Structure Removal

Edited by
Andrew Murphy

Institut für Linguistik
Universität Leipzig
2019
Contents

Structure Removal: A Synopsis
Andrew Murphy 1

Structure Removal Counterfeeds Status Government
Johanna Benz 35

The Short Life Cycle of External Arguments in German Passive Derivations
Gereon Müller 55

Long-Distance Passives by Structure Removal
Gereon Müller 113

Tough-Displacement without Movement
Marie-Luise Schwarzer 167

Neutralization in Balinese Subjects by Structure Removal
Imke Driemel & Sören E. Tebay 215

Removing Clausal Determiners in Kwa
Sampson Korsah & Andrew Murphy 247

A Remove-Based Theory of the Complementarity Effect in Breton
Philipp Weisser 295

Implementing Minimalist Syntax and Remove
Johannes Englisch 323

The Fleeting DP in Bulgarian and Macedonian: The View From Left-Branch Extraction
Jelena Stojković 343

Feature Deletion by Head Movement – A New Solution to Agreement Asymmetries in Modern Standard Arabic
Anke Himmelreich 385

The Third Construction and Strength of C: A Gradient Harmonic Grammar Approach
Gereon Müller 419
Structure Removal: A Synopsis

Andrew Murphy

1. Introduction

In current Minimalist approaches to syntax, structures are built bottom-up in an incremental fashion. Consequently, syntactic derivations involving cyclic application of Merge at the root are assumed to involve a monotonic increase in hierarchical structure, with little or no modification of already built structure (cf. No Tampering Condition; Chomsky 2005:13). This all follows from the assumption that the only structure-building tool available to syntax is Merge (Chomsky 1995 et seq.).

That said, there is a long tradition, going back to the earliest transformational work, of assuming that syntactic operations can also remove parts of an existing syntactic representation. This began with early transformational approaches such as S-Pruning in Ross (1967), and has been a recurrent theme ever since. This volume presents a collection of papers exploring the concept of syntactic Structure Removal from varying empirical and methodological perspectives. In this introduction, I will trace the historical development of some core concepts of Structure Removal that have been proposed so far.

2. Tree Pruning (Ross 1967)

In his dissertation, Ross (1967, 1969) proposed a process of Structure Removal, referred to by him as Pruning, that deletes non-branching S nodes in a structure.

(1) S-Pruning (Ross 1967:44):
Delete any embedded node S which does not branch (i.e. which does not immediately dominate at least two nodes).

Ross (1967:44) notes that pruning is not an ordinary transformational rule, but instead a ‘condition upon the well-formedness of trees’. Nevertheless, we will
see that this should be viewed as a grammatical operation, since it interacts with other rules in the grammar.

2.1. Ross’ arguments for Pruning

An important point to bear in mind when understanding Ross’ motivation for S-Pruning is his assumption that adjectives and adnominal modifiers are actually reduced relative clauses. So, the sentence in (2b), for example, would be derived from (2a) by a Relative Clause Reduction Rule that deletes the relative pronoun who and the copula (a process he later called ‘whiz-deletion’; Ross 1972:65).

(2) a. I met a man who was from Boston
   b. I met a man from Boston

Ross provided no fewer than eight arguments in favour of S-Pruning. We will restrict ourselves to three of them in what follows.

2.1.1. Particle movement

The first argument we will consider comes from particle verbs in English. It is well-known that particles associated with verbs can appear before or after the direct object (3).

(3) a. The shock touched off the explosion.
   b. The shock touched the explosion off. (Ross 1967:48)

Ross (1967:47f.) assumes that the sentence in (3b) is derived from (3a) by means of a ‘particle movement’ rule that re-orders the particle and object so that the NP precedes the particle. A simplified version of this rule is given in (4).

(4) Particle movement:
   V PRT NP → V NP PRT
   Condition: blocked if NP is complex (i.e. dominates an S node)

Importantly, this rule has the condition that it is blocked if the NP is complex. Ross (1967:49) assumes that an NP counts as complex if it dominates an S node. This is then supposed to account for the contrasts in (5) and (6).

(5) a. *I am going to call \([\text{NP somebody [S who is strong ]]}\) up.
   b. ?I am going to call \([\text{NP somebody strong } ]\) up.
(6) a. *I polished [NP the vase [S which was from India ]] up.
   b. ?I polished [NP the vase from India ] up.

