │      ┌───┴───┐
                 ohne    DP       T′
                         │     ┌──┴──┐
                        PRO    T     vP
                                 ┌───┴───────────┐
                                vP               CP
                             ┌───┴────┐       ┌───┴────┐
                             v       FP₁      C        FP₂
                                  ┌───┴──┐         ┌────┴───┐
                                 VP₁     F₁       VP₂       F₂
                               ┌──┴──┐   │      ┌──┴──┐     │
                              t_CP   V₁  zu     DP   V₂    zu
                                     │          │    │
                                  bedauern     mich mögen
```

---

[20] The present account presupposes that extraposition takes place in syntax. However, a PF-movement account of extraposition would also be possible as long as it precedes *zu*-cliticization.

b. restructuring predicate, ex. (36-b)
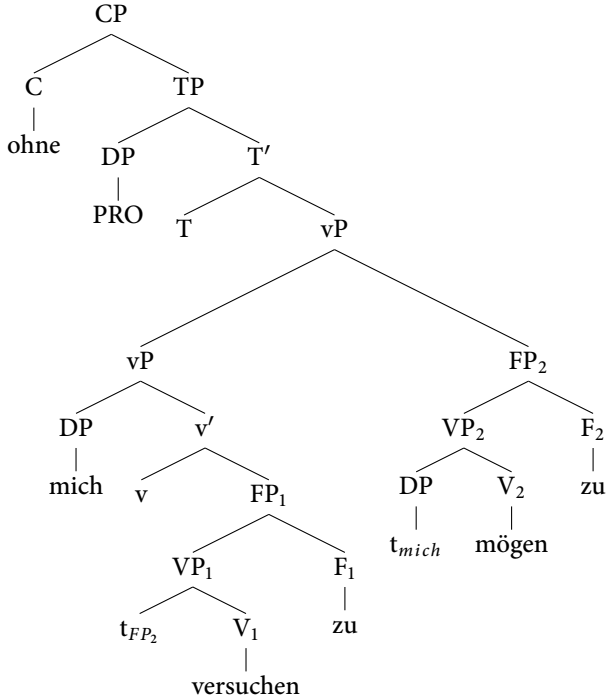
This provides the correct input for the *zu*-cliticization rule: After linearization, both *zu*s are adjacent to the verb which they are supposed to be associated with and cliticization is successful (again, the traces of *zu* are only employed for expository purposes):

(41) a. CP-complement

ohne $[_{vP} [_{vP} [_{FP_1} [_{VP_1} t_{CP} \textbf{zu}_1+\text{bedauern}_1] \ t_{zu1}]] [_{CP} [_{FP_2}$
without to+bedauern.INF
mich $\textbf{zu}_2+\text{mögen}_2 \ t_{zu_2}]]]$
me to+like.INF

b. VP-complement/3rd construction

ohne mich $[_{vP} [_{vP} [_{FP_1} [_{VP_1} t_{FP_2} \textbf{zu}_1+\text{versuchen}_1] \ t_{zu_1}]] [_{FP_2}$
without me to+try.INF
$[_{VP_2} \textbf{zu}_2+\text{mögen}_2] \ t_{zu_2}]]$
to+like.INF

Zu-placement thus provides new evidence that the 3rd construction in Standard German involves (remnant) extraposition and therefore has to be derived differently than VPR-structures.

Note that extraposition is also necessary for finite complement clauses as e.g. in the following example:

(42)   *Standard German*
ohne   zu glauben [$_{CP}$ dass Peter kommt]
without to believe   that Peter comes
'without believing that Peter will come'

If the complement CP starts as a left-hand complement of V and reaches its post-verbal position by means of inversion with V (or is directly linearized as a right-hand-complement, as is sometimes proposed even under left-branching approaches), one would expect *zu* to end up on the rightmost element of the complement clause, contrary to fact. Consider the following tree structure:

(43)   CP-complement without extraposition, ex. (42)
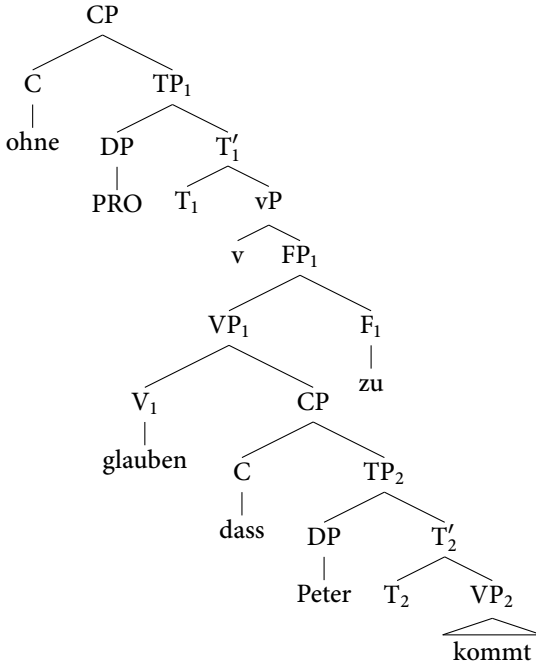
If *zu*-cliticization applied to this structure, the result would be ungrammatical:

(44)   *Standard German*
        *ohne   glauben, [$_{CP}$ dass Peter **zu** kommt]
        without believe   that Peter to comes
        'without believing that Peter will come'

Instead, the CP has to be extraposed (adjoined to matrix vP):

(45)
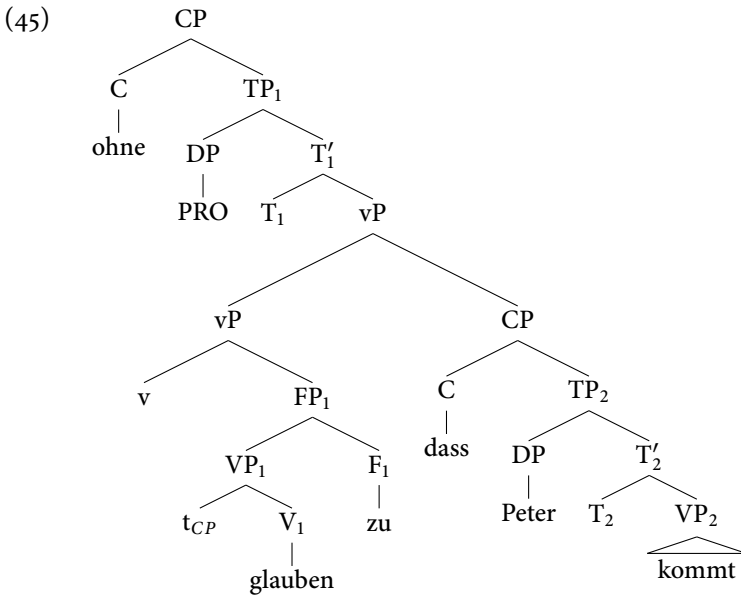
This provides the correct context for *zu*-cliticization: it now targets V_1.

Note incidentally that the *zu*-placement facts tend to argue against extraposition analyses of ascending/VPR-structures as e.g. proposed in Haegeman (1992): The basic idea is that the dependent VP is right-adjoined to the higher VP. Applied to a 3-verb cluster with 132 order, the structure would look as follows (since extraposition involves VP_2, there can be non-verbal material between V_1 and V_3, thereby deriving VPR-structures):

(46) $[_{VP_1} [_{VP_2} [_{VP_3} V_3] V_2] V_1] \Rightarrow [_{VP_1} [_{VP_1} t_{VP_2} V_1] [_{VP_2} [_{VP_3} V_3] V_2]]$

Suppose that this is a cluster within non-finite VP es e.g. in the misplaced *zu* examples in Standard German, cf. (11-c). If extraposition of VP_2 targets matrix vP as in the previous cases of extraposition, we obtain the following structure:

(47)
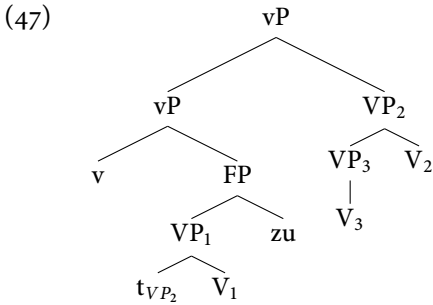
If we then apply *zu*-cliticization, *zu* ends up on V_1, as in the ungrammatical (11-b). This can only be avoided if extraposition targets a position below *zu*, i.e. involves adjunction to VP_1. Consequently, if all ascending/VPR-structures arise by means of extraposition, one has to assume that extraposition may target different nodes depending on whether a VP (adjunction to VP), or an FP/CP (adjunction to vP) is extraposed. In the approach pursued here, there are two mechanisms that derive ascending structures, VP-inversion and extraposition. It is not so easy to choose between these two options, but there is one piece of empirical evidence that argues that the latter approach is superior.

I will concentrate in what follows on the difference between the 3rd construction and VPR. They are similar in that both constitute monoclausal domains for processes like scrambling and pronoun fronting (vor VPR cf. e.g. Haegeman 1992: 110). Additionally, when an argument of the lexical verb is scrambled to a higher verbal projection, the movement does not show the hallmarks of regular scrambling (like inducing freezing effects and blocking focus projection), cf. Salzmann (2011), Geilfuss-Wolfgang (1991: 25f.). This may suggest that they should be derived in the same way. However, there is a striking asymmetry in the domain of scope: While scrambled elements in VPR-constructions can reconstruct, this does not seem to be possible in the 3rd construction (cf. Salzmann 2011: 454 for VPR and Bobaljik and Wurmbrand 2005: 810,831):

(48) a. dass er [$_{VP_1}$ 2 Manager wett$_1$ [$_{VP_2}$ t$_{2\ Manager}$ vo siine Idee
          that he   2 managers wants   of his   ideas
          überzüüge$_2$]]
          convince.INF
          'that he wants to convince two managers of is ideas'

$(2 >$ want; want $> 2)$

   b. weil   er [$_{VP_1}$ alle Fenster   vergass$_1$ [$_{VP_2}$ t$_{alleFenster}$ zu
       because he   all windows forgot   to
       schliessen$_2$]]
       close.INF
       'because he forgot to close all the windows' (all $>$ forget; *forget $>$ all)

Under an extraposition analysis of the 3rd construction, there is a straightfor-
ward explanation for the absence of reconstruction: What is extraposed is a
remnant VP. Importantly, remnant VPs have been shown to induce scope freez-
ing effects (Barss 1986: 517-542), as expressed in the following generalization:

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

This is exactly the configuration that obtains in remnant (VP-) movement: α
is A-moved out of a VP; VP is then A$'$-moved to a position above α so that
α no longer c-commands its trace β. An example pair illustrating the effect
in the domain of verb clusters is the following (slightly adpated from Haider
2003: 101):

(50) a. dass ihr   [$_{VP_1}$ niemand [$_{VP_2}$ t$_{niemand}$ zu beleidigen$_2$]
          that her.DAT   no.one   to insult.INF
          gelang$_1$]
          succeeded
          'that she managed to insult no.one'   ($\neg\exists >$ succeed; succeed $> \neg\exists$)

   b. [$_{VP_3}$ t$_{niemand}$ zu beleidigen$_3$] ist$_1$ ihr   [$_{VP_1}$ niemand [$_{VP_2}$
                      to insult.INF   is her.DAT   no.one
       t$_{VP_3}$ gelungen$_2$] t$_{ist}$].
              succeeded
       'that she managed to insult no.one'   ($\neg\exists >$ succeed; *succeed $> \neg\exists$)

It is immaterial at this point how the scope freezing effect is derived (cf. e.g. Sauerland and Elbourne 2002 for a recent proposal); descriptively, it seems that A-movement can in principle reconstruct, leading to ambiguity in intraposed verb cluster constructions like (50-a) and VPR (48-a), but not if remnant movement is involved as in remnant topicalization of part of a verb cluster (50-b). The lack of ambiguity in the third construction (48-b) then finds a straightforward explanation if it is analyzed as an instance of remnant extraposition as well.

To conclude this digression: The facts from *zu*-placement suggest that ascending structures in VPR and the 3rd construction are derived differently. This correlates with a scope asymmetry. While an approach solely based on extraposition cannot capture the scope differences in a natural way, similarities and differences fall out directly if it is assumed that the two constructions are the result of different mechanism, viz. VP-inversion vs. extraposition.[21]

3.5.   Ordering of Operations: Synopsis

I have discussed a number of operations that interact in intriguing ways. It turned out to be possible to determine a strict order between these operations without encountering any contradictions. They are summarized in the following table:

(51)   Ordering of operations (left-branching): synopsis

extraposition (RC, CP/FP$_2$) > topicalization (VP) >	syntax
te-lowering > VP-inversion > cluster formation (+/-inversion) > < *zu*-cliticization	PF

The ordering in syntax is intrinsic in that it follows from the Strict Cycle Condition/the Extension Condition. The ordering between *te*-lowering and VP-inversion is probably extrinsic. The ordering between cluster formation and *zu*-cliticization again follows from cyclicity (if that concept is adopted for PF). Their ordering with respect to VP-inversion is also intrinsic if it is assumed that

---

[21]If as discussed in section 3.7, a right-branching VP-structure is adopted, the two structures can be distinguished in that the 3rd construction involves extraposition while in VPR the basic linearization is retained.

hierarchy-sensitive operations take place in a different subcomponent of PF than adjacency-sensitive operations.

3.6. Previous Accounts

The previous sections have shown that an analysis in terms of rule ordering is successful in accounting for the placement of *zu*. But before concluding that this is the optimal solution, I would like to briefly discuss previous analyses of misplaced *zu*. They can be divided into syntactic/derivational accounts where *zu* is an independent syntactic element and morphological/realizational approaches where *zu* is just a feature of a non-finite complement that receives morphological expression according to specific rules. I will discuss the two types of approaches in turn.[22]

3.6.1. *Syntactic/Derivational Accounts of Misplaced zu*

The idea that cluster formation has to precede placement of *zu* can be found in a number of analyses. For instance, von Stechow (1990: 159) argues that *zu* is generated in INFL and incorporated into the verbal complex after re-analysis (which is taken to be the mechanism that generates complex heads and ascending orders, cf. Haegeman and van Riemsdijk 1986). The account seems to involve lowering/affix hopping of *zu*. It is explicitly assumed that this takes place before PF, which means it should be subject to syntactic locality conditions; as a consequence, one would expect *zu* to end up on V_1 and not on V_2. It seems that lowering is sensitive to adjacency in this account, but that seems implausible for a syntactic operation (quite apart from the fact that it violates the c-command constraint on incorporation). Arguably, the underlying intuition was the same as that for the rule of *zu*-cliticization proposed here, but given the framework of that time, a solution by means of a PF-operation was apparently not obvious. Whether this account can be extended to other cases of *zu*-placement and how it would deal with Dutch remains open. A somewhat different proposal is found in Sternefeld (1990: 251) who first argues that it is the rightmost verb that moves to INFL, where *zu* is base-generated. Since movement to INFL follows cluster formation, *zu* ends up on the correct verb

---

[22] A hybrid solution is proposed in Sternefeld (2006) who treats *zu* as a lexical feature of infinitives, which, however, can undergo migration to a different part of the cluster to derive cases of misplaced *zu* like (11-c).

(viz. V_2). But it remains mysterious why it is not the head of the verb cluster V_1 that moves to INFL as would be expected under a syntactic account where locality constraints apply (e.g. minimality). Furthermore, the account requires excorporation of V_2 from the cluster. Sternefeld also considers an incorporation solution as in von Stechow (1990) but points out that this raises problems for *te*-placement in Dutch as in (9-b) where incorporation would have to precede cluster formation/reordering. The issue is eventually left open, and it remains unclear to what extent the placement of the infinitival particle can be handled in a systematic way both within German and cross-linguistically. The solution that comes closest to the current proposal is the one by Hinterhölzl (2009: 208) who argues that *zu* is a phrasal affix and fuses with the adjacent infinitive at Morphological Form. This is clearly a PF-operation and the basic intuition is arguably the same as in the present account; unfortunately, the workings of the operation are not spelled out in much detail so that it is not clear to what extent it can be applied to other cases of *zu*-placement (for instance, Hinterhölzl only discusses misplaced *zu* in 132 clusters like (11-c) but does not address 312 clusters like (12-a)). Furthermore, it is not clear whether the cross-linguistic variation can be derived in a natural way.

3.6.2.   *Realizational Approaches*

In Bader (1995) and Vogel (2009), *zu* is treated as a phrasal affix/an instance of edge inflecton. It is not an independent syntactic head but rather a morpho-syntactic feature assigned to an infinitival complement. Its realization on the last element of the verb cluster is the result of special realizational rules (an EDGE-feature in Bader's HPSG-approach and an alignment constraint in Vogel's OT-account). Both approaches successfully account for misplaced *zu/z*: *zu/z* is realized on the rightmost terminal node of the verb cluster. This captures in a very different way the intuition that the position of *zu* depends on the surface order in the verbal complex. The question that arises, though, is whether this rule of *zu*-placement successfully accounts for other cases of *zu*-placement as well. There are no problems with a single verb cluster as in (9-a). But clearly, the realization of *zu* on some dependent element of a non-finite complement must be restricted. For instance, it has to be ruled out that *zu* is realized on the rightmost element of a verb cluster that is embedded under the non-finite complement, as e.g. in (42) above and in the following example from Vogel (2009: 329):

(52)   *Standard German*
    *Ich bin froh,   nicht haben   hören   **zu** müssen, dass du   dich
    I   am happy not   hav.INF hear.INF to must.INF that you self
    geärgert **zu** hast.
    annoyed to have.2SG
    'I am happy that I did not have to hear that you were annoyed.'

Vogel (2009: 329) proposes that *zu* appears on the right-most element of the extended projection marked for *zu*. This correctly rules out (52) since the embedded clause constitutes a separate domain with its own extended projection. Other cases are less trivial. Consider, for instance, the placement of *zu* in VPR-cases like (14), repeated for convenience:

(53)   1X2 *Swiss German*
    ohni   mi [welle$_1$ [uf d   bullesite **z** stelle$_2$]], im   gegeteil,
    without me want.INF on the cops.side to put.INF   on.the contrary
    aber ...
    but
    'without wanting to side with the cops, on the contrary, but ...'

Here, the complement of *welle* 'want' contains non-verbal material. In principle, VPR-complements can also contain external arguments and adverbials related to tense, which suggests that they are larger than bare VPs. There is no general consensus on their size apart from the fact that they are smaller than CPs but larger than VPs; for instance, based on West-Flemish data with expletive subjects within the VPR-complement, den Dikken (1996) proposes that they are TPs. Whether this constitutes a separate extended projection in Vogel's terms is hard to say. With respect to transparency effects like scrambling and pronoun fronting, VPR-complements tend to behave like monoclausal units, cf. e.g. Haegeman (1992: 110). This might indicate that VPR-complements do not constitute a separate domain, but note that this equates 'separate extended projections' with transparency effects and the presence of a CP. In other words, Vogel's approach predicts that *zu* must be realized within the CP containing the non-finite complement marked with *zu*. As pointed out in Haider (2011: 250), this may make the wrong prediction for restructuring infinitives like (10-a), repeated here for convenience:

(54)   321 *Standard German*
dass er [$_{VP_1}$ [$_{VP_2}$ [$_{VP_3}$ das Buch **zu** lesen$_3$]   **zu** versuchen$_2$]
that he   the book to read.INF to try.INF
versprach$_1$]
promised
'that he promised to try to read the book'

There is no doubt that the lower *zu*-phrase shows transparency effects; according to the reasoning above, this would suggest that it does not have any extended projections. It is not fully clear what Vogel's account would predict in this case. Since the structure is left-branching, one expects a *zu* on V$_1$ *versuchen* 'try' as it is the right-most element of the VP. The question is whether the embedded VP can be marked for *zu* at all if it does not have any functional projections above VP. Arguably, the realizational rule has to be adjusted: *zu* must be realized on the right-most element *within* the non-finite constituent marked vor *zu*, irrespective of whether the non-finite constituent bears any functional projections above VP. To rule out (52), one has to stipulate that *zu* has to be realized within the same CP that contains the non-finite XP marked for *zu*.[23] Another possible complication arises with the third construction as in (36-b) and the corresponding tree structure in (40-b). I repeat the final structure for convenience:

(55)   VP-complement/3rd construction
ohne   mich [$_{vP}$ [$_{vP}$ [$_{FP_1}$ [$_{VP_1}$ t$_{FP_2}$ **zu+versuchen**$_1$] t$_{zu_1}$]] [$_{FP_2}$ [$_{VP_2}$
without me   to+try.INF
zu+mögen$_2$] t$_{zu_2}$]]
to+like.INF

Here it crucially depends on which XP is assigned the feature *zu*. If it is the projection that the extraposed VP is adjoined to, viz. vP as in (40-b), one would probably expect *zu* to remain unrealized in VP$_1$ because the right-most terminal of the non-finite XP assigned *zu* would be the right-most verb of the extraposed FP$_2$. Perhaps this can be avoided by the restriction introduced above that *zu* has to be realized within the phrase marked for *zu*, under the

---

[23]This stipulation can perhaps be avoided if CP-complements are extraposed, as was assumed in (45) above. Unfortunately, Vogel does not discuss the structural position of CP-complements. However, as we will see below, extraposition may still not be sufficient.

assumption that the adjoined FP_2 is not sufficiently part of FP_2.[24] As should have become clear, once other *zu*-placement phenomena are taken into account, the intuitively simple solution proposed in Vogel (2009) has to be modified and restricted in rather specific ways to attain observational adequacy.

Examples like (55) are particularly interesting given the observations about so-called missing *z* in Swiss German in Bader (1995). I will discuss this phenomenon in a separate subsection.

3.6.3. *Missing* z *in Swiss German*

Bader (1995: 22,26) discusses Swiss German verb clusters that seem structurally very parallel to the 3rd construction data in the standard language. They also contain two non-finite phrases marked for *zu*. What is remarkable in this construction is that the *zu* assigned to the higher VP can go missing (*schiine* 'seem' and *probiere* 'try' require a *zu*-infinitive):

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

As opposed to the examples from the standard language like (36-b), there is no *zu* on V_2 *probiere* 'try'. Bader accounts for both missing *z* and misplaced *z* by means of a realizational rule that is very smilar to the one proposed in Vogel (2009). Although the technical details differ, the result is the same: *zu/z* is realized on the last element of the XP assigned/marked with *zu*. This accounts for misplaced *zu*.
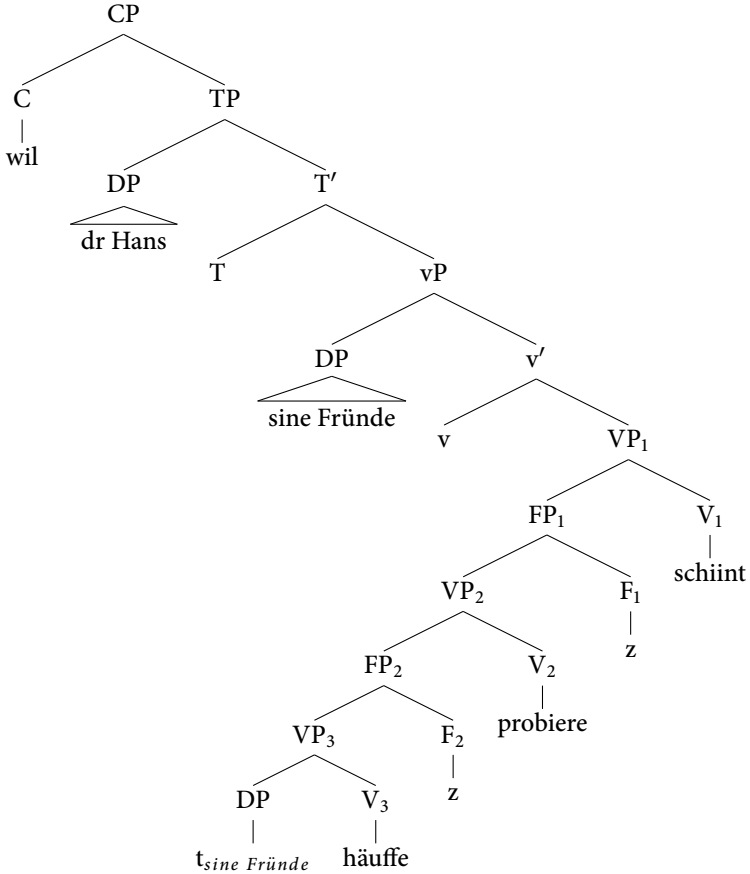
There is no provision in this system for preventing the feature from percolating downwards, and this is exactly what is exploited to account for missing *z*: In (56), the XP headed by *häuffe* 'help' is analyzed as a complement of *probiere* 'try'. The *zu*-feature assigned to the constituent headed by *häuffe* is, of course, realized on *häuffe*. Crucially, the *zu*-feature assigned to the constituent headed by *probiere* is also realized on *häuffe* because this is the rightmost element contained in that XP. In other words, missing *z* is not a separate phenomenon under this analysis, it is simply a side-effect of the workings of the realizational

---

[24]An obvious alternative consists in adjoining the extraposed VP to a higher node, but this would have to be motivated independently.

rule. This seems to be a very elegant solution, but problems arise in other areas of *zu*-placement (basically as in Vogel's approach). Since downwards percolation is in principle unlimited, it could also wrongly derive (52) where *zu* assigned to the matrix VP ends up on the right-most verb of the embedded VP. Again, downwards percolation has to be restricted somehow. While the left-branching cases with two *zu*s as in (54) can probably be handled (the EDGE-feature has to be realized at the right edge of every constituent bearing the feature), problems also arise with the 3rd construction in Standard German, cf. (36-b) where one might also expect a missing *z*, contrary to fact. It seems that Bader would have to resort to extraposition of both VPs and CPs, which is perhaps not so obvious in his more surface-oriented HPSG-approach. In sum, then, realizational approaches to *zu*-placement may seem very elegant at first sight. However, once the entire empirical domain of *zu*-placement is taken into account, it becomes obvious that they need to be restricted in rather specific ways to derive the facts.

Before concluding, I would like to briefly discuss missing *z* in the rule-based framework adopted here. The starting point for the derivation of an example like (56) will again be a left-branching structure with two FPs as in the 3rd construction in Standard German (37-b):

(57)   Missing *z* before inversion, cf. ex. (56)

```
                        CP
                  _____|_____
                 C            TP
                 |        _____|_____
                wil     DP          T′
                      __|__      ____|____
                    dr Hans    T          vP
                                      _____|_____
                                    DP          v′
                                  __|__      ___|___
                                sine Fründe v      VP₁
                                              _____|_____
                                            FP₁          V₁
                                        _____|_____       |
                                      VP₂         F₁    schiint
                                   ____|____       |
                                 FP₂      V₂       z
                              ____|____    |
                            VP₃      F₂  probiere
                         ____|____    |
                       DP       V₃    z
                        |        |
                  t_sine Fründe häuffe
```

Then suppose that the ascending structure is not derived by means of extraposition but by means of PF-inversion of V_1 with FP_1 and V_2 with FP_2:

(58)   Missing *z* after VP-inversion, cf. ex. (56)
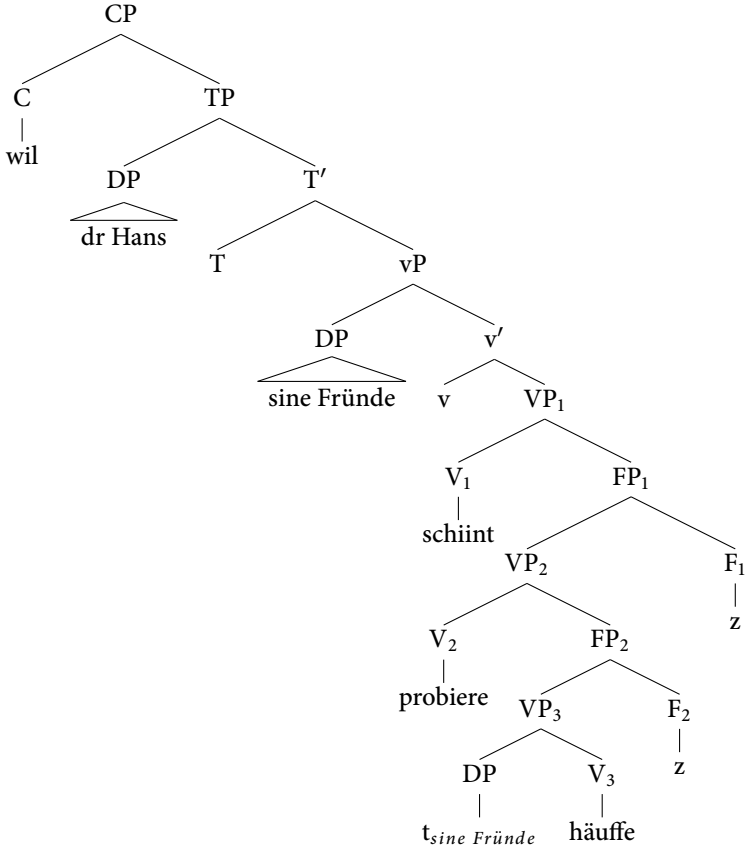
Note that the two *zu*s are now adjacent to each other. Suppose that a rule of haplology can reduce them into one *zu*. Then, *zu*-cliticization can apply in regular fashion, deriving the missing *z* example. The difference between the 3rd construction in Standard German and the missing *z*-construction would therefore reside in the operation that derives the ascending structure: extraposition vs. VP-inversion. Why there should be this difference remains to be investigated; perhaps, the possibility of VP-inversion with *zu*-infinitives in Swiss German can be related to the pervasiveness of ascending structures in these varieties.[25]

---

[25]Note that an analysis of the 3rd construction in Swiss German in terms of VP-inversion makes clear predictions with respect to scope reconstruction. Since no remnant movement is involved, we would expect scope reconstruction to be possible, as in VPR-structures like (48-a).

Before concluding, I would like to add a few empirical details concerning the missing *z* construction. The phenomenon appears to be very subtle. I have not been able to elicit it in an informal survey with linguists speaking a Swiss German dialect by means of a translation task, and in judgment tasks, examples with missing *z* were often rejected. This may be related to the above-mentioned preference for finite subordinate clauses in Swiss German. Many speakers also allow for the Standard German variant with two *zu*s, suggesting that extraposition may also be an option in their grammar. According to Cooper (1995: 188f.), who discusses the phenomenon in some detail, missing *zu* is limited to Verb Raising cases and is blocked whenever there is non-verbal material between V_2 'try' and V_3. Since her example on p. 189 strikes me as very unnatural, I've constructed a pair based on (56) with the judgments taken from Cooper:

(59) a. VPR/3rd: 2x *z Swiss German*
          De Hans schiint$_1$ *(z) probiere$_2$ [siine Fründe] z hälffe$_3$.
          the John seems   to try.INF   his.DAT friends to help.INF
          'John seems to try to help his friends.'
      b. VR: 1 *z Swiss German*
          De Hans schiint$_1$ [siine Fründe] (z) probiere$_2$ z hälffe$_3$.
          the John seems his.DAT friends to try.INF   to help.INF
          'John seems to try to help his friends.'

I have not been able to replicate this contrast in my survey.[26] Data from the internet are only of limited help as there are only very few hits; in fact, I have been able to find only two relevant examples:

(60) a. … au ohni   probiere z wohrsagere
              also without try.INF   to prophesy.INF
              'without trying to prophesy'[27]

The facts are subtle, and I will limit myself to pointing out that Cooper (1995: 197, 199, fn. 39) argues that scope reconstruction *is* possible in the 3rd construction in Zurich German.

[26] The same goes for Cooper's examples 87 on p. 193 and 91 on p. 194f.; her data generally seem somewhat dubious to me.

[27] http://thats-me.ch/forum/em-gewinner/20/31, found on March 28, 2013.

  b.   ... ohni   öpe jeh mau säuber probiere, Dütsch z rede
        without PRT ever once self   try.INF   German to speak.INF
        'without ever trying to speak German oneself'[28]

Example (60-b) is a counterexample to Cooper's claim. If missing *z* were indeed restricted to VR-structures, they would constitute a problem for the present account since inversion of V_2 with FP_2 would have to depend on whether VP_2 contains non-verbal material; obviously, there is no simple way of ensuring this. Bader (1995) and Vogel (2009) also do not predict a VPR/VR-asymmetry. Future research will have to determine the precise properties of the construction, but for now I will simply conclude that it is another phenomenon that can be covered by the rule-based approach. The empirical coverage between the derivational/rule-based proposed here and the realizational approach is similar, but in my view the derivational account is superior in providing a more interesting account of the cross-linguistic variation and in requiring fewer stipulations to rule out unlimited downward percolation of the *zu*-feature.

3.7. PF-Rules in a Right-Branching Sructure

Until now, I have presupposed a left-branching structure as the input for the PF-operations. In this subsection, I will briefly evaluate the prospects of an analysis based on a right-branching structure. I will ignore the placement of non-verbal elements such as objects; for what follows, it is immaterial whether they obtain their preverbal position by means of movement as in strongly anti-symmetric approaches (e.g. Zwart 1994) or whether they are directly linearized as left-hand sisters of V as in approaches that employ linearization parameters that are sensitive to syntactic category (cf. e.g. Cooper 1995, Barbiers 2000, Schmid and Vogel 2004).

  I will start with Dutch where things remain straightforward. For the simple ascending cases like (9-b) the starting point will be stacked right-branching VPs. As a language-particular property, all non-verbal constituents have to move out of the VP (for discussion of the evacuation operation, cf. Salzmann 2011).

---

[28] http://www.chefkoch.de/forum/2,22,296109/An-alle-CHer-Wir-zelebrieren-den-Kantoenligeist.html, found on March 28, 2013.

(61) *Standard Dutch*
      zonder [$_{vP}$ het boek [$_{FP}$ te [$_{VP_1}$ hebben$_1$ [$_{VP_2}$ t$_{het\ boek}$ gelezen$_2$]]]]
      without   the book   to   have.INF   read.PRT
      'without having read the book'

Te will then lower onto V$_1$. Dutch also allows descending orders to a limited degree, especially with participles. In those cases, *te* also appears on V$_1$:

(62) *Standard Dutch*
      zonder het boek gelezen$_2$ **te** hebben$_1$
      without the book read.PRT to have.INF
      'without having read the book'

This implies that *te*-lowering has to precede reordering – as under the left-branching analysis. Since descending orders in Dutch are as impenetrable as their German counterparts, I will assume that they are also the result of cluster-formation based on adjacency. Note that given a right-branching base, descending orders no longer involve string vacuous cluster formation but cluster formation with reordering. String-vacuous cluster formation is still needed for 312 clusters, cf. fn. 29.

   Turning now to German and starting with simple descending structures like (9-a), repeated for convenience:

(63) 321 *Standard German*
      Er dachte, das Buch [lesen$_3$ können$_2$ **zu** müssen$_1$].
      he thought the book read.INF can.INF   to must.INF
      'He thought he had to be able to read the book.'

The starting point will be a right-branching VP-structure with *zu* at the beginning of the verb cluster (as in Dutch, the object has scrambled out of the lexical VP):

(64)   ... [$_{vP}$ das Buch [$_{FP}$ zu [$_{VP_1}$ müssen$_1$ [$_{VP_2}$ können$_2$ [$_{VP_3}$ t$_{das\ Buch}$
          the book   to   must.INF   can.INF
      lesen$_3$]]]]]
      read.INF

The major challenge for a right-branching account is the pre-final position of *zu*. I postulate an inversion rule that inverts *zu* (= F) with VP_1 so that it occurs after the last element of the verb cluster:

(65)   ... [$_{vP}$ das Buch [$_{FP}$ [$_{VP_1}$ müssen$_1$ [$_{VP_2}$ können$_2$ [$_{VP_3}$ t$_{das\ Buch}$
           the book   must.INF   can.INF
        lesen$_3$]]] zu]]
        read.INF to

The descending order in the verb cluster is derived by means of reordering cluster formation. Reordering has to precede *zu*-cliticization since *zu* ends up on V_1 in this case. The result is illustrated in the following example:

(66)   [$_{vP}$ das Buch [$_{FP}$ [$_V$ lesen$_3$+können$_2$+**zu**+müssen$_1$] t$_{zu}$]]
           the book   read.INF+can.INF+to+must.INF

The cases of misplaced *zu* work similarly. I will illustrate the workings on the basis of the ascending VPR-example (14), repeated for convenience:

(67)   1X2 *Swiss German*
        ohni   mi [welle$_1$ [uf d bullesite **z** stelle$_2$]], im   gegeteil,
        without me want.INF on the cops.side to put.INF   on.the contrary
        aber ...
        but
        'without wanting to side with the cops, on the contrary, but ...'

The starting point is again a stacked VP-structure:

(68)   12, ex. (14) *Swiss German*
        ohni   mi [$_{FP}$ z [$_{VP_1}$ wele$_1$ [$_{VP_2}$ t$_{mi}$ uf d   Bullesiite stelle$_2$]]]
        without me   to   want.INF   on the cops.side put.INF

Then, F is inverted with VP_1, placing *z* at the end of the VP:

(69)   12, ex. (14) *Swiss German*
        ohni   mi [$_{FP}$ [$_{VP_1}$ wele$_1$ [$_{VP_2}$ t$_{mi}$ uf d   Bullesiite stelle$_2$]] z]
        without me   want.INF   on the cops.side put.INF to

Then, *zu*-cliticization applies and "mis"places *zu* on V_2.[29]

The left-branching cases with two *zu*s as in (10-a) are next. I repeat the relevant example for convenience:

(70)   321 *Standard German*
    dass er das Buch **zu** lesen$_3$ **zu** versuchen$_2$ versprach$_1$
    that he the book to read.INF to try.INF   promised
    'that he promised to try to read the book'

The starting point would look as follows (with the object having scrambled to Spec$_V$):

(71)   dass er [$_{vP}$ das Buch [$_{VP_1}$ versprach$_1$ [$_{FP_1}$ zu [$_{VP_2}$ versuchen$_2$ [$_{FP_2}$ zu
    that he   the book   promised   to   try.INF   to
    [$_{VP_3}$ t$_{das\ Buch}$ lesen$_3$]]]]]]
                read.INF

Then, F_1 inverts with VP_2 and F_2 inverts with VP_3 to put *zu* at the end of the respective non-finite phrase:

(72)   dass er [$_{vP}$ das Buch [$_{VP_1}$ versprach$_1$ [$_{FP_1}$ [$_{VP_2}$ versuchen$_2$ [$_{FP_2}$ [$_{VP_3}$
    that he   the book   promised   try.INF
    t$_{das\ Buch}$ lesen$_3$]   zu]] zu]]]
          read.INF to   to

We argued above that (reordering) cluster formation takes place before *zu*-cliticization. However, if we first form the entire cluster consiting of $V_3+V_2+V_1$, we would end up with both *zu*s next to each other, which may trigger haplology. Then, the remaining *zu* would arguably be affixed onto V_1, which is, of course, the wrong result:

(73)   321 *Standard German*
    *dass er das Buch lesen$_3$ versuchen$_2$ **zu** versprach$_1$
    that he the book read.INF try.INF   to promised
    'that he promised to try to read the book'

---

[29] The analysis of the 'scandal construction' (12-a) is similar, but slightly more complex: To derive a 312 order from a linear structure, we have to allow for string-vacuous cluster formation between V_2 and V_3 followed by reordering cluster formation between V_1 and [V_2+V_3]. Both processes need to precede *zu*-cliticization, which is unproblematic under a cyclic PF-derivation.

It rather seems that we have to intersperse cluster formation with *zu*-cliticization so that *zu* ends up on V_2 and V_3. This is exactly the result that obtains under a cyclic bottom-up derivation: If we assume that the bracketed structure in (72) is still available and derive the structure bottom-up, the first step will be to cliticize F_2 onto V_3. Then, the complex zu_2+V_3 would undergo reordering cluster formation with V_2, leading to $zu_2+V_3+V_2$. Then, *zu*-cliticization of F_1 onto V_2 would apply, leading to $zu_2+V_3+zu_1+V_2$. Finally, the entire complex would undergo reordering cluster formation with V_1, producing the correct output $zu_2+V_3+zu_1+V_2+V_1$. Importantly, a cyclic derivation also produces the correct result for simple descending cases like (9-a) where the entire verbal complex is formed before *zu*-cliticization can apply.

The missing *z* facts in (56) can be handled quite straightforwardly: I repeat the relevant example from above:

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

The starting point will be the following structure:

(75)   wü   dr Hans [$_{vP}$ sine   Fründe [$_{vP_1}$ schiint$_1$ [$_{FP_1}$ z [$_{VP_2}$
    because the John   his.DAT friends   seems   to
    probiere$_2$ [$_{FP_2}$ z [$_{VP_3}$ t$_{sine Fründe}$ häuffe$_3$]]]]]]
    try.INF   to   help.INF

Then, F_1 inverts with VP_2 and F_2 with VP_3:

(76)   wü   dr hans [$_{vP}$ sine   Fründe [$_{vP_1}$ schiint$_1$ [$_{FP_1}$ [$_{VP_2}$ probiere$_2$
    because the John   his.DAT friends   seems   try.INF
    [$_{FP_2}$ [$_{VP_3}$ t$_{sine Fründe}$ häuffe$_3$] z]] z]]]
                      help.INF to to

Then, a haplological rule reduces the two *zu*s to one and *zu*-cliticization applies, deriving the desired result.

What remains to be discussed is the 3rd construction in Standard German with 2 *zu*s as in (36-b). I repeat the relevant example:

(77)   VP-complement/3rd construction, *Standard German*
ohne   mich [$_{VP}$ zu versuchen$_1$ [$_{VP}$ zu mögen$_2$]]
without me   to try.INF   to like.INF
'without trying to like me'

The basis will be the following structure:

(78)   ohne   [$_{vP}$ mich [$_{FP_1}$ zu [$_{VP_1}$ versuchen$_1$ [$_{FP_2}$ zu [$_{VP_2}$ t$_{mich}$
without   me   to   try.INF   to
mögen$_2$]]]]]
like.INF

It must not be derived like the missing *z* case because both *zu*s have to be
retained. The only possibility to derive the correct result is extraposition of FP$_2$
to a position above FP$_1$, e.g. vP. Then, F$_1$ can invert with VP$_1$ and F$_2$ with VP$_2$,
leading to the following tree structure:

(79)   The 3rd construction in Standard German under a right-branching
VP-structure
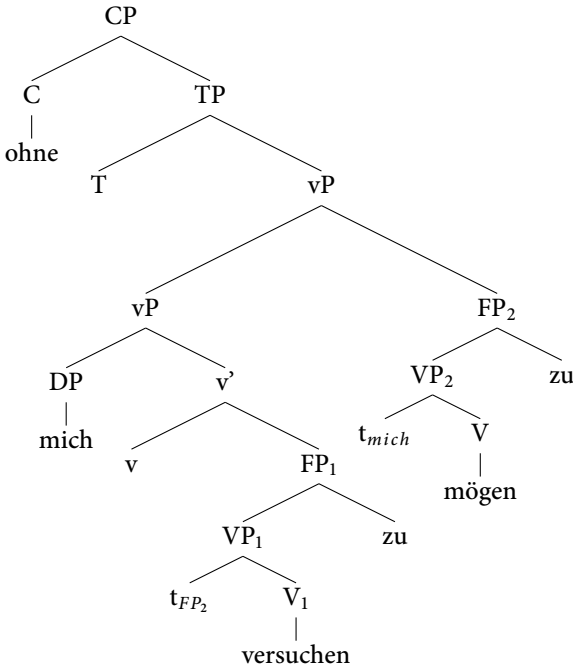
Then, *zu*-cliticization can apply in both VPs. Note that for the same reason extraposition must also be assumed for CP-complements because otherwise the *zu* assigned to matrix VP$_1$ would end up on the last verb of the complement CP, i.e. the problem is exactly the same as under the left-branching structure in (42).

As has been shown in this subsection, it is also possible to determine a non-contradictory ranking on the basis of a right-branching structure. The complete ordering is as follows:

(80)   Ordering of operations (right-branching): synopsis

extraposition (RC, CP/FP$_2$) > topicalization (VP) >	syntax
te-lowering/*zu*-inversion > cluster formation (+/-inversion) > < *zu*-cliticization	PF

The result is quite similar to the one in table (51): The major difference is that VP-inversion is no longer necessary while a rule of *zu*-inversion had to be added. The relative ordering of *te*-lowering and *zu*-inversion cannot be determined for obvious reasons. As under a left-branching structure, the ordering in syntax follows from cyclicity, and the same goes for the relative ordering between cluster formation and *zu*-cliticization. In other words, the entire ordering in (80) is intrinsic: either because of cyclicity in syntax (extraposition vs. topicalization) and PF (reordering cluster formation and *zu*-cliticization) or because the operations take place in separate subcomponents of PF (*te*-lowering has to take place before the adjacency-sensitive operations). In the ordering in (51), however, the ordering between *te*-lowering and VP-inversion requires extrinsic ordering.

This may constitute slight advantage for a right-branching approach. But there are two aspects that seem suboptimal: First, extraposition for the 3rd construction and CP-complements is still necessary (at least in German) even though they can be directly linearized as right-hand sisters of V. This undermines one – independent – argument in favor of a right-branching structure (Zwart 1994); but since a left-branching structure requires obligatory extraposition as well (and is faced with the same questions w.r.t. a plausible trigger), this is probably not too detrimental. The only rule that seems quite stipulative and which can be avoided under a left-branching approach is *zu*-inversion. While this leads to a complication for German, it should be pointed

out that generating the infinitival particle in a functional projection above (and thus before) the verb cluster has advantages for West-Flemish where *te* is not associated with the verb, recall (28-a). Under a left-branching approach, one probably has to postulate a corresponding rule of *te*-inversion to move it to the beginning of the verb cluster in West Flemish while no such rule is necessary under a right-branching account.

The following table lists the relevant phenomena with the rules required under both a left-branching and a right-branching approach ('svCF' stands for string-vacuous cluster formation, 'rCF' for reordering cluster formation, and 'hapl' for haplology; the other abbreviations should be self-explanatory):

(81)　Phenomena and derivations

		left-branching	right-branching
32*zu*1	(9-a)	svCF > *zu*-clit	*zu*-inv > rCF > *zu*-clit
*zu*3*zu*21	(10-a)	*zu*-clit > svCF > *zu*-clit > svCF	*zu*-inv (2x) > *zu*-clit > rCF > *zu*-clit > rCF
13*zu*2	(11-c)	VP-inv > *zu*-clit > svCF	*zu*-inv > rCF > *zu*-clit
31*zu*2	(12-a)	rCF > svCF > *zu*-clit	*zu*-inv > svCF > rCF > *zu*-clit
*zu*1*zu*2	(36-b)	extrapos > *zu*-clit (2x)	extrapos > *zu*-inv (2x) > *zu*-clit (2x)
*zu*1CP	(42)	extrapos > *zu*-clit	extrapos > *zu*-inv > *zu*-clit
12*zu*3	(56)	VP-inv (2x) > hapl > *zu*-clit	*zu*-inv (2x) > hapl > *zu*-clit
*te*123	(9-b)	*te*-lowering> VP-inv (2x)	*te*-lowering
2*te*1	(62)	*te*-lowering > svCF	*te*-lowering > rCF

This table suggests that fewer operations are needed under a left-branching account. This tends to be correct for German and its varieties where descending orders are frequent. Once we look at Dutch, where ascending orders predominate, a right-branching VP-structure provides the better input for the PF-rules. Consequently, once the larger picture is taken into account, the price to be paid seems to be similar in both approaches. While a left-branching VP-structure provides the simpler solution for descending structures, a right-branching VP-structure is superior for ascending structures.

4.　Conclusion

In this paper, I have addressed two puzzles in the domain of verb cluster formation in terms of rule ordering. I first discussed the extraposition paradox where extraposition to a VP that is part of a verb cluster is blocked when the VP is in-situ but not when it is topicalized. I have argued that the verbs have to form a complex head when adjacent in descending order. In contrast to earlier

approaches, cluster formation takes place post-syntactically and is subject to an adjacency requirement. This explains why extraposition to the non-final VP in V-final structures leads to ungrammaticality: cluster formation is blocked as the adjacency is disrupted; extraposition thus bleeds cluster formation. Since cluster formation takes place after topicalization, nothing prohibits extraposing to the lexical VP if it is later moved to the beginning of the clause. In that case, topicalization bleeds cluster formation as the context for the latter operation is destroyed (there is no sequence of verbs anymore).

The second puzzling phenomenon discussed was the placement of the infinitival particle in Dutch and German. I have assumed that the particle is an independent syntactic element in both languages. The cross-linguistic differences result from the fact that the operation that associates the particle with the verb takes place at different points of the derivation in the two languages. While it is an early PF-process in Dutch that is still sensitive to hierarchical structure, it is a late process in German because it is sensitive to linear order and adjacency. Thus, while *te*-placement represents an instance of Lowering, *zu*-cliticization is best described as an instance of Local Dislocation in the framework of Embick and Noyer (2001).

I have shown that these two processes interact with other PF-rules such as cluster formation and VP-inversion. It turned out to be possible to determine a strict and non-contradictory ordering between these rules. Furthermore, with one exception under a left-branching VP-structure, the ordering is fully intrinsic – either because it follows from cyclicity (in syntax and PF) or because the operations take place in different components (syntax vs. PF or different subparts of PF). I take this to lend support to the rule-based approach pursued here. Furthermore, it can be seen as an initial attempt to provide insight into the articulation of the PF-component with earlier rules being more sensitive to hierarchical structure while later rules operate on linear structure.

On a more general theoretical level, a comparison between a left-branching and a right-branching approach has not revealed any significant advantages for either of the approaches. Rather, while a left-branching approach is best suited to derive descending orders, a right-branching approach provides a simpler account of ascending structures. Finally, a consistent account of *zu*-placement requires a remnant extraposition analysis of the 3rd construction (at least in Standard German) and more generally obligatory extraposition of right-hand VP and CP-complements of V.

References

Arregi, Karlos and Andrew Nevins (2012): *Morphotactics. Basque Auxiliaries and the Structure of Spellout*. Springer, Dordrecht.

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:* Z. Penner, ed., *Topics in Swiss German syntax*. Peter Lang, Bern, pp. 19–28.

Barbiers, Sjef (2000): The Right Periphery in SOV Languages: English and Dutch. *In:* P. Svenonius, ed., *The Derivation of VO and OV*. John Benjamins, Amsterdam, pp. 181–218.

Barss, Andrew (1986): Chains and Anaphoric Dependencies. Doctoral dissertation, MIT.

Bayer, Josef, Tanja Schmid and Markus Bader (2005): Clause Union and Clausal Position. *In:* M. d. Dikken and C. M. Tortora, eds, *The function of function words and functional categories*. John Benjamins, Amsterdam, pp. 79–113.

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

Bobaljik, Jonathan and Susi Wurmbrand (2005): 'The Domain of Agreement', *Natural Language & Linguistic Theory* **23**(4), 809–865.

Broekhuis, Hans (1993): 'Verb Projection Raising', *Spektator* **22**, 28–47.

Chomsky, Noam (1973): Conditions on Transformations. *In:* S. R. Anderson and P. Kiparsky, eds, *A Festschrift for Morris Halle*. Holt, Rinehart and Winston, New York, pp. 232–286.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Massachusetts.

Cooper, Kathrin E. (1995): *Topics in Zurich German Syntax*. Germanistisch Instituut, Groningen.

Dikken, Marcel den (1996): The Minimal Links of Verb (Projection) Raising. *In:* W. Abraham, S. D. Epstein, H. Thrainsson and C. J.-W. Zwart, eds, *Minimal Ideas. Syntactic Studies in the Minimalist Framework*. Benjamins, Amsterdam, pp. 67–96.

Dikken, Marcel den and Eric Hoekstra (1997): 'Parasitic Participles', *Linguistics* **35**(6), 1057–1090.

Donaldson, Bruce C. (1993): *A Grammar of Afrikaans*. Mouton de Gruyter, Berlin/New York.

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.

Geilfuss-Wolfgang, Jochen (1991): *Verb- und Verbphrasensyntax*. Arbeitspapier des SFB 340 11, University of Tübingen, Tübingen.

Haegeman, Liliane (1992): *Theory and Description in Generative Syntax. A Case Study in West Flemish*. Cambridge University Press, Cambridge.

Haegeman, Liliane (1998): 'Verb Movement in Embedded Clauses in West Flemish', *Linguistic Inquiry* **29**(4), 631–656.

Haegeman, Liliane and Henk van Riemsdijk (1986): 'Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs', *Linguistic Inquiry* **17**(3), 417–466.

Haider, Hubert (1993): *Deutsche Syntax – Generativ: Vorstudien zur Theorie einer projektiven Grammatik*. Narr, Tübingen.

Haider, Hubert (2003): V-Clustering and Clause Union: Causes and Effects. *In:* P. Seuren and G. Kempen, eds, *Verb Constructions in German and Dutch*. John Benjamins, Amsterdam, pp. 91–126.

Haider, Hubert (2010): *The Syntax of German*. Cambridge University Press, Cambridge.

Haider, Hubert (2011): 'Grammatische Illusionen – Lokal wohlgeformt – global deviant', *Zeitschrift für Sprachwissenschaft* **30**(2), 223–257.

Halle, Morris and Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. *In:* K. Hale and S. J. Keyser, eds, *The view from building 20. Essays in linguistics in honour of Sylvain Bromberger*. MIT Press, Cambridge, pp. 111–176.

Hinterhölzl, Roland (2009): 'The IPP-Effect, Phrasal Affixes and Repair Strategies in the Syntax-Morphology Interface', *Linguistische Berichte* **2009**(218), 191–215.

Hodler, Werner (1969): *Berndeutsche Syntax*. Francke Verlag, Bern.

Höhle, Tilman N. (2006): Observing Non-Finite Verbs: Some 3v Phenomena in German-Dutch. *In:* P. Brandt and E. Fuß, eds, *Form, Structure, and Grammar. A Festschrift Presented to Günther Grewendorf on Occasion of His 60th Birthday*. Akademie Verlag, Berlin, pp. 55–77.

Kayne, Richard S. (1994): *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.

Neeleman, Ad and Fred Weerman (1993): 'The Balance between Syntax and Morphology: Dutch Particles and Resultatives', *Natural Language & Linguistic Theory* **11**(3), 433–475.

Richards, Marc (2008): 'Desymmetrization: Parametric Variation at the PF-interface', *Cancadian Journal of Linguistics* **53**(2/3), 275–300.

Riemsdijk, Henk van (1998): 'Head Movement and Adjacency', *Natural Language & Linguistic Theory* **16**(3), 633–679.

Salzmann, Martin (2011): Resolving the Movement Paradox in Verb Projection Raising. *In:* O. Bonami and P. Cabredo Hofherr, eds, *Empirical Issues in Syntax and Semantics 8*. CNRS, Paris, pp. 453–485.

Salzmann, Martin (to appear): 'New Evidence for Postsyntactic Verb Cluster Formation and a Right-Branching Base Order', *Proceedings of NELS* **42**.

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.

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.

Stechow, Arnim von (1990): Status Government and Coherence in German. *In:* G. Grewendorf and W. Sternefeld, eds, *Scrambling and barriers*. John Benjamins, Amsterdam, pp. 143–198.

Sternefeld, Wolfgang (1990): Scrambling and Minimality. *In:* G. Grewendorf and W. Sternefeld, eds, *Scrambling and barriers*. John Benjamins, Amsterdam, pp. 239–257.

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

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*. Schweizer Spiegel Verlag, Zürich.

Williams, Edwin (2004): The structure of clusters. *In:* K. É. Kiss and H. v. Riemsdijk, eds, *Verb Clusters: A Study of Hungarian, German, and Dutch*. John Benjamins, Amsterdam, pp. 173–201.

Wurmbrand, Susanne (2004*a*): 'Two Types of Restructuring–Lexical vs. Functional', *Lingua* 114(8), 991–1014.

Wurmbrand, Susanne (2004*b*): West Germanic Verb Clusters: The Empirical Domain. *In:* K. . Kiss and H. v. Riemsdijk, eds, *Verb clusters: A study of Hungarian, German, and Dutch*. John Benjamins, Amsterdam, pp. 43–85.

Wurmbrand, Susanne (2005): Verb Clusters, Verb Raising, and Restructuring. *In:* M. Everaert and H. v. Riemsdijk, eds, *The Blackwell companion to syntax. Volume 5*. Blackwell, Oxford, pp. 227–341.

Wurmbrand, Susanne (2007): 'How Complex Are Complex Predicates?', *Syntax* 10(3), 243–288.

Zwart, Jan-Wouter (1994): 'Dutch is Head Initial', *The Linguistic Review* 11(3/4), 377–406.