1. Background

A requirement for any minimalist approach to structure-building (as in Chomsky (2001; 2008; 2013)) is that it can be decided whether a given Merge(α,β) operation that combines two categories α and β (each of which may be a lexical item or internally complex) is legitimate. There are basically two options, viz., approaches in terms of feature-driven Merge and approaches relying on free Merge. On the one hand, with feature-driven Merge of two items α, β, it can be assumed that one of the two items (say, α) is equipped with an intrinsic formal property requiring (or permitting) the other item (β) to be its sister. On this view, designated features for structure-building on α must be matched by β, and can be assumed to be discharged as a consequence of carrying out the operation. On the other hand, in a free Merge approach, Merge applies without restrictions throughout, which initially leads to massive overgeneration. Subsequently, filters check an output representation generated by free Merge and decide about the legitimacy of the operation. These filters (i.e., representational constraints) can in principle be of various types: syntactic, semantic, prosodic, information-structural, even stochastic – thus, they do not need to be syntax-internal.

The two approaches are often extensionally equivalent. However, there can in principle be contexts where they make different predictions. (1) illustrates a critical configuration. Here, Merge first combines two items α and β, with α acting as the head of the new projection (cf. (1-a)), and in a following step, β is removed from the structure again, as a consequence of an operation X applying to some item γ and β (cf. (1-b)).

(1) a. Merge(α,β) → [α α β]
   b. X(γ,β) → ... [α α]

Under feature-driven Merge, the legitimacy of Merge(α,β) in (1-a) can be correctly determined: The operation is well formed if a structure-building feature of α is matched by β in the derivation, and later operations which undo the configuration are unproblematic. Thus, counter-bleeding takes place (see Chomsky (1951; 1975), Kiparsky (1973)): Removing β in the second step comes too late to bleed the original Merge operation. In contrast, in the free Merge approach, where only the final output representations are checked, problems can arise: In particular, bleeding of Merge(α,β) may now wrongly be
predicted because the justification for this operation cannot be read off the output structure in (1-b) (i.e., the output representation is \textit{opaque}, in Kiparsky’s terminology).

Against this background, the goal of the present paper is to pursue the question of whether an operation \( X \) with the properties sketched in (1-b) can plausibly be assumed in syntactic theory. I will argue that this is indeed the case; consequently, there is an argument for feature-driven Merge as opposed to free Merge.

What could the operation \( X \) consist of? A first candidate might be movement, i.e., internal Merge. Internal Merge by definition always presupposes some earlier external Merge operation, and internal Merge may also follow another internal Merge operation that has applied to the same item (successive-cyclic movement). Thus, there is a potential problem for the free Merge approach arising as a consequence of output opacity because internal Merge might undo the configuration generated by earlier structure-building. However, in most current theories of movement this potential problem does not become an actual problem because it is assumed that if \( \beta \) is moved from one position to another one, it is not actually removed from the first position; rather, movement is generally taken to leave behind a copy (or a trace), or to merely create a second occurrence of the same item: Thus, the original configuration required by the output filters is preserved.\(^1\)

However, I would like to suggest that there is another candidate for \( X \), viz., removal of structure. For concreteness, suppose that Merge\((\alpha, \beta)\) is followed by another operation that does not \textit{build} structure by merging \( \beta \) anew but rather \textit{removes} structure by eliminating \( \beta \) from the derivation; the empirical evidence for such an operation will be such that whereas some syntactic operations (taking place before removal) require the presence of \( \beta \), subsequent syntactic operations require the complete absence of \( \beta \).

Intermediate scrambling is a case in point. In a free Merge approach to scrambling in German, this operation is not assumed to be feature-driven but to be licensed by information-structural and prosodic constraints referring to the position where the moved item is overtly realized. Consequently, cases of \textit{intermediate} scrambling are a priori unexpected under this view. However, intermediate scrambling has been argued to underlie the absence of superiority effects with clause-bound wh-movement in German (as in (i-a); see Fanselow (1996) and Grohmann (1997)), and the occurrence of superiority effects with non-clause bound wh-movement in German (as in (i-b); see Büring & Hartmann (1994), Fanselow (1996), Heck & Müller (2000), Pesetsky (2000), pace Fanselow & Féry (2008), Fanselow (2015)).

(i) a. Was2 hat t2 wer1 t2 gesagt ?
   what\textsubscript{acc} has who\textsubscript{nom} said
   “Who said what?”

   b. *Was2 hat wer1 gesagt [\textsc{cp} dass der Fritz t2 mag ] ?
      what\textsubscript{acc} has who\textsubscript{nom} said that the Fritz likes
      “Who said that Fritz likes what?”

\(^1\) There is a caveat, however. Suppose that the output filter determining the legitimacy of an internal Merge operation applying to some item \( \beta \) is such that it requires the phonological realization of \( \beta \) in some designated position that corresponds exactly to the position reached after Merge\((\alpha, \beta)\). Then, subsequent internal Merge\((\gamma, \beta)\) moving \( \beta \) to another position (with concomitant phonological realization in this latter position) will invariably create output opacity, in the sense that the trigger for the first movement step cannot be checked anymore.
(so that assuming $\beta$ to be merely PF-invisible would not suffice). If such an operation exists, the legitimacy of the original Merge operation cannot be checked by output filters – by definition, there can be no structural reflex (copy, occurrence, etc.) of the structure removal operation --; and consequently, there is an argument for feature-driven (as opposed to free) Merge.\(^2\) To establish this argument, I will proceed as follows. First, I will outline a principled theory of structure-removal in section 2 that centers around an elementary operation Remove. After that, in sections 3 and 4, I present evidence from a number of different empirical domains of German syntax (passive, applicative, restructuring, and complex prefield constructions) that suggests the existence of an operation like Remove for phrases and heads, respectively. Finally, section 5 draws a conclusion.

2. Remove

I would like to contend that syntactic derivations employ two elementary operations modifying representations: In addition to an operation that builds structure – Merge (Chomsky (2001; 2008; 2013)) --, there is a complementary operation that removes structure: Remove.

Empirical support for such an operation comes from incompatible structure assignments in syntax. As a matter of fact, there is substantial evidence for conflicting representations in syntactic derivations. The standard means to account for this phenomenon is movement (internal Merge): If some item $\alpha$ shows properties associated both with position P and position Q, then this is due to the fact that $\alpha$ has moved from Q to P. Addressing conflicting representations in terms of movement is often straightforward (cf., for instance, $\theta$-assignment in the base position, accompanied by satisfaction of a criterial movement constraint in the derived position, as with wh-movement of an object), sometimes less obviously so (see, e.g., Weisser (2015) on medial clauses and asymmetric coordination, derived by correlating base-generated subordination (Q) and surface coordination (P) by movement of the clause to a Spec& position). However, there are many cases of conflicting representations that do not lend themselves to analyses in terms of movement; and it is these latter cases that can be taken to empirically motivate the existence of structure removal.

If Remove exists as the mirror image of Merge, it is expected to show similar properties and obey identical constraints. I will adopt the following four assumptions about Merge. First, Merge is feature-

\(^2\) There is one further qualification. If the evaluation of representations by output filters in a free Merge approach can take place iteratively, based on smaller structures, then the legitimacy of Merge operations applying to material that is eventually removed might in principle be correctly determined in such an approach after all. For the time being, I will simply presuppose that this is not an option; I will return to this question at the very end of this paper, in section 5.
driven. It is triggered by designated features (here rendered as \([\bullet F\bullet]\)), which are ordered on lexical items (signalled by \(\succ\) in what follows), thereby determining the sequence of operations triggered by a given head (see, among others, Svenonius (1994), Collins (2002), Adger (2003), Lechner (2004), Kobele (2006), Sternefeld (2006), Pesetsky & Torrego (2006), Heck & Müller (2007), Müller (2014), Abels (2012), Stabler (2013) and Georgi (2014)). Second, Merge may apply to heads (incl. head movement in cases of internal Merge) or phrases (incl. XP movement in cases of internal Merge). The difference between the two cases must be formally encoded in any theory; I will assume that this is accomplished by designated indices accompanying the structure-building features: \([\bullet F_0\bullet]\), \([\bullet F_2\bullet]\) (with 0=min, and 2=max). Third, Merge obeys the Strict Cycle Condition in (2) which precludes syntactic operations from solely applying within embedded domains (see Chomsky (1973; 1995; 2001; 2008); also cf. the Extension Condition and the No Tampering Condition). Fourth, Merge can be external or internal.

(2) \textit{Strict Cycle Condition (SCC)}:

Within the current XP \(\alpha\), a syntactic operation may not exclusively target some item \(\delta\) in the domain of another XP \(\beta\) if \(\beta\) is in the domain of \(\alpha\).

(3) \textit{Domain} (Chomsky (1995)):

The domain of a head \(X\) is the set of nodes dominated by XP that are distinct from and do not contain \(X\).

Clearly, if Remove exists, it is expected to obey exactly the same restrictions. I will assume that this is the case: First, Remove is feature-driven. It is triggered by designated \([-F\bullet]\) features, which are ordered on lexical items. Second, Remove may apply to heads or phrases: \([-F_0\bullet\], [-F_2\bullet\].\footnote{With both Merge and Remove, 0 and 2 are mere diacritics that stand for “minimal” and “maximal” projection, respectively, and thus do not actually instantiate a reference to bar levels (\([-F_1\bullet\) is not available as it would truly require reference to a certain bar level.). Furthermore, both operations presuppose a conservative approach to labelling, where the label is directly accessible to the selecting head, and the selecting head invariably determines the label of the current root node after the operation (Merge or Remove) has been carried out.} Third, Remove obeys the Strict Cycle Condition in (2). And fourth, Remove can be external or internal – that said, all the cases I will be concerned with in this article involve internal Remove, i.e., removal of items that are part of the syntactic structure that Remove applies to.\footnote{External Remove amounts to removal of material that is not present in syntactic structure. See Müller (2015a) on how this paradox can be resolved, and on potential empirical evidence for this operation in the areas of adjectival passive and object drop in German. (Basically, the idea is that external Remove provides a new approach to truly implicit arguments – i.e., those arguments which play a role for semantic interpretation and must in some sense exist, but which do not participate in syntactic operations.)}

To illustrate how Remove works in syntactic derivations, let me first consider the case where the operation applies to phrases, beginning with the removal of a complement. In (4), a head \(X\) starts out
with a two-membered list of features for structure manipulation that need to be discharged one after the other. First, in (4-a), X is merged with YP, triggered by a structure-building (subcategorization) feature [•Y•] on X. In the next step in (4-b), YP is removed again from the derivation, triggered by [–Y₂–] on X.

5

(4) Remove and phrases: complements

\[\text{a. Merge}(X[•Y•] \rightarrow [–Y₂–], YP): \quad \text{b. Remove}(X[–Y₂–], YP):\]

```
X'
  \[
X[–Y₂–]
\]
  \[
YP
\]
  \[
ZP
\]
  \[
Y'
\]
  \[
Y
\]
  \[
WP
\]
```

Note that YP is in fact the only phrase in (4-a) that is accessible for removal at this point. If X were to bear a feature [–Z₂–] or a feature [–W₂–], the derivation would crash: ZP, WP cannot be removed by X because of the Strict Cycle Condition (YP is in the domain of the current root projection, ZP and WP are in the domain of YP, and removal would exclusively target a position in a domain embedded in the domain of the root).

6

Specifiers can be removed in the same way, by discharging a designated feature on the head. In (5-a), an X' projection (resulting from prior Merge of X with some UP) is merged with YP which therefore becomes X’s specifier. As shown in (5-b), feature-driven Remove can then subsequently get rid of YP again.

5 Since I am almost exclusively concerned with Merge operations targeting XPs in this paper, I will uniformly use [•Y•] instead of [•Y₂•].

6 Thus, (4) essentially qualifies as a Duke-of-York derivation (see Pullum (1976), McCarthy (2003), and Lechner (2010), among others). As is generally the case with this type of interaction of operations, it is far from vacuous – crucially, as will be shown below, the intermediate representation can have an influence on the applicability of other processes before it is undone again.

7 Note that this would not hold for internal Merge: Movement of, say, ZP to SpecX would be possible because this operation would not exclusively affect an embedded domain; it would also affect SpecX, hence XP.
\[\text{a. } \text{Merge}(X' [\ast Y \ast] > [-Y_2-], YP): \]

\[\text{b. } \text{Remove}(X' [-Y_2-], YP): \]

Again, ZP and WP cannot be removed by X because of the Strict Cycle Condition. However, in principle, X (bearing [\ast U \ast]) might also remove UP in a configuration like (5-a), i.e., after YP has been merged. To avoid this outcome, the Strict Cycle Condition could be strengthened (from phrases to projections). However, I will assume such a derivation to be permitted, even though this issue will not affect anything that follows below.\(^8\)

Next consider the situation where Remove applies to a head rather than a phrase (triggered by [–F\(_0\)–] rather than by [–F\(_2\)–]). (6) illustrates a case where the head of a complement is removed.

\[^8\text{ There are two reasons for this. First, this kind of derivational step is exactly what is needed to reconcile the option of tucking in-movement (see Richards (2001)) with the Strict Cycle Condition; assuming tucking in to be well motivated with internal Merge, and assuming Merge and Remove to obey the same constraints then implies that X can target UP in (5-a). Second, if ellipsis constructions are to be addressed in terms of structure removal (rather than mere PF deletion), as argued by Murphy (2015), it is unavoidable that in sluicing constructions like (i) in German, removal of the TP by a [–T\(_2\)–] feature on C must take place after wh-movement to SpecC has occurred.}\]

(i) Fritz hat irgendwen gesehen, aber ich weiß nicht [CP wen\(_1\) C [TP der Fritz seen has]]

Fritz has someone seen, but I know not whom the Fritz seen has

“Fritz saw someone, but I don’t know who.”
(6) **Remove and heads: complements**

a. Merge\(X, [Y \rightarrow -Y_0 -], YP)\):

\[
\begin{array}{c}
X' \\
X[-Y_0 -] \\
YP \\
Y \\
ZP \\
\end{array}
\]

b. Remove\(X, [-Y_0 -], Y)\):

\[
\begin{array}{c}
X' \\
X \\
ZP \\
\end{array}
\]

Since \([-F_0 -]\) removes the head, it takes away the highest projection (given a bare phrase structure approach, a head’s projection does not exist independently of the head), but only this. More deeply embedded material (like ZP in (6)) is not affected by structure removal in this case. The question then is what happens with the material that was originally included in the removed projection. The obvious assumption would seem to be that it is reassOCIated with the main projection, i.e., with the projection of the head responsible for structure removal, thereby effectively replacing the original item (YP). Basically, this works like tree pruning (see Ross (1967, ch. 3)); and the same assumption is also made by Stepanov (2012) in his approach to head movement (where the projection of a moved head is assumed to disappear, and material in the head’s original projection is reassOCIated). If there are two or more items in YP (e.g., ZP, WP), the null hypothesis clearly is that they reassemble in their original hierarchical and linear order in the XP domain, so that structural changes induced by the operation are minimized.9

For concreteness, let me be a bit more specific about the reassOCIation operation that is required under removal applying to heads. First, it is clear that reassOCIation cannot be an instance of Merge: It only applies to phrases (not to heads), the external/internal distinction does not make sense here, and, perhaps most importantly, reassOCIation is not feature-driven; rather, it is a last resort operation triggered by the need to reintegrate material into the present tree that is floating around as a consequence of Remove. Second, material that is temporarily unassOCIated to the current tree as a result of Remove cannot be assumed to be part of the workspace (because then it would be expected to only optionally re-enter the structure, and always as an adjunct; see below); rather, it is still in the same domain as the main tree from which it has temporarily been split off. Third, given that reassOCIation must respect the pre-Remove order of items, minimal memory is required for carrying out the operation: If \(\alpha, \beta\) are in the minimal domain of XP, X is subject to head removal, and \(\alpha\) c-commands (precedes) \(\beta\), then \(\alpha\) c-commands (precedes) \(\beta\)

---

9 For Remove applying to a complement head Y, as in (6), but with a complement WP of Y and a specifier ZP of Y, this implies that WP becomes the complement of X, and that ZP becomes the specifier of X. This will become relevant below, in the discussion of wh-movement from restructuring infinitives based on (23). Also see Müller (2016b) for direct empirical evidence for such dislocation without movement (which suggests that a reassOCIated ZP occupies an position exhibiting the properties of base positions) based on restructuring infinitives and tough-movement constructions.
after reassocation. Fourth, nothing needs to be stipulated concerning the locus of reassocation: Given that reassocation, like all syntactic operations, obeys the Strict Cycle Condition, reassigned material will have to show up in the projection of the head that brought about the removal, and can never show up in a lower domain.

Finally, the case where Remove applies to the head of a specifier is shown in (7). In the abstract example chosen here, the head to be removed (Y) has a specifier (ZP) and a complement (WP); consequently, these two items become reassociated as two specifiers of the head X that has triggered the operation.

(7)  Remove and heads: specifiers

a.  Merge(X′[•Y•][−Y₀−],YP):

\[ \text{XP} \quad \text{YP} \quad X' \quad ZP \quad Y' \quad X[−Y₀−] \quad UP \]

b.  Remove(X′[−Y₀−],Y):

\[ \text{XP} \quad ZP \quad X' \quad WP \quad X \quad UP \]

To sum up, Remove applying to YP removes the whole YP constituent, including all other material included in it, whereas Remove applying to Y only takes out the YP shell, leaving all other material included in it intact and attaching it to the triggering head’s projection in a maximally structure-preserving way.

Because of the Strict Cycle Condition, material that is subject to Remove is predicted to exhibit what one might call short life cycle effects (with a principled qualification that I will discuss momentarily). Some other operation \( \Gamma \) can be interspersed between Merge(X,YP) and Remove(X,Y) or Remove (X,YP). However, a YP or YP shell removed by \([−F−]\) is only accessible for other processes for a small part of the derivation: As soon as the derivation moves on and combines XP with some other head, YP ceases to be a possible target for removal. Given incremental, bottom-up derivations, this implies that a YP that is subject to removal at some point of the derivation is expected to be accessible from below (downward accessibility) and inaccessible from above (upward inaccessibility); Remove counter-bleeds \( \Gamma \) but bleeds subsequent operations. Empirical evidence for short life cycle effects of this type can thus be taken to support the hypothesis that structure removal exists. That said, there is one systematic exception to short life cycle effects with structure removal: In those cases where Remove applies to a specifier (as in (5) and (7)), it is actually irrelevant whether this specifier is introduced by external Merge (as presupposed so far) or by internal Merge; consequently, movement should be able to extend the life cycle
of material that is subject to removal, by transporting it to a higher domain where it can be targeted by a head with a \([-F-]\) feature. (I will address this issue in subsection 4.2.)

A final question that needs to be addressed is where the material goes that is subject to Remove. Merge takes a (possibly complex) item from the workspace of the derivation (with the original numeration as a subpart containing only non-complex linguistic expressions taken from the lexicon), and combines it with the current tree. Accordingly, Remove puts a (possibly complex) item back into the workspace. I will suggest below that such a removed item can re-enter the original tree as an adjunct in certain cases (by-phrases in passive contexts and demoted theme arguments in applicative contexts); but in general it does not have to do so. Clearly, this approach then presupposes that a workspace is not necessarily reduced to a single tree by the end of the derivation. In order to distinguish between “active” material in the workspace that must be subject to a syntactic operation and “inactive” material in the workspace that arises as a consequence of structure removal and does not have to re-enter the tree, it can be assumed that there are two separate domains of the workspace reserved for the two different types of linguistic expressions. (On this view, external Remove (see footnote 4) would amount to moving material from the active part of the workspace into the inactive part.)

With all theoretical assumptions in place that tell us what an operation Remove that acts as the counterpart of Merge should look like, let me now turn to empirical evidence in support of it. My strategy will be to address a number of different kinds of phenomena from a single language (German) that suggest removal of phrases or heads, with the properties just laid out (downward vs. upward accessibility, short life cycle effects aside from movement), rather than just one phenomenon, even if that means that it will not be possible to develop the analyses in as much detail as would ultimately be required. Section 3 will be concerned with evidence for removal of XP based on German passive and applicative constructions; section 4 will address evidence for removal of X in German restructuring and complexprefield constructions.

3. Removal of YP: Grammatical Function-Changing

A class of phenomena that lend themselves to analyses in terms of structure removal involves grammatical function-changing. In what follows, I will discuss (verbal) passive and applicative constructions in German from this perspective.

3.1. Passive

Abstracting away from by-phrases for the moment, there is no overt realization of the external argument in German passive constructions; as a matter of fact, this is the core property of passive in general. Still, there is some evidence for a syntactically accessible external argument DP (see Chomsky (1957), Baker, Johnson & Roberts (1989), Sternefeld (1995), Collins (2005), and Merchant (2013), among others). Thus, (8-a) shows that the external argument of a passive construction (rendered as DP_{ext} in what
follows) can exert control into a purpose clause; (8-b) shows that DP_{ext} can control a subject-oriented secondary predicate; and (8-c) shows that DP_{ext} can effect binding of a reciprocal pronoun.\footnote{Williams (2015, ch. 12) concludes for English analogues of (8-a) that the syntactic presence of a controller DP_{ext} does not have to be postulated. A core argument is that instances of remote control as in (i) cannot possibly be accounted for by postulating a local DP_{ext} as a controller; but whatever accounts for remote control might the perhaps be extended to local control as in (analogues of) (8-a).}

(8) a. Der Reifen wurde DP_{ext1} aufgepumpt [{\textsc{cp} \text{ um PRO}_1 \text{ die Fahrt fortzusetzen }}] the tire was inflated in order the journey to continue “The tire was inflated in order to continue the journey.”

b. Das Handout wurde DP_{ext1} [\textsc{sc PRO}_1 übermüdet ] verfasst the handout was tired written “The handout was written while tired.”

c. Es wurde DP_{ext1} einander\textsubscript{1} gedankt it was each other thanked “People thanked each other.”

Assuming that control (of PRO) and binding (of a reciprocal) involve Agree operations (Chomsky (2001)), the conclusion can be drawn that DP_{ext} is syntactically active in (8) and can be accessed, such that Agree can take place between DP_{ext} and an item that it c-commands.

On the other hand, there is also evidence against a syntactic accessibility of DP_{ext} in German passive constructions. For instance, DP_{ext} cannot be interpreted as a variable bound by a quantified DP in a higher clause (cf. (9-a)); DP_{ext} cannot itself be controlled by a higher subject (cf. (9-b), see Stechow & Sternefeld (1988)); and, in contrast to other non-overt material (as in topic drop constructions or with extraction from verb-second clauses), DP_{ext} cannot satisfy a criterial movement constraint like the verb-second requirement (cf. (9-c)).

(9) a. *Kein Student\textsubscript{1} gibt zu [{\textsc{cp} dass DP_{ext} schlecht gearbeitet wurde }] no student admits that badly worked was “No student admits that he did not work well.”

b. *Er versucht [{\textsc{cp} DP_{ext} gearbeitet zu werden }] he tries worked to be “He tries to ensure that work is being done.”

c. *Ich denke [{\textsc{cp} DP_{ext} ist gut gearbeitet worden }] I think is well worked been “I think that people worked well.”

\footnote{Williams (2015, ch. 12) concludes for English analogues of (8-a) that the syntactic presence of a controller DP_{ext} does not have to be postulated. A core argument is that instances of remote control as in (i) cannot possibly be accounted for by postulating a local DP_{ext} as a controller; but whatever accounts for remote control might the perhaps be extended to local control as in (analogues of) (8-a).}

(i) Two outfielders were traded away. The goal was to find a better pitcher.

However, the experimental study reported in McCourt et al. (2015) suggests that there might be two distinct mechanisms involved in local vs. remote control in passive contexts after all. Apart from that, it is not a priori clear that there could not be DP-internal controller in the second clause in (i).
Assuming, as before, that the processes involved in (9) (viz., quantifier binding, control, and movement) require syntactic accessibility of DP_{ext} (for Agree or Merge), the conclusion can be drawn that DP_{ext} is in fact not accessible in the contexts in (9) (signalled by the DP_{ext} notation). Taken together, (8) and (9) suggest that DP_{ext} in German passive constructions is accessible from below and inaccessible from above.\footnote{See Müller (2016c) for further evidence in support of downward accessibility and upward inaccessibility of DP_{ext} in German passive constructions (related, i.a., to principle C effects, non-occurrence of minimality effects, and transparency for anaphoric binding); and for arguments against approaches that postulate full accessibility of DP_{ext} (and account for the evidence in (9) in some other way, cf. the references given at the beginning of this subsection), and against approaches that postulate full inaccessibility (or absence) of DP_{ext} (and accordingly need to reanalyze the evidence in (8), cf. Chomsky (1981), Kiparsky (2013), Schäfer (2012), Alexiadou & Doron (2013), and Alexiadou, Anagnostopoulou & Schäfer (2015), among others). Also, see Alexiadou & Müller (2015) for discussion of a principled exception to upward inaccessibility – DP_{ext} permits extremely local binding by an adverb of quantification, as in (i).}

The simplest, most straightforward way to account for this generalization is to assume that accessibility results from the syntactic presence of DP_{ext}, and that inaccessibility is due to the fact that DP_{ext} is removed from the structure; alternative analyses cannot easily derive the systematic pattern underlying the generalization. For concreteness, the analysis developed in Müller (2016c) works as follows. Passive is triggered by the optional addition of a [–D\_2] feature to v in the numeration (i.e., to the very same head that introduces the external argument DP). [–D\_2] on v will remove an existing DP specifier of v. Furthermore, the system is myopic and exerts instantaneous repair: Removal of an argument DP immediately triggers removal of the next case feature from v; this accounts for absorption of structural case. This is essentially a consequence of whatever derives Burzio’s generalization. For concreteness, suppose that a head assumes that the number of DPs and case features is balanced; undoing the effect of a [•D•] feature by discharging a [–D\_2] feature therefore invariably implies removal of a [*case*] feature on a head in the syntax (if such a feature is present).\footnote{This implies that probes can be deleted locally when the need arises; see Béjar & Řezáč (2009), Preminger (2014), and Georgi (2014), among others. As noted by two reviewers, alternatives to this approach to case absorption are readily available under present assumptions. For instance, it could simply be stipulated as a pre-syntactic restriction on v heads that [*acc*]} On this view, the derivation of a typical German passive construction like (10-a) involves the steps in (10-b).

(10)  a. dass das Buch gelesen wurde
     that the book\textsubscript{nom} read was
     “that the book was read.”

**(i)** Es wurde größtenteils DP_{ext} geschlafen beim Vortrag
     it was for the most part slept at the talk
     ‘Most people slept through the talk.’
In (10-b-i), there is a \( v \) with structure-building features for Merge operations with \( \text{VP} \) and \( \text{DP}^{\text{ext}} \), plus a \([-\text{D}_2-]\) feature for \( \text{DP} \) removal (this is why it qualifies as a passive head), plus, initially, a structural case probe feature for accusative assignment (\([\ast \text{acc}\ast]\)). In addition, there is a \( \text{VP} \) in the workspace with an internal argument \( \text{DP} \) (\( \text{das Buch} \)) and the lexical verb (\( \text{gelesen} \)). In (10-b-ii), \( v \) has undergone Merge with \( \text{VP} \), thereby discharging \([\bullet \text{V}\bullet]\). Next, in (10-b-iii), \( \text{DP}^{\text{ext}} \) is introduced, and \([\bullet \text{D}\bullet]\) is discharged. At this point, the short life cycle of \( \text{DP}^{\text{ext}} \) starts; it becomes syntactically accessible for syntactic processes like those in (8), which require Agree operations into the c-command domain of \( \text{DP}^{\text{ext}} \). However, \( \text{DP}^{\text{ext}} \) is then quickly removed again from the derivation; cf. (10-b-iv). Finally, \( v \)'s structural case probe is deleted, yielding (10-b-v) (where the object \( \text{DP} \) does not have case yet – it will later pick up nominative case via Agree with \( T \)). Crucially, from (10-b-iv) onwards, \( \text{DP}^{\text{ext}} \) cannot be accessed anymore by syntactic operations, for the simple reason that it is not present anymore; this accounts for the observations underlying data such as those in (9).

Note that the short life cycle of \( \text{DP}^{\text{ext}} \) that is indicated in (10) is not an accidental property brought about by a specific initial feature specification of \( v \) but follows systematically from subjecting Remove to the Strict Cycle Condition: A \( \text{DP} \) that is merged in some projection \( \text{XP} \) can only be removed again within that very same projection.\(^{13}\) This derives the ban on passivization of unaccusative verbs (Perlmutter & Postal (1983), pace Primus (2010), Kiparsky (2013)) without further ado; see (11-a) (with an unergative verb, and \( \text{DP} \) merged in Spec\( v \)) vs. (11-b) (with an unaccusative verb, and \( \text{DP} \) merged in \( \text{VP} \)).

(11)  
\( a. \) Hier wird jetzt gearbeitet  
here is now worked

\( b. \)  
(i) \( v(\bullet \text{V}\bullet)[\ast \text{D}\ast][-\text{D}_2-][\ast \text{acc}\ast]; [\text{VP das Buch gelesen}] \)
(ii) \( [v' \bullet \text{V}\bullet][\ast \text{D}\ast][-\text{D}_2-][\ast \text{acc}\ast]; [\text{VP das Buch gelesen}] \)
(iii) \( [v\text{DP}^{\text{ext}} \bullet \text{V}\bullet][-\text{D}_2-][\ast \text{acc}\ast]; [\text{VP das Buch gelesen}] \)
(iv) \( [v\text{VP} \bullet [\ast \text{acc}\ast]; [\text{VP das Buch gelesen}] \)
(v) \( [v\text{VP} \bullet [\text{VP das Buch gelesen}] \)

\(^{13}\) Assuming that movement of \( \text{DP} \) cannot feed removal of \( \text{DP} \) in German passive constructions, it follows that argument removal cannot be attributed to a higher head – say Pass or Voice – than the one that introduces \( \text{DP}^{\text{ext}} \) (i.e., \( v \)). Accordingly, evidence in support of a split Pass/Voice-\( v \) structure (as in Harley (2013), Merchant (2013), and Sundaresan & McFadden (2014)) needs to be reanalyzed. See, again, Müller (2016c). I will return to the issue of movement feeding removal (and thereby extending the life cycle of an item that is subject to removal) in section 4.2 below.
“People are working here now.”

b. *Es wurde angekommen
   it was arrived
   “People arrived.”

Thus, \([-D_2-]\) on \(v\) does not intrinsically stipulate that it is the external argument \(\text{DP}_{\text{ext}}\) that is removed as a consequence of Remove, rather than some VP-internal object \(\text{DP}\). Rather, this effect follows from the Strict Cycle Condition: Structure-building and structure-removal can only take place in the root domain (cf. discussion of (4)). Thus, if \([-D_2-]\) on \(v\) were to target \(\text{DP}\) in the structure \([v[-D_2-]\ [VP\ \text{DP\ V}]]\) that underlies (11-b), Remove\((v[-D_2-],\text{DP})\) would violate the Strict Cycle Condition because \(\text{DP}\) is in the domain of \(\text{VP}\), and \(\text{VP}\) is in the domain of \(v\) (cf. (2), (3)).

To complete this sketch of a Remove-based analysis of passive, it should be pointed out that this analysis does not make it necessary to assume that \(\text{DP}_{\text{ext}}\) is some designated kind of empty category (say, \(\text{pro}\)). As a matter of fact, \(\text{DP}_{\text{ext}}\) can in principle be anything: A referential expression, a pronoun, a \(\text{DP}\) without phonological features, and so on. A permanently removed \(\text{DP}_{\text{ext}}\) typically triggers existential quantification as a default operation once the phase is concluded (which can, however, be overridden under certain circumstances; cf. footnote 11); otherwise fatal recoverability problems would arise. Alternatively, a \(\text{DP}_{\text{ext}}\) that is removed from the structure via a \([-D_2-]\) feature on \(v\), and placed in the workspace, can be remerged into the structure in the only way that is available without structure-building features, viz., as an adjunct. (Incidentally, this mechanism is similar to the renumeration procedure proposed by Johnson (2003) for all subjects and adjuncts.) This then gives rise to \(by\)-phrases; and as one might expect, a \(\text{DP}_{\text{ext}}\) that shows up in a remerged \(by\)-phrase is in principle accessible for operations triggered by higher heads; compare, e.g., (9-a) with (12).

(12) Kein Student\(_1\) gibt zu [\(CP\) dass schlecht [\(PP\ von ihm\_1\ ] gearbeitet wurde ]
   no student admits that badly by him worked was
   “No student admits that he did not work well.”

14 Of course, if \(V\) itself is equipped with a \([-D_2-]\) feature, a VP-internal argument can be affected by Remove. This underlies both antipassives (see Müller (2015a, sect. 3.3)) and applicatives (see the next section).

15 This ensures that a proper name \(\text{DP}_{\text{ext}}\) merged in the matrix clause cannot carry out binding.

16 Some technical questions arise with this approach to \(by\)-phrases. E.g., it must be ensured that the \(\text{DP}_{\text{ext}}\) bears the external argument \(\theta\)-role before before removal and after remerge as a \(by\)-phrase; and it must be ensured that that \(\text{DP}_{\text{ext}}\) accompanied by \(by\) is identified as the original \(\text{DP}_{\text{ext}}\) to begin with. As for the latter question, some indexing device recording the information that the \(DP\) in question was at an earlier stage merged as the external argument of the verb could of course do the job; however, if it is assumed that the derivation proceeds phase-based, with the original numeration partitioned into subarrays (see Chomsky (2000; 2001)), there is in fact no other \(DP\) left in the workspace after removal of \(\text{DP}_{\text{ext}}\), so ambiguities cannot arise. In the same vein, a \(\text{DP}_{\text{ext}}\) will maintain its original \(\theta\)-role, assigned as part of the first, feature-driven, Merge operation, throughout the derivation – the \(by\)-phrase \(\text{DP}_{\text{ext}}\) that re-enters the derivation as an adjunct is the same object as the \(\text{DP}_{\text{ext}}\) that is originally merged in Spec\(v\), it is just accompanied by a preposition that makes it immune from case absorption effects.
Finally, a note on locality. A $\text{DP}_{\text{ext}}$ that is removed from Specv and then subsequently remerged into the structure as an adjunct must do so before the derivation moves on to the next phase.\footnote{This excludes cases as in (i-a), where $\text{DP}_{\text{ext}}$ is remerged as an adjunct in the matrix clause (or, in fact, where it is remerged in the embedded CP phase, and then moved to the matrix domain), and in (i-b) (or (9-a)), where $\text{DP}_{\text{ext}}$ is not base-generated in the matrix clause, but subject to removal in the embedded clause and then remerged as an argument in the matrix domain (however, this latter option is independently ruled out if a DP cannot receive two different $\theta$-roles – i.e., if it cannot be merged twice as a result of two separate $\bullet\text{DP}\bullet$ features).} More generally, it is plausible to assume that the phase is the decisive locality domain for $\text{DP}_{\text{ext}}$ interpretation: A removed argument must be remerged as an adjunct, or triggers default existential quantification, by the end of the phase.

To conclude, there is evidence for downward accessibility and upward inaccessibility of $\text{DP}_{\text{ext}}$ in German passive constructions, and this systematic pattern provides empirical evidence for postulating Remove operations restricted by the Strict Cycle Condition. Clearly, if $\text{Remove}(v', \text{DP}_{\text{ext}})$ exists, there is no way of determining the legitimacy of the earlier $\text{Merge}(v', \text{DP}_{\text{ext}})$ operation by inspecting the resulting output representation (as required under the free Merge approach) because the relevant information has categorically, and irrevocably, been lost (if some trace-like diacritic were retained after structure removal, upward inaccessibility could not be ensured anymore); in contrast, no such problem arises under a feature-driven Merge approach.

3.2. Applicative

Instances of $\text{be}$-prefixation are usually viewed as a canonical case of applicative constructions in German (see, e.g., Stechow (1992), Wunderlich (1993)). In (13-a), $V$ ($\text{laden}$, ‘load’) takes a goal argument realized by a PP ($\text{auf den Wagen}$, ‘onto the wagon’) and a theme argument realized by an accusative DP ($\text{Heu}$, ‘hay’). In (13-b), $\text{be}$-prefixation leads to argument reversal. The theme argument is demoted – it is either realized by a preposition ($\text{mit}$, ‘with’) or does not show up at all; the goal argument loses its preposition and is assigned structural accusative case.

(13) a. dass wir $\text{Heu}$ auf den Wagen laden
   that we$_{\text{nom}}$ hay$_{\text{acc}}$ onto the wagon load
   “that we load hay onto the wagon.”

b. dass wir den Wagen (mit Heu) be-laden
   that we$_{\text{nom}}$ the wagon$_{\text{acc}}$ with hay ‘be’-load

Thanks to a reviewer for bringing up this scenario.
“that we load the wagon with hay.”

In what follows, I will adopt a version of an approach to applicative formation going back to Baker (1988) and (for German) Stechow (1992). On this view, the structure of vP in (13-a) looks roughly as in (14), generated on the basis of a verb laden\([\bullet P\text{-}\bullet D\text{-}}\):

\[
(14) \quad \text{vP} \rightarrow \text{DP} \rightarrow \text{VP} \rightarrow \text{vP}
\]

\[
\text{DP} \rightarrow \text{DP} \rightarrow \text{VP} \rightarrow \text{VP} \rightarrow \text{vP}
\]

\[
\text{DP} \rightarrow \text{PP} \rightarrow \text{V} \rightarrow \text{V} \rightarrow \text{vP}
\]

\[
\text{DP} \rightarrow \text{P} \rightarrow \text{DP} \rightarrow \text{laden}
\]

The structure in (14) basically also functions as the input to (13-b). Under the Baker-Stechow approach, the sole difference is that P is be instead of auf and needs to incorporate into V. Incorporation of P then implies that the goal DP den Wagen cannot receive case from P anymore, so v steps in and assigns case to it, which in turn means that the theme DP Heu must become oblique. However, on this view it is not quite clear in what sense the theme DP can be said to be demoted in the applicative – it occupies exactly the same structural position as before, the only difference being that it needs to be supported by a case-assigning preposition. Furthermore, it is unclear why the theme argument should become optional in (13-b). Both problems are solved if structure removal is added to the approach: Under this assumption, the applicative is triggered by a co-occurrence of P incorporation and a \([-D_2-]\) feature added to V in the numeration, yielding \(V[\bullet P\text{-}\bullet D\text{-}][\bullet D\text{-}][\bullet D_2\text{-}]\). The resulting structure looks as in (15), with the theme argument removed from the clause.

---

18 I will not consider an approach where applicatives can be traced back to specific functional heads (like Appl) that introduce arguments (see Pylkkänen (2000), among many others). While such an approach (or a modification of it, as in Hole (2014)) may well be correct for other constructions in German that can be called “applicative” (e.g., free dative constructions), it cannot straightforwardly capture the argument reversal effect with be-prefixation.

19 As noted by Stechow (1992), be can be viewed as a reduced form of bei (‘with’), which can still be used as a local preposition instead of auf in (13-a) in non-standard German varieties.
As before, it is neither necessary nor possible to specify which DP will be removed by the [–D2–] feature on V: The Strict Cycle Condition ensures that only the theme DP can be targeted in (14). As a consequence of Remove(V',DP), the theme argument Heu is taken out of the structure and put in the workspace of the derivation. Optionally, it may then re-enter the structure as an adjunct to VP, accompanied by the appropriate preposition (see Baker (1988) on what motivates the choice).

With these assumptions in place, let me now turn to the predictions that the analysis makes for the accessibility of the theme argument in German applicative constructions: Applicatives as in (13-b) are expected to exhibit short life cycle effects, with downward accessibility and upward inaccessibility. And indeed, the available empirical evidence points to this conclusion. Thus, (16-a) shows that in the absence of applicative formation, the theme DP can control the PRO subject of a secondary predicate. Crucially, (16-b) illustrates that such control is still possible when applicative formation applies, and the theme DP is removed from the VP (it may or may not subsequently re-enter the structure as an adjunct).

Two further remarks. First, if only left-adjunction is an option, or if V does not move to v, a further scrambling operation applying to the goal DP is then required to derive the unmarked order in (13-b). Second, the analysis just sketched presupposes that P incorporation and DP removal co-occur so as to trigger applicative formation. Given that both these operations are in principle optional, the question arises of what happens if one occurs without the other. Suppose first that be is the P head (i.e., incorporation takes place) but [–D2–] does not show up on V. In that case, there will be two DPs that need to be assigned case, but there is only one case available (viz., [∗acc] on v). This accounts for the ungrammaticality of (i-a).

(i) a. *dass wir Heu den Wagen beladen
   that we nom hay acc the wagon acc load
   “that we load hay onto the wagon.”

b. *dass wir (mit Heu) auf den Wagen laden
   that we nom (with hay) onto the wagon load
   “that we load hay onto the wagon.”

Alternatively, [–D2–] occurs on V but there is no P incorporation, as in (i-b). In this case, there will not be any DP left that requires accusative case from v, and this can be taken to violate a constraint like the Inverse Case Filter (see Bošković (2002)). Note that this reasoning is compatible with case probe deletion as assumed above for the passive (see footnote 12) if it is assumed that case probe deletion must be extremely local, involving information within the same head only. Still, as noted by a reviewer, attributing the illformedness of (i-b) to an Inverse Case Filter is not an innocuous assumption, as there are several well-
“One pours the water over the freshly chilled berries when it is hot.”

b. Man be-gießt DP\textsubscript{theme} dann die gut gekühlten Beeren [SC PRO\textsubscript{1} heiss ] (mit dem Wasser) the water)
“One pours the water over the freshly chilled berries when it is hot.”

The same pattern shows up with the examples in (17-ab); (17-b) can have a reading where what is loaded onto the wagon is wet (in addition to the obvious alternative readings where we are wet, or where the wagon is wet).

(17) a. Wir laden das Heu\textsubscript{1} [SC PRO\textsubscript{1} nass ] auf den Wagen
“We load the hay onto the wagon when it is wet.”

b. Wir be-laden DP\textsubscript{theme} den Wagen [SC PRO\textsubscript{1} nass ]
“We load the wagon with something wet.”

Data like (16-b) and (17-b) strongly suggest that the theme argument is accessible for c-command in applicative constructions even though it does not have to be overtly realized (and if it is, it is embedded in a PP which should block c-command). This follows from the approach to applicatives in terms of structure removal: Control is effected after the theme DP has been merged, and before it is removed.\textsuperscript{21}

\textsuperscript{21} Reciprocals (and reflexives) fail to provide an argument for downward accessibility of the theme DP in German applicatives, see (i-a) vs. (i-b).

(i) a. Wir setzen die Spielfiguren\textsubscript{1} auf einander\textsubscript{1}
“We put the pawns on top of one another.”

b. *Wir besetzen DP\textsubscript{int}\textsubscript{1} einander\textsubscript{1} (mit den Spielfiguren)
“We put the pawns on top of one another.”

I take the illformedness of (i-b) to have an independent source (that is possibly related to a combination of recoverability problems and the general markedness of reciprocal/reflexive binding among objects in German). Similar considerations apply to control constructions involving adjunct clauses (as in (ii-a), brought up by a reviewer) rather than secondary predicates, where a DP\textsubscript{theme} of an applicative cannot effect control even though it should be accessible at the relevant point of the derivation. The
In contrast, the theme DP is inaccessible for operations triggered by higher heads; for instance, as shown in (18-a) vs. (18-b), variable binding by a matrix clause quantified DP is impossible unless the theme argument is reintroduced into the structure as part of a PP.

(18) a. *Kein Student will [CP dass man DP-wagen den Wagen belädt] no student wants that one the wagon loads
   “No student wants that one loads the wagon with him.”

   b. Kein Student1 will [CP dass man den Wagen mit ihm belädt] no student wants that one the wagon with him loads
   “No student wants that one loads the wagon with him.”

To conclude, the fact that the theme argument in German applicative constructions exhibits downward accessibility and upward inaccessibility provides an independent argument for an approach to applicatives in terms of structure removal. However, it is clear that if an approach along these lines is on the right track, there is no way how the legitimacy of an initial Merge operation that introduces the theme DP could be checked by inspecting the output representation once Remove has applied (as required in the free Merge approach): One would wrongly expect bleeding. Again, the feature-driven Merge approach does not face any problem since Merge(V',DP) is counter-bled by Remove(V',DP).

4. Removal of Y: Reanalysis

While Remove (X(′),YP) takes whole constituents out of syntactic structures, Remove(X(′),Y) merely results in the elimination of the top layers of constituents. This offers a new approach to various phenomena that provide evidence for conflicting representations which seem to require some concept of reanalysis. The existing models of reanalysis either involve unconstrained reanalysis rules (cf., e.g., Bach & Horn (1976) and Chomsky (1977) on extraction from DP, Chomsky (1981) on S-bar deletion, or De Kuthy & Meurers (2001) on verbal complexes), or they rely on multidimensional representations (see Huybregts (1982), Bennis (1983), Haegeman & Riemsdijk (1986), Di Scullo & Williams (1987), Sadock (1991), and Pesetsky (1995)), which are both extremely powerful and empirically problematic (see Chomsky (1982)). In contrast, a removal-based approach to reanalysis phenomena is highly constrained (given the Strict Cycle Condition, and given the limited effects on existing structures that it can have), and it makes hypothesis that the illformedness of (ii-a) has an independent source is confirmed by the observation that an overt DPtheme in a minimally different example without applicative formation also cannot control into the adjunct clause; cf. (ii-b).

(ii) a. *Man be-gießt DPtheme die Elektrode [CP um PRO1 zu verschwinden] one ‘be’-pours the electrode in order to disappear
   “One pours some liquid on the electrode in order to make the liquid disappear.”

   b. *Man gießt Wasser1 auf die Elektrode [CP um PRO1 zu verschwinden] one pours water on the electrode in order to disappear
   “One pours water on the electrode in order to make it disappear.”

In both cases, it looks as though object control into the adjunct clause is blocked by the availability of subject control.
systematic predictions concerning accessibility of material that is subject to reanalysis.

In this section, I will discuss two pertinent phenomena of German syntax, viz., restructuring infinitives and complex prefields.

4.1. Restructuring

Whereas non-restructuring infinitives behave in virtually all relevant respects like finite embedded clauses and thus uniformly demand a biclausal analysis in terms of CP embedding, with restructuring infinitives there is both evidence for monoclausality (i.e., for the absence of at least a CP shell, possibly also of a TP or vP shell) and evidence for biclausality. Among the well-known pieces of evidence in favour of a monoclausal analysis of restructuring infinitives are the following properties (see, e.g., Stechow & Sternefeld (1988), Grewendorf (1988), Fanselow (1991), Bayer & Kornfilt (1994), and Haider (2010)): Restructuring infinitives cannot undergo extraposition; a negative item in the infinitive can optionally take wide scope; items may scramble out of the infinitive into the matrix domain; there is status government (cf. Bech (1955/1957), Stechow (1990)) – i.e., “verbal case assignment” (cf. Fabb (1984), and Adger (2003) for a more recent technical implementation) – among the verbs participating in the construction; there is pied piping of infinitives; verb projection raising may occur; and the intonation may signal monoclausality. Let me just focus on two of these properties here. First, a matrix verb like versuchen (‘try’) (see (19-a)) that optionally brings about restructuring can trigger wide scope of an embedded negative element (cf. the reading in (19-a-i)), in addition to the more marked option of embedded negative scope (cf. (19-a-ii)); as indicated in (19-b), a non-restructuring matrix verb like bedauern (‘regret’) cannot do so (the wide scope reading in (19-b-i) is unavailable, in contrast to the embedded reading in (19-b-ii)).

(19) a. Sie hat nichts zu sagen versucht
   she has nothing to say tried
   (i)  She did not try to say anything.
   (ii) She tried not to say anything.

b. Sie hat nichts gesagt zu haben bedauert
   she has nothing said to have regretted
   (i)  #She did not regret that she had said something.
   (ii) She regretted that she had not said anything.

Note that the amalgamation of nicht (‘not’) and an indefinite pronoun, as in nichts (‘nothing’) (also known as a “kohäusive Verbindung” in the German literature on the topic), is confined to membership in the same clause.

---

22 As noted by Fanselow (1989; 1991), there is some variation among speakers as to which verbs count as (non-)restructuring predicates in German; informally, an emerging generalization would seem to be that the younger the speaker, the more verbs (s)he accepts as a restructuring predicate. This does not affect the point of the main text.
Second, as shown in (20-ab), scrambling is known to be a clause-bound process in German (see Ross (1967)).

(20) a. *dass ihn1 der Oberförster sagte [ dass Peter t1 treffen soll ]
    that him the head forester said that Peter meet shall
    “that the head forester said that Peter should meet him.”

    b. *dass ihn1 der Oberförster sagte [ solle Peter t1 treffen ]
    that him the head forester said should Peter meet
    “that the head forester said that Peter should meet him.”

However, with restructuring infinitives scrambling of items subcategorized by the embedded predicate in front of matrix material is unproblematic; see (21-a) (with the restructuring verb versuchen (‘try’)) vs. (21-b) (with the non-restructuring verb bezweifeln (‘doubt’).

(21) a. dass sich1 der Oberförster1 t1 zu rasieren versuchte
    that REFL the head forester to shave tried
    “that the head forester tried to shave himself.”

    b. *dass sich1 der Oberförster1 [ t1 rasiert zu haben ] bezweifelte
    that REFL the head forester shave to have doubted
    “that the head forester doubted that he had shaved himself.”

Thus, there is evidence for a monoclausal analysis. On the other hand, there is also evidence for a bi-clausal analysis of restructuring infinitives in German. A first argument goes back to Stechow & Sternefeld (1988); it consists in the observation that every control verb that permits restructuring can optionally also show up in a non-restructuring context. This implicational generalization must remain a mystery if restructuring predicates can simply optionally involve TP-embedding, vP-embedding or VP-embedding, but it is directly accounted for if the only way to end up with such a smaller complement size is via an initial CP embedding that is then subject to some reanalysis operation. A second traditional argument emerges from the generalization that the subject of a restructuring control infinitive can never be realized by an overt DP; this restriction can be tied to the presence of a CP shell.23 A third, more empirical, argument is based on the observation that restructuring never creates new binding domains. To see this, consider the examples in (22). The restructuring verb versprechen (‘promise’) is a subject control verb. As one would expect, an embedded object reflexive pronoun can be locally bound by the non-overt subject PRO; see (22-a). The matrix object ihm (‘him’) cannot act as an antecedent for the reflexive; see (22-b). However, under a monoclausal approach, this fact actually raises severe problems: If there is no

---

23 Following Adger (2003), it can be assumed that control predicates embedding infinitival clauses select a special type of complementizer which in turn assigns a case-like feature [null] to the embedded subject that requires a non-overt realization not just of the ending, but of the whole argument DP (as PRO). (This implies that even after a CP shell has been removed in restructuring contexts, an embedded subject DP in a control infinitive can never be case-marked by an item in the matrix clause, and be overtly realized, even if it becomes structurally accessible to it; see below.)
local binding domain which clearly separates the arguments belonging to the embedded predicate (PRO, *sich*) from the arguments belonging to the matrix predicate (*der Oberförster* (‘the head forester’), *ihm* (‘him’)), with all arguments belonging to one and the same local domain, given restructuring, then one would expect the reflexive pronoun *sich* to be able to freely pick its antecedent from the set of accessible items in the same way that this is possible for an (accusative) object reflexive in a double object construction; cf. (22-c) (cf. Sternefeld & Featherston (2003), based on the experimental study reported in Featherston & Sternefeld (2003), according to which binding of a reflexive by a dative is possible, and actually preferred if the antecedent is pronominal). Of course, this problem is only amplified if one assumes that a restructuring infinitive does not even have a PRO subject.

(22) a. Der Oberförster₁ hat ihm₂ (PRO₁) sich₁ zu waschen versprochen
   the head forester has him₂ REF to wash promised
   “The head forester promised him to wash himself.”

   b. *Der Oberförster₁ hat ihm₂ (PRO₁) sich₂ zu waschen versprochen
      the head forester has him₂ REF to wash promised
      “The head forester promised him to wash him.”

   c. Der Oberförster₁ hat ihm₂ sich₁/₂ im Spiegel gezeigt
      the head forester has him₂ REF in the mirror shown
      “The head forester showed him himself in the mirror.”

Thus, (22-b) poses a challenge for a purely monoclausal approach, but it is directly accounted for under a biclausal approach, where CP acts as a locality domain for reflexivization.

As with the passive, it would seem that most existing approaches to restructuring strictly rely on either a monoclausal approach (see Haider (1993; 2010), Kiss (1995), and Wurmbrand (2001), among many others) or a biclausal approach (see Baker (1988), Sternefeld (1990), Müller & Sternefeld (1995), and Koopman & Szabolcsi (2000)). Evidence that points in the opposite direction is then typically accommodated by additional stipulations, or an attempt is made to invalidate it. Alternatively, a genuine reanalysis approach can be pursued according to which a regular CP embedding is optionally reanalyzed as a monoclausal configuration, via one of the unrestricted mechanisms mentioned above (see Rizzi (1982), Aissen & Perlmutter (1983), Haegeman & Riemsdijk (1986), Di Sciullo & Williams (1987)).

From the present perspective, a simple resolution of the conflict created by incompatible structure assignments required in restructuring contexts suggests itself. Evidence for monoclausality implies inaccessibility of CP (TP, ...) shells for syntactic operations; evidence for biclausality implies accessibility of the CP shell for syntactic operations; and as before, structure removal in the course of the derivation can reconcile the conflicting demands in a principled way. Here, then, is a sketch of a new reanalysis approach based on structure removal: Suppose that restructuring verbs uniformly embed CPs, just like non-restructuring verbs. However, they optionally come equipped with Remove-triggering features that can then successively peel off CP (TP, ...) layers from the complement of the restructuring verb: \([-C_0=-]\)
([−T0−], ...). The clausal shells thus affected are therefore predicted to exhibit short life cycles.

Evidence that presupposes biclauasality implies operations that need to be carried out and/or checked before structure removal (they are counter-bled and counter-fed by structure removal). This includes subcategorization of CP (via [●C●]) by all restructuring verbs (which accounts for the fact that there are no control restructuring verbs that cannot optionally preserve full biclauasality). It also holds for the non-extendability of binding domains by restructuring: A reflexive pronoun picks an antecedent in the minimal CP, and the embedded subject will always qualify as such a potential antecedent, thereby providing an index for the reflexive pronoun, via Agree – subsequent removal of the CP shell cannot change matters anymore because it cannot lead to overwriting of an existing index. Finally, given that the question of overt vs. non-overt realization of a subject DP in infinitives is decided on the basis of the absence or presence of a (specific type of) CP projection, the CP that is initially present in restructuring contexts ensures non-overt realization (as PRO); the embedded subject DP cannot change its feature [null] again. In all these three cases, there is thus counter-bleeding or counter-feeding by subsequent Remove(V,CP).

In contrast, evidence that suggests monoclausality involves operations that apply after Remove(V,CP) since they also involve structure on top of the matrix VP (given the Strict Cycle Condition). This is patently evident with long-distance scope of negation (see (19)) and long-distance scrambling (see (21)), but it holds more generally for all arguments in favour of monoclausality that have been given in the literature. So, all evidence for monoclausality involves transparent bleeding and feeding by Remove in the present analysis.

The question arises of what predictions this approach makes for a combination of wh-movement and scrambling from restructuring infinitives, as in (23).

(23) Wem1 hat ihn2 Maria t1 t2 vorzustellen versucht?
whom dot has him acc Maria introduce tried
“What did Maria try to introduce to him?”

Under standard assumptions, the wh-phrase must first undergo movement to the embedded SpecC position at the point when the embedded CP is generated; such a step will be required by a locality constraint like the Phrase Impenetrability Condition (Chomsky (2001)) in any theory that adopts it. Thus, suppose that the next phase edge required for an intermediate wh-movement step is Specv, and suppose furthermore that SpecV is not accessible for intermediate wh-movement steps. Then, wh-movement cannot take place from SpecC before Remove(V[−C0−],CP). This latter operation then brings about a structure-preserving reassociation of the wh-phrase (i.e., the original specifier of C) and the TP (i.e., the original complement of C) with the matrix VP. Subsequently, potential further removal operations triggered by V may take place. After that, on the vP cycle, the external argument Maria is merged; the wh-phrase wem1 moves from the SpecV position which it now occupies (without ever having been moved there) to
Specv; and the unstressed pronoun *ihn* is attracted to a position in front of the subject.\(^{24}\)

To end this subsection, I would like to highlight an orthogonal but potentially interesting property of the approach to restructuring in terms of structure removal just sketched: It is perfectly conceivable that different kinds of restructuring verbs can have differently many features for structure removal (e.g., just \([-C_0-]\), or \([-C_0-]\) and \([-T_0-]\), or \([-C_0-], [-T_0-], \) and \([-v_0-]\)), which will (ultimately) result in restructuring infinitives of different sizes, depending on the amount of structure that is successively removed by the matrix verb; and this has in fact been argued for in the literature (see, e.g., Fanselow (1991) and Wurmbrand (2001; 2015)).\(^{25}\)

From the more general point of view of deciding between feature-driven Merge and free Merge, it should be clear that to the extent that structure removal is well motivated for restructuring, this domain, too, provides an argument against the latter approach: After removal of a complement CP shell, it cannot be decided whether the original Merge(V,CP) operation is legitimate by solely inspecting the output representation.

4.2. Complex Prefields

Normally, only one item may show up in the area before the finite verb in German main clauses (the verb-second property). However, in the complex prefield construction, two (or more) items can occur in the domain preceding the finite verb in C; see (24-ab).\(^{26}\)

\[(24)\]
\[\text{a. } \text{DP } \text{Den Fahrer } \text{PP } \text{zur Dopingkontrolle } \text{begleitete ein Chaperon to the doping test accompanied a chaperon} \]

“A chaperon accompanied the rider to the doping test.”

\(^{24}\) This analysis is reminiscent of Chomsky’s (1981) analysis of wh-movement from raising constructions in English, as in (i).

\[\text{(i) What}_1 \text{ do you believe John to have done } t_1?\]

Based on a suggestion from Luigi Rizzi, Chomsky (1981, 66) argues that (i) involves a combination of intermediate wh-movement to (what would now be called) SpecC (as required by the Subjacency Condition) followed by V-governed S-bar deletion (i.e., removal of the CP shell, in current terms). However, there is a crucial difference: S-bar deletion automatically removes an intermediate trace (note that SpecC and C were assumed to be one single category COMP, so this operation must apply very late in the derivation, after wh-movement to the final landing site, and will therefore invariably qualify as massively counter-cyclic); in contrast, structure removal just removes the head C and reassociates the wh-phrase in the matrix VP domain.\(^{25}\)

\(^{25}\) Note that no restrictions are needed on the possible combinations and orders of Remove features on restructuring verbs. A \(V[\{c\}*][-C_0-][-T_0-]\) first takes a CP complement, next cuts its back to TP status, and finally removes the TP shell, resulting in vP status of the complement. However, a minimally different \(V[\{c\}*][-T_0-][-C_0-]\) that reverses the Remove features, or a minimally different \(V[\{c\}*][-T_0-]\) that does without \([-C_0-]\) altogether, will never result in successful structure removal: On the VP cycle, V cannot bring about removal of TP via an intervening CP because TP is too deeply embedded, and the operation is blocked by the Strict Cycle Condition.

\(^{26}\) The construction frequently shows up in live sports broadcasts, perhaps particularly so with bike races; this is reflected by lexical choices in the examples of this subsection. Also note that whereas most of the examples in this section are confined
b. [pp Mit dem Hauptfeld] [pp ins Ziel] kamen auch Fernando Escartín und Aitor Garmendia into the finish came also Fernando Escartín and Aitor Garmendia

“Escartín and Garmendia arrived at the finish with the peloton.”

There are two competing analyses of the phenomenon. On the one hand, it has been assumed that prefields can be truly complex under certain circumstances. On this view, there are two (or more) separate constituents in the prefield in (24), as a consequence of an option of multiple fronting (cf. Lötscher (1985), Speyer (2008)); cf. (25). On the other hand, it has been argued that prefield complexity is only apparent. Under this approach, there is a single constituent in the prefield in (24), viz., a fronted VP with an empty head; cf. (26). This empty head may then be a trace resulting from prior head movement (cf. Müller (1998)), or it may be a separate empty head that does not (directly) participate in a displacement configuration (cf. Fanselow (1992), Müller, St. (2005)).

(25)

```
(25)  CP
     \   /
   XP_1 C'  
   |   |
   YP_2 C'
   |   TP
     ..t_1..t_2..
```

(26)

```
(26)  CP
     \   /  
   VP_0 C'  
   |   |
   XP_1 V'  
   |   C  
     TP
     ..e..t_0..
```

Again, closer inspection reveals that there is evidence both for single constituency and for multiple constituency in complex prefields in German. An argument for single constituency (as in (26)) is based on the fact that the items that show up in a complex prefield must be clause-mates (cf. Fanselow (1992)); see (27-a) (where the two fronted items are clause-mates) vs. (27-b) (where the two items originate in different clauses and thus cannot be part of a single VP lacking an overt head). This follows if it is a single VP constituent that undergoes the movement, but not if two items can move separately.

(27) a. [CP [VP_0 Fahrern_1 EPO_2] [C_1 sollte man besser nicht t_0 geben]]

    “One should not give riders EPO.”

b. *[CP Fahrern_1 EPO_2] [C_1 sollte man besser nicht t_1 nachsagen [CP dass sie t_2 genommen haben]]

    “One should not say about riders that they have taken EPO.”

to two items preceding the finite verb, the construction can in principle accommodate arbitrarily many items (see Fanselow (1992); Müller, St. (2005); Müller, St. et al. (2015)).
Similarly, Müller, St. (2005) observes that the ordering restrictions among multiple items in complex prefields are identical to those in the middle field; see (28-ac) (with unmarked order of dative and accusative object) vs. (28-bd) (with a marked order). This generalization follows directly if the prefield constituent is the middle field constituent but would have to qualify as spurious if there were separate movements of two items to SpecC positions.

(28)  
a. \begin{array}{l}
[CP \text{ Fahrern} EPO_2 \ \\
\quad \text{ riders}_{dat} \text{ erythropoietin}_{acc} \text{ sollte \ man \ besser \ nicht \ geben }] \end{array}

“One should better not give riders EPO.”

b. \begin{array}{l}
?[CP \ EPO_2 \ \\
\quad \text{ Fahrern}_1 \text{ should \ man \ besser \ nicht \ geben }] \end{array}

“One should better not give riders EPO.”

c. \begin{array}{l}
\text{ dass } \text{ man } \text{ Fahrern}_1 \text{ Epo}_2 \text{ gegeben hat } \ \\
\text{ that one } \text{ gave riders EPO.”} \end{array}

d. \begin{array}{l}
?[\text{ dass man } \text{ Fahrern}_1 \text{ Epo}_2 \text{ gegeben hat } ] \end{array}

“that one gave riders EPO.”

However, there is also evidence for multiple constituency. A first argument for this comes from freezing effects (see Ross (1967), Wexler & Culicover (1980)), according to which moved items are islands for further extraction even if these items are transparent for extraction in situ. Indeed, extraction from an item in a complex prefield exhibits a freezing effect. To see this, consider the examples in (29). (29-a) is a complex prefield construction with a DP and a PP headed by zu (‘to’). (29-bc) show that this type of PP permits postposition stranding, with an R-pronoun da topicalized to a (non-complex) prefield position and to a middle field-internal scrambling position, respectively. In (29-d), such postposition stranding takes place via scrambling within a fronted regular VP (with an overt V head), with PP uncontroversially in its in situ position, from which extraction via scrambling is unproblematic, exactly as in (29-c). Against this background, (29-e) illustrates a freezing effect in the complex prefield position: PP does not permit extraction here even though it does in other contexts. This strongly suggests that PP does not occupy a base position in (29-e), which in turn favours the multiple constituency analysis in (25).

As noted by a reviewer, (i) (= (29-a), with zum Erfolg (‘to the success’) replaced by da-zu (‘there-to’) already exhibits reduced acceptability.

(i)  
?[[CP Dem Team \ [pp \ da-zu ] \ [C’ gratulierte Bernard Hinault ]] \ 
\quad \text{ the team}_{dat} \text{ there-to } \text{ congratulated Bernard Hinault}_{nom} \]

“Hinault congratulated the team on the success.”

However, it is unlikely that suffices as an account of (29-e). First, there is still a marked difference in acceptability between the two examples. And second, without attempting to go into details here, it seems likely that the reduced acceptability of (i) is due to difficulties with imposing an intonational pattern on the fronted items that is typically required by the complex prefield
A second argument for multiple constituency involves Barss’ generalization (cf. Barss (1986), Sauerland & Elbourne (2002), Bhatt & Dayal (2007), Neeleman & van de Koot (2010), Heck & Assmann (2014)), according to which a quantified item \( \gamma \) contained in a moved XP \( \alpha \) cannot take scope, via reconstruction, over an item \( \beta \) that c-commands \( \alpha \)’s trace and is c-commanded by \( \alpha \); see (30).

(30)  
\[ \text{Barss’ generalization scenario:} \]
\[ [\alpha \ldots \gamma \ldots] [\ldots \beta \ldots [\ldots t_\alpha \ldots]] \]

a. Scope: \( \beta \succ \gamma \): \( \checkmark \)

b. Scope: \( \gamma \succ \beta \): *

It follows from Barss’ generalization that whereas (31-a) is ambiguous (with either wide or narrow scope of the fronted universal quantifier object DP \textit{jeden Fahrer} (‘every rider’)), (31-b) is not: the object DP cannot have wide scope over the existential quantifier subject DP. Importantly, as shown in (31-c), complex prefields do not trigger Barss’ generalization effects: The universal quantifier object can have scope over the existential quantifier.\(^{28}\) This suggests that the object DP \textit{jeden Fahrer} is not part of a fronted VP in (31-c), unlike what is uncontroversially the case in (31-b).

(31)  
a. \[ \text{[DP Jeden Fahrer ] begleitet ein Chaperon zur Dopingkontrolle every rider, acc. accompanies a chaperon, nom, to the doping test} \]
\[ \forall \succ \exists: \text{“For each rider, there is a chaperon who accompanies him to the doping test.”} \]
\[ \exists \succ \forall: \text{“There is a chaperon who accompanies each rider to the doping test.”} \]

\(^{28}\) Judgements are subtle here, but very clear for almost all speakers I have consulted, and hold under benevolent modulation of intonation contours in all cases. Note also that the unavailable reading in (31-b) is the vastly more salient one, given what is known about chaperons.
Given these observations (as well as several others, related, inter alia, to weak crossover, negative polarity items, left dislocation, and extraposition, which are highlighted in the much more comprehensive study of the phenomenon developed in Müller (2015b)), the conclusion can be drawn that there is conflicting evidence as to what the structure of complex prefields in German looks like: The observations based on (27) and (28) support a VP fronting structure as in (26), whereas the observations in (29) and (31) favour a multiple movement structure as in (25). By now, it should be clear how this conflict can be resolved systematically: An initial VP topicalization structure gets reanalyzed as a multiple fronting structure, as a consequence of a [–V₀–]-induced operation that removes the VP shell in SpecC.

As a first step towards such an analysis, recall from the discussion of (7) that there is nothing in the approach to structure removal sketched in section 2 above that would preclude internal Merge (movement) of some item to a specifier position feeding subsequent Remove of this item; as noted above, this is the only way how material that is subject to removal can extend its life cycle beyond what would otherwise be expected given the Strict Cycle Condition. For concreteness, suppose that in complex prefield constructions, remnant VP fronting (triggered by [V•] on C, or by some other movement-triggering feature on C targeting the VP) feeds removal of the VP shell (triggered by [–V₀–] on C). The derivation given in (32) shows how reanalysis in complex prefields is brought about. The first step is that V has left the VP, thereby creating a remnant VP from which the verb is missing; see (32-a). Next, in (32-b) VP topicalization takes place. Finally, structure removal takes place. In (6) and (7) above, I have illustrated this by a single representation. This time, for the sake of clarity, the two steps that are required for this are indicated in two separate representations, viz., (32-c) (where the VP shell is removed as a consequence of C’s [–V₀–] feature, thereby creating two floating phrases that were part of VP’s minimal domain) and (32-d) (where the floating daughters XP₁ and YP₂ of the original VP are reassociated with the triggering

---

29 Also see Murphy (2014) on such an interaction of movement and structure removal. – Note incidentally that in order to maintain the ban on passivization of unaccusatives in German (cf. discussion of (11) above), it must be assumed that the internal argument DP cannot undergo movement to Specv in this context, at least not prior to [–D₂–] discharge by v. For the time being, I will leave open the question of why this should be so, and whether it might ultimately reflect a deeper asymmetry between [–F₀–] and [–F₂–] features.

30 In (32), e is the trace of a moved lexical V. V may be in C or in a TP-internal right-peripheral position adjoined to some functional head; this must hold irrespective of whether V is finite or non-finite (e.g., a past participle).
head’s projection in a structure-preserving way).

(32) a. *Pre-movement structure:*

```
C'    
C[STOP]→[−V₀−]  TP
   /\  
  ... VP₀ ...  
 /   \      /   \  
XP₁  V'    VP₀  YP₂  V
   \   \   /   \   /   \  
   e   e  ... t₀ ...
```

b. *VP fronting:*

```
CP  
  /     \  
 VP₀   C'    
   /\      /   \  
  XP₁  V'  C[−V₀−]  TP
   /   \   /   \   /   \  
  YP₂  V  ... t₀ ...  e
```

c. *Structure removal:*

```
CP  
  /     \  
  C'    C   TP
   /\      /   \  
  XP₁  C    C  YP₂
   /\      /   \   /   \  
   ... t₀ ...  ... t₀ ...
```

d. *Reassociation:*

```
CP  
  /     \  
 XP₁  C'    C'    
   /\  /\  /\  
  YP₂ C  C  TP
   /\   /\  /   \  
   ... t₀ ...
```

Thus, movement of an item that is eventually targeted by structure removal (here: the VP) can extend its life cycle somewhat. However, downward accessibility/upward inaccessibility of the item is ensured
as before. Consequently, the prediction is that the evidence for a single VP constituent involves earlier (lower) stages of the derivation (cf. (32-ab)); evidence for multiple constituents involves later (higher) stages of the derivation (cf. (32-d)). The seemingly contradictory properties of complex prefields in German can now be accounted for.

First, the clause-mate condition (see (27)) follows from the assumption that root C has only one structure-building feature for topicalization in German; so only a single constituent (like VP) can move to the prefield.

Second, order restrictions are identical in VP and in the prefield (see (28)) because the item is identical: The only option for VP-internal material to undergo reordering (e.g., by scrambling) is when VP is still in situ. Movement of, say, YP₂ within VP after VP topicalization in (32-b) would violate the Strict Cycle Condition; and movement of YP₂ within CP after VP removal in (32-d) is impossible because all structure-building operations must be triggered by designated features (including, on this view, scrambling), and given that root C has only one structure-building feature for movement to begin with (which it has discharged by attracting a VP), there can be no [F] feature left that might trigger XP₁-YP₂ reordering.

Third, the freezing effect (see (29)) follows if the locality constraint that ultimately derives freezing in general is not derivational but applies to output representations (cf. Browning (1991), among many others). The reason is that after structure removal, YP₂ in (33) occupies a (derived) specifier position that is representationally indistinguishable from a position occupied as a consequence of movement (or other specifier positions which also block extraction, for that matter) – in this way, removal of one category (VP) can result in a structural placement of another category (YP) that is otherwise only attainable under movement. Thus, if the freezing effect can be viewed as an instance of a general prohibition against extraction from specifiers (cf. Huang (1982)), its presence in (29-e) is accounted for.³¹

Finally, concerning Barss’ generalization (see (31)), relative scope is an LF-related phenomenon that is determined on the basis of output representations, i.e., after structure removal. Hence, at the relevant stage, there is no VP anymore that might prevent a prefield item from taking scope over a middle-field internal item.³²

³¹ An alternative account of the freezing effect in complex prefields that is based on a strictly derivational rather than representational approach to freezing is developed in Müller (2016a), and argued there to be superior in view of the absence of freezing effects with remnant movement. However, I refrain from laying out the derivational approach here because its presentation would take an inordinate amount of space, and the question is somewhat orthogonal to my main concerns.

³² Note that this approach does not license multiple overt wh-dislocation via VP fronting and V removal in German, as in (i).

(i) *[CP Wem₁ was₂ [C₁w] hat ] Karl t₁ t₂ geschenkt ]?
   whom.dat what.acc has Karl.nom given
   “Whom did Karl give what?”
Although there are several further issues that will eventually need to be addressed on the basis of this new reanalysis-based approach to complex prefields, I will leave it at that for present purposes.\(^{33}\) As before, the more general conclusion I would like to draw is that there is good empirical evidence for postulating structure removal with complex prefields; and since structure removal leads to opacity (because important information of an earlier stage of the derivation is ultimately lost), this then favours feature-driven Merge over free Merge.

5. Conclusion and Outlook

In sections 3 and 4, I have presented empirical evidence in support of a Remove operation that functions as a counterpart of Merge. A common property of all the relevant data (from passive, applicative, restructuring, and complex prefield constructions in German) is that they suggest conflicting representations at work, where neither one can be dispensed with in favour of the other, and which do not lend themselves to accounts in terms of movement. To conclude this paper, I would like to briefly consider some conceptual issues raised by an operation of structure removal, and then address the consequences for the overarching question of feature-driven Merge vs. free Merge more generally.

First, one might ask whether an operation like Remove that radically alters syntactic representations violates basic syntactic principles. This does not seem to be the case. As a matter of fact, the only well-established constraint that Remove violates is the Projection Principle (Chomsky (1981)), which bans removal of thematically relevant structure. However, the Projection Principle has arguably always qualified as dubious since it can only be formulated as a global rule (see Lakoff (1971)), in the sense that in order to find out whether it is respected or not, non-adjacent steps of the derivation must be compared; thus, it is clear that it cannot be maintained in a current minimalist approach for principled reasons.

A related question concerns semantic interpretation. Here I would like to acknowledge that structure removal may indeed lead to incompatibilities with the standard concept of transparent logical forms as laid out, e.g., in Heim & Kratzer (1998); but the questions that this raises are not qualitatively different from questions raised by cyclic spell-out to LF (and PF) as it is standardly adopted in minimalist work (Chomsky (2001; 2013)). For concreteness, let me enumerate the requirements that an approach to semantic interpretation must meet in order to accommodate the assumptions made in the present paper. First, referential indices (exist and) are invariantly assigned during the syntactic derivation; variable

Whereas (i) could indeed be derived on the basis of a fronted VP (assuming that C could bear a \([-V_2–]\) feature in this context), VP fronting as such is never an option in wh-clauses in German: An interrogative C is equipped with a feature \([\bullet \text{wh} \bullet]\) (rather than a categorial feature \([\bullet \text{X} \bullet]\), as is the case with declarative root C), and VPs never permit wh-pied piping in German (cf. Heck (2008)), so VP can never qualify as a wh-phrase.\(^{33}\) To name just one obvious question: It seems that structure removal by C is both possible and obligatory only if the head of VP is empty. How can this be derived? In Müller (2015b), I develop a last resort-based account; simplifying a bit, it looks as though C can have \([-V_0–]\) features only if this is the only possibility to accommodate information-structural requirements demanding two separate constituents in the prefield.

\(^{33}\) To name just one obvious question: It seems that structure removal by C is both possible and obligatory only if the head of VP is empty. How can this be derived? In Müller (2015b), I develop a last resort-based account; simplifying a bit, it looks as though C can have \([-V_0–]\) features only if this is the only possibility to accommodate information-structural requirements demanding two separate constituents in the prefield.
binding relations established in the derivation persist throughout the derivation. Second, if an argument remains in the workspace after the phase from which it has been removed (or for which it has been selected without being merged, see footnote 4) is completed, and has not found a binder in that phase (see footnote 11), it is interpreted in the clause in which it does not structurally show up anymore as being bound via default existential quantification. And third, relative scope relations are determined based on final output representations (see discussion of Barss’ generalization). All of this can be derived if the object of semantic interpretation is not a complex syntactic representation at the level of logical form (as in Heim & Kratzer (1998)) but the *derivation tree* that records all operations that have applied throughout the derivation; see Kobele (2015).

Another conceptual question that might be raised is whether it “makes sense” for syntactic derivations to first build structure and then remove it again. Here I would like to argue that asking the question means falling victim to a teleological fallacy: According to standard minimalist assumptions, it is emphatically not the case that Merge exists *so that* syntactic structures can be built. Rather, Merge exists (as a consequence of a sudden, accidental evolutionary step, according to Chomsky’s view), and so it can be used for structure-building. In the same way, I suggest that Remove exists, and can be used for structure removal (which in turn makes it possible to resolve conflicting requirements for syntactic structures).

Next, I would like to point out that that there is a case to be made that an operation like Remove is not only expected in a system based on Merge for reasons of symmetry; operations of this type are in fact already widely assumed to be present as part of the faculty of language, albeit in slightly different form: To wit, *feature deletion* (with uninterpretable features, before transfer to semantics) is widely adopted in minimalist analyses, both as part of Agree operations and in the form of impoverishment operations that are morphologically motivated (with impoverishment qualifying as a postsyntactic operation that is nevertheless very close to core syntax in Arregi & Nevins (2012), and, in fact, as an operation that can also take place within syntax in Keine (2010) and Doliana (2013)). The relevant insight here is that the difference between features or feature bundles on the one hand and heads and phrases on the other hand is a quantitative rather than qualitative one – syntactic categories are composed of nothing but features.34

Finally, let me return again to the main question posed at the outset, viz., what consequences the existence of a Remove operation (with the properties laid out in section 2) has for distinguishing between feature-driven Merge and free Merge. I have tried to develop a simple argument against free Merge, and in support of feature-driven Merge: First, there is evidence for an operation Remove that, by its very nature, does not leave a reflex in the structure to which it has applied (if it did, strict inaccessibility of the item that it has affected could not be ensured). Second, this implies that the legitimacy of an earlier

34 In line with this, it has been shown by Hornstein (2014) that Chomsky’s (2014) recent approach to complementizer-trace effects in English presupposes that the CP shell is *structurally removed*, and not just PF-deleted (also see Chomsky (2015, 24); and, in fact, Chomsky (1981, 303) for an early version of this analysis); this corresponds directly to [–F0–] feature-driven Remove operations as envisaged above.
Merge operation involving the item that undergoes removal cannot be checked by inspecting output representations; but output representations are the only structures that a free Merge approach can access. However, recall from footnote 2 that I have so far presupposed that it is the final output representations that are accessed in a free Merge approach. In principle, the free Merge approach might be compatible with Remove after all if intermediate output representations are accessed. On such a view, the order of operations would have to be (i) Merge, (ii) Check legitimacy of Merge based on output filters, (iii) Remove. To execute this idea, suppose that a phase-based model of free Merge evaluation is adopted; Remove, by definition, would then take place once a phase is otherwise completed (and the legitimacy of Merge operations in the phase has been checked by output filters). Given that CP and vP are phases (and TP and VP are not), such an approach might accommodate the evidence for Remove of DP by v in passive constructions (section 3.1, cf. (33-a)), and for Remove of VP shells by C in complex prefield constructions (section 4.2, cf. (33-b)). However, this would not work for Remove of DP by V in applicative constructions (section 3.2, cf. (33-c)), and it would also fail with Remove of CP (TP, vP) by V in restructuring contexts (section 4.1, cf. (33-d)); in these latter two cases, Remove by V must take place before the phase head is even merged, given the Strict Cycle Condition – and the Strict Cycle Condition also ensures that it cannot be the phase head itself that is responsible for Remove in these contexts.

(33) a. \([vP \text{DP}_\alpha [v'_{-D_2-}] \text{VP}]]\)
b. \([\text{CP VP}_\alpha [c'_{-V_0-}] \text{TP}]]\)
c. \([vP \text{DP}_\alpha [v'_{-D_2-}] \text{PP}]]\)
d. \([vP \text{V}_{-C_0-}] \text{CP}_\alpha ]\)

A possible way out (from a free Merge perspective) might then be to reduce the size of output representations even further, from phases to phrases, so that the legitimacy of the Merge operations introducing XP\(_\alpha\) in (33-a-d) can be checked by output filters in the same phrase in which XP\(_\alpha\) is merged. However, this would seem to come dangerously close to being a notational, but arguably more complex, variant of the feature-driven Merge approach: The central remaining difference would be that the order of (ii) output filter evaluation and (iii) Remove would have to be stipulated (because (ii) is still not strictly part of Merge (i), as it is the case with the feature-driven Merge approach). In addition, there may well be more general issues with radically reducing the domain for output filter evaluation for (internal and external) Merge operations. For instance, the filters that have been proposed in the literature are often quite surface-oriented, and thus do not necessarily lend themselves to evaluation at intermediate stages; minimally, they need a substantial amount of material to work on. Furthermore, it seems clear that the domain used for output evaluation of free Merge operations (XP) would have to be different from the domain standardly taken to be the spell-out domain of phases (or phrases, given (33-cd)), viz., the com-
plement domain of X; otherwise, only complements could be affected by Remove. This consequence may be viewed as conceptually unattractive. For these reasons, I would like to conclude that the main argument of the present paper remains valid even when iterative output evaluation in a phase-based model is adopted: If Remove exists, and if it has the properties I assume it to have, then it provides an argument for feature-driven Merge.
References


Massachusetts, Amherst.