In each case, the unacceptable (a) examples involve a complex NP that contains a relative clause and therefore cannot undergo the transformation in (4). The more acceptable (b) examples do not count as complex and re-ordering can apply.

However, recall that Ross assumes that the modifiers here are instances of reduced relative clauses. So, example (6b) is derived from (6a), as shown in (7).

(7) Relative Clause Deletion Rule:

Importantly, Relative Clause Deletion in (7) results in a non-branching S node, which is prohibited by (1). Consequently, S-Pruning applies here to delete the S node (8).

(8) S-Pruning:
After S has been removed, the NP no longer counts as complex, and the condition on the Particle Movement in (4) is satisfied. Crucially, if the S node in (8) were not removed by S-Pruning then the derived NP the vase from India would count as ‘complex’ for the rule in (4) and it could not apply. The rule of Relative Clause Deletion feeds subsequent S-Pruning, since it creates a non-branching S node. This then presupposes a certain extrinsic or ‘parochial’ ordering to these rules (Pullum 1979), namely \textit{Relative Clause Deletion} \(\rightarrow\) S-Pruning \(\rightarrow\) Particle Movement.

2.1.2. Latin adjectives

Another example of S-Pruning comes from adjectives in Latin. Here, Ross is concerned with examples such as (9) where the subject and object are modified by postnominal adjectives.

\begin{equation}
\text{(9) } \text{Homō bonus amat fēminam pulchram}
\end{equation}

\begin{equation*}
\text{man good loves woman beautiful}
\end{equation*}

\begin{equation*}
\text{‘The good man loves the beautiful woman.’ (Ross 1967:75)}
\end{equation*}

Again, Ross assumes that these adjectives are reduced relative clauses, derived by the \textit{Relative Clause Deletion Rule} that we saw in the previous section (10).

\begin{equation}
\text{(10) } \text{Relative Clause Deletion Rule:}
\end{equation}

\begin{equation*}
\begin{array}{c}
\text{S} \\
\text{NP} \\
\text{V} \\
\text{Adj} \\
\text{Fēminam} \\
\text{V} \\
\text{Est} \\
\text{Pulchra} \\
\text{Ø} \\
\text{Ø}
\end{array}
\end{equation*}
As in the English examples discussed previously, the resulting structure involves non-branching S nodes that should be subject to S-Pruning (11).

(11) \[
\begin{array}{c}
S \\
\downarrow \\
NP \\
\downarrow \\
N \\
homō \\
\downarrow \\
Adj \\
bonus \\
\downarrow \\
S \\
\downarrow \\
V \\
amat \\
\downarrow \\
N \\
fēminam \\
\downarrow \\
Adj \\
pulchra \\
\end{array}
\]

Ross (1967:76ff.) provides two arguments for S-Pruning here. First, Latin is a relatively free word order language that allows adjectives to be permuted by a rule of Scrambling (12).

(12) *Scrambling in Latin* (Ross 1967:75):

\[
Pulchram₂ \ homō₁ \ amat \ fēminam₂ \ bonus₁ \\
\text{beautiful} \ man \ \text{loves} \ \text{woman} \ \text{good} \\
\text{The good man loves the beautiful woman.}
\]

Importantly, Scrambling must be sensitive to the presence of an S node. It could not, for example, take place out of a relative clause. For this reason, the S node must be removed to allow for the adjective *pulchram* (‘beautiful’) in (12) to be re-ordered.

Another argument comes from case marking. Ross notes that the unreduced relative clause does not allow for the adjective to receive the same case marking as the head noun (13).

(13) *No case marking into a relative clause* (Ross 1967:80):

\[
\begin{array}{c}
homō \ qui \ est \ bonus \ amat \ fēmina-m \ quae \ est \ pulchra(-m) \\
\text{man} \ \text{who} \ \text{is} \ \text{good} \ \text{loves} \ \text{woman-ACC} \ \text{who} \ \text{is} \ \text{beautiful}(-\text{ACC}) \\
\text{The man who is good loves the woman who is beautiful.}
\end{array}
\]

One interpretation of this is that the rule of case marking (or concord) is
sensitive to the presence of an S node. Postnominal adjectives modifying direct objects do so show accusative, such as *pulchram* in (9), meaning that the structure of such an example must differ from that of (13) in no longer having an S node. S-Pruning therefore feeds this rule of Case Marking or concord.

### 2.1.3. Serbo-Croatian clitics

The final of Ross’ arguments to be discussed here comes from second position clitics in Serbo-Croatian (see Rivero 1970 for a parallel argument from Spanish). The crucial observation that he attributes to Wayles Browne (published as Browne 1974) is that clitics in Serbo-Croatian must occupy the ‘second position’ in their clause. Thus, the arguments of the verb *čita* (‘read’) in (14) appear directly after the complementizer of the embedded clause (14).

(14) Ivan želi [š da mi je Ivan čita ]
    Ivan wanted that me it Ivan read
    ‘Ivan wanted Ivan to read to it to me.’

Ross assumes that (14) is derived from the (ungrammatical) underlying structure in (15) by a rule of Clitic Placement that places the clitics in second position of the clause in which they are contained (Ross 1967:90). Importantly, this rule is sensitive to the presence of an S node.

(15) Ivan želi [š da Ivan čita je mi ]
    Ivan wanted that Ivan read it me
    ‘Ivan wanted Ivan to read to it to me.’

It is also possible to have what we could a control structure in which the embedded subject and complementizer are absent (16).

(16) Ivan mi je želi čitati
    Ivan me it wanted read
    ‘Ivan wanted to read it to me.’

Here, it is striking that the second position clitics *mi* and *je* surface before matrix verb *želi*. Ross (1967:88) proposed that control structures were derived
by a rule of *Equi NP Deletion* that deleted the identical embedded subject as well as the complementizer (17).\(^1\)

(17) *Equi NP Deletion*:

Note that, relative to the output structure in (17), the rule of Clitic Placement would not move the clitics since they are already in the second position of their local S clause. In order to derive clitic climbing, the embedded S node must be absent. Like the Relative Clause Deletion rule in the preceding two sections, Equi NP Deletion also creates a non-branching S node. Thus, S-Pruning will apply to (17). Subsequently, Clitic Placement will move the clitics to the second position of the matrix clause, as in (16).

2.2. Perlmutter (1968)

Another example of Pruning can be found in Perlmutter (1968). Perlmutter is concerned with the correlation of the two apparently independent properties of pro-drop and *that*-trace effects. He shows that, in languages like English and French, neither pro-drop (18a) nor *that*-trace violations are possible (18b).

(18) a. *pro have worked all day
    b. *Who, did you say [ that ___ left early ] ?

However, there are other languages such as Spanish and Italian which allow

---

\(^1\)Note that Ross assumed that the embedded clause was dominated by an NP node. I have omitted this here.
pro-drop and subject extraction in the presence of a complementizer. This is shown for Spanish in (19).

(19) a. *pro hemos trabajado todo el día
   have.1PL worked all the day
   ‘We have worked all day.’
   b. Quién dijiste [ que —, salió temprano ]?
   who say.2sg that left early
   ‘Who did you say left early?’ (Perlmutter 1968:207)

In order to account for this, Perlmutter proposes the constraint in (20) that we can refer to as the Null Subject Constraint.²

(20) Null Subject Constraint (Perlmutter 1968:204):
    Any sentence (other than an imperative) in which there is an S that does
    not contain a subject in surface structure is ungrammatical.

This constraint is assumed to be parametrized, it holds in English, but not in Spanish. This can unify both pro-drop and that-trace configurations, since neither of these have overt subjects dominated by an S node. One issue puzzle faced by Perlmutter’s analysis, and indeed all accounts of the that-trace effect, is why the null complementizer obviates the effect (21b).

(21) a. *the woman who he said [ that —, hid the rutabaga ]
   b. the woman who he said [ —, hid the rutabaga ]
      (Perlmutter 1968:214f.)

Perlmutter assumes that complementizer-less embedded clauses are derived by a rule of that-deletion, as in (22).³

---

²This constraint also accounts for the possibility of expletives in languages in which (20) holds (Perlmutter 1968:208f.).

³I have modified Perlmutter’s original examples by using the wh-pronoun. Perlmutter (1968:219) analyzes the relative pronoun that as the relativized subject, however it is unclear whether this is best viewed as a complementizer or a pronoun in relative clauses.
(22) that-deletion:

```
NP
   NP
    the woman
  NP  NP
   who  he
   V
  S
    said
   that
Ø  hid the rutabaga
```

The resulting structure is given in (23). However, the constraint in (20) would still be predicted to rule out this structure as ungrammatical, since the embedded S node does not contain an overt subject.

(23)

```
NP
   NP
    the woman
  NP  NP
   who  he
   V
  S
    said
Ø  hid the rutabaga
```

Perlmutter (1968:219f.) proposes that Ross’ rule of S-Pruning applies here. Note that the rule of that-deletion together with relativization of the subject gives rise to precisely the same configuration as in Ross’ examples, namely a non-branching S node.

2.3. É. Kiss (2008)

É. Kiss (2008) invokes Tree Pruning in a contemporary setting in her analysis of
non-configurationality in Hungarian. Her observation is that V-to-T movement leads to free word order within the VP, as shown in (24).

(24) \textit{Free order in Hungarian VP with V-to-T movement} (É. Kiss 2008:443):

\begin{enumerate}
\item [TP \textit{Össze} \textit{[T' vesztek}, \textit{VP t, a fiúk egymással \textit{]]]}]
\textit{out fell the boys each other with}
\item [TP \textit{Össze} \textit{[T' vesztek}, \textit{VP t, egymással a fiúk \textit{]]]}]
\textit{out fell each other with the boys}
\end{enumerate}

\textit{‘The boys fell out with each other.’}

\begin{enumerate}
\item [TP \textit{Könyvet \textit{[T' vett}, \textit{VP t, Péter Éván-ak \textit{]]]} book.ACC bought Peter Eva-for}]
\item [TP \textit{Könyvet \textit{[T' vett}, \textit{VP t, Éván-ak Péter \textit{]]]} book.ACC bought Eva-for Peter}]
\end{enumerate}

\textit{‘Peter bought some book(s) for Eva.’}

Her proposal is that a pruning-like process, which she calls \textit{domain flattening}, applies to phrases whose heads have been moved (25).

(25) \textit{Domain flattening} (É. Kiss 2008:462):

When the head of a phase is moved into the head position of the next higher phase, the silent copies of the moved head and their projections are pruned.

As represented in (26), this leads to a flat structure within the VP.

\begin{enumerate}
\item (26)
\item [TP \textit{Össze [T' vesztek, VP t, a fiúk egymással \textit{]]]}]
\item [TP \textit{Könyvet [T' vett, VP t, Péter Éván-ak \textit{]]]}]
\item [TP \textit{Könyvet [T' vett, VP t, Éván-ak Péter \textit{]]]}]
\end{enumerate}

\textit{This result of this is that, in the absence of asymmetric c-command between VP-internal arguments, linearization of the VP becomes free. Similar to Müller...}

2.4. Stepanov (2012)

Stepanov (2012) extends É Kiss’ idea to a broader range of cases in which a phrase headed by a trace results in it being ‘pruned’. One such case involves what Stepanov (2012:687) calls the ‘gate opening’ effect of head movement in removing barriers. This generalization can be stated in (27).

(27) Trace-Barrier Generalization (Bošković 2005:35; Bošković 2011:16):

A phrase that is normally a barrier to movement ceases to be a barrier if headed by a trace.

A good example of this comes from Galician. DPs with a free-stranding determiner are to be islands for extraction of PP complements (28a,b). However, if the determiner cliticizes to the verb as in (28c,d), extraction becomes possible.

(28) Incorporation feeds extraction in Galician (Uriagereka 1988:81):

a. *[PP De quén ] viche [DP o retrato tPP ] ?
   of whom saw.2SG DET portrait
   ‘Who have you seen the portrait of?’

b. *[PP De cal ] liche [DP a reseña tPP ] ?
   of which read.2SG DET review
   ‘Which one have you read the review of?’

c. [PP De quén ] viche-loi [DP t₁ retrato tPP ] ?
   of whom saw.2SG-DET portrait
   ‘Who have you seen the portrait of?’

d. [PP De cal ] liche-la [DP t₁ reseña tPP ] ?
   of which read.2SG-DET review
   ‘Which one have you read the review of?’

The intuition here is that DP constitutes a barrier for extraction (29) (see Bošković 2005 for an analysis of how this follows from the PIC and Anti-Locality).
If cliticization of the determiner is analyzed as head movement (30), then the generalization in (27) would mean that the DP ceases to be a barrier and extraction becomes possible.

This can be unified with Kiss’ proposal if we assume that projections whose heads undergo movement are subsequently pruned. Importantly, this must take place in the syntax directly in order to be able to feed subsequent movement processes.
Stepanov (2012:690) argues that pruning as an additional operation is not necessary. Instead, he suggests that head movement does not leave a trace and that this, coupled with Bare Phrase Structure assumptions, causes a phrase whose head has moved to ‘collapse’ (Stepanov 2012:686). It is, however, not entirely clear how this works and what happens to specifiers/adjuncts, etc. Furthermore, it seems to predict that headless phrases cannot move (since they do not exist). While this has been claimed (Funakoshi 2012), it faces a challenge from multiple fronting constructions in German, (for example (see S. Müller 2005; G. Müller 2018b and section 5 for discussion).

2.5. Fitzpatrick (2006)

Another approach to pruning can be found in Fitzpatrick (2006), who discusses auxiliary drop in English (31).

(31) a. **Has** Anyone seen John today?
   b. **Do** You want chicken or beef? (Fitzpatrick 2006:400)

He provides various arguments that these are not inherently reduced clauses and must include a full CP. One of the more straightforward arguments involves the NPI subject in (31a). Since these NPIs are licensed in downward-entailing contexts such as polar questions, the relevant features in C must be present to license it. This can be seen by comparing it to declaratives with raising intonation. In (32), the inflection on the verb shows that we do not have subject-auxiliary inversion here, and consequently the NPI is not licensed.

(32) *Anyone wants a hot dog? (Fitzpatrick 2006:409)

Fitzpatrick argues that subject-auxiliary inversion allows for deletion of the verb by simply not spelling-out the final root CP phase. As (33) shows, there is to T-to-C movement resulting in do-support. When the CP phase is complete, its complement TP is sent to the interfaces for interpretation. The standard assumption is normally that the final CP edge undergoes the final process of Transfer, but Fitzpatrick (2006:419) suggests that this ‘need not apply in all cases, and aux-drop is one case where it fails to apply’. If the root CP edge is not interpreted, then we simply have the TP structure without the moved auxiliary – this gives the impression of deletion of the auxiliary (subject to semantic recoverability).
Fitzpatrick (2006:429) notes that this is, in essence, a Structure Removal analysis, however ‘the small amount of tree pruning at the root is done not through an *ad hoc* truncation operation, but rather finds a more natural place in a theory of cyclic spell-out’. Thus, like Stepanov (2012), he seeks to derive the effect of syntactic deletion without appealing to a dedicated operation for it. That said, some of the details of the implementation are vague. For example, the optionality of suspending the final Transfer operation is not made explicit, nor is its restriction to the root CP. Regarding the latter, there are proposals, such as S-Pruning and ¯S-deletion (to be discussed in the following section), that identify a similar process applying at the embedded CP edge.

3. **S-bar Deletion** (Chomsky 1981)

An early example of Structure Removal is the operation of *¯S-Deletion* discussed in Chomsky (1981). In proposing ¯S-deletion, Chomsky (1981:66f.) was concerned with cases of so-called Exceptional Case-Marking (ECM), where the subject of an embedded clause is case-marked by the matrix verb (34).

(34) I believe him to be smart

At this time, a finite embedded clause was an ¯S constituent, which is analogous to CP in more modern parlance. The C-head and its specifier correspond to a single COMP domain, and the equivalent of TP or IP was S. The structure of the standard embedded finite clause is given in (35).
I believe [S [COMP that] [S he is smart]]

The topmost node of an embedded clause, Š, was assumed to constitute a barrier to government, and concomitantly Case assignment, from the matrix verb. This effectively means that, in the presence of Š, accusative case can not be assigned to the subject of the embedded clause by the matrix verb in (35).

In order for accusative case to be assigned in ECM examples such as (34), Chomsky (1981:66) proposed what he called a ‘marked rue of Š-deletion’ that would be available to ECM-verbs such as believe. This transformational rule would remove the Š node on an embedded clause, and therefore allow for the verb to assign Case to the embedded subject under government (36).

(36) Š-deletion:

\[ \text{I believe [Š [S him to be smart]]} \rightarrow \text{I believe [S him to be smart]} \]

From a modern perspective, it might seem strange to posit a deletion operation for this purpose – why not simply assume that ECM complements are inherently S, rather than Š, constituents? There was good reason for this, however. Assuming ECM complements to be S-constituents was incompatible with the prevailing theory of locality at the time, Subjacency.

To see this, consider the definition of Subjacency in (37).

(37) Subjacency (Chomsky 1977:73):

A phrase cannot move from position Y to position X in

\[ \ldots X \ldots [\alpha \ldots [\beta \ldots Y \ldots ] \ldots ] \ldots X \ldots, \]

where \( \alpha \) and \( \beta \) are cyclic nodes \([= S, NP]\).

This has the effect that a movement step cannot cross more than one cyclic, or ‘bounding’, node, which we will assume are S and NP in our present terminology. This could derive various island effects, but had the effect that long-distance movement had to apply in a successive-cyclic fashion, as is now widely accepted. Movement in ‘one fell swoop’ (38a) would cross two bounding nodes, namely the matrix and embedded S, which is precluded by (37). Instead, movement could stop off, leaving a trace in the intermediate COMP domain (38b). Each

---

4Later in Chomsky (1981), he considers an alternative in which Š is transformed into S (Chomsky 1981:303ff.).
of the movement steps now only crosses a single bounding node, in compliance with Subjacency.

\[(38) \text{ a. } *\text{Who}_1 \text{ do } [\text{S you think } [\text{S [COMP that ] } [\text{S John loves } t_1 ]]] \] ?

\[ \text{b. } \text{Who}_1 \text{ do } [\text{S you think } [\text{S [COMP } t_1 \text{ that ] } [\text{S John loves } t_1 ]]] ? \]

Subjacency also offered an explanation for why English shows wh-island effects. In (39), the subject wh-phrase who blocks movement from stopping at the edge of the embedded clause.\(^5\) As a result, movement must apply in one fell swoop in (39), crossing two S-nodes as in (38a) and therefore violating the Subjacency Condition.

\[(39) *\text{What}_2 \text{ do } [\text{S you think } [\text{S [COMP who}_1 ] [\text{S } t_1 \text{ ate } t_2 ]]] ? \]

With this in mind, let us return to our ECM example. If we did adopt the alternative approach in which ECM complements were simply S, rather \(\hat{S}\)-constituents, then wh-movement from the complement of an ECM verb would violate Subjacency in crossing two S nodes (40).

\[(40) *\text{Who}_1 \text{ do } [\text{S you believe } [\text{S him to have seen } t_1 ]]] ? \]

Thus, such extraction is predicted to be ungrammatical, as noted by Chomsky (1981:172): ‘Rizzi points out that we could not assume that the structures in question are base-generated with S rather than \(\hat{S}\) complements, or it would follow that such sentences as [(41)] would be on a par with wh-island violations.’ Consequently, it seems that we need ECM complements to contain an \(\hat{S}\) layer to facilitate successive-cyclic movement, with later \(\hat{S}\)-deletion removing the barrier to Case government from the matrix verb (41).

\(^5\)Since the COMP domain could, in principle, host multiple elements, the inability for a trace to co-occur with the wh-phrase in COMP would have to be more or less stipulated. This problem is less acute in a CP analysis where the trace and wh-phrase would compete for a single Spec-CP position (see Rudin 1988).
(41) Who do you believe [S [\text{COMP} t] [S him to have seen t₁]]?

Thus, the motivation for Š-deletion as a syntactic operation stems from the conflicting evidence for the presence vs. absence of Š. While the Š node is required for successive-cyclic movement, it must be absent for the purposes of Case assignment to the embedded subject. This seemingly paradoxical situation is resolved by first having the Š layer present in the structure and then later deleting it.⁶

4. **C-deletion (Chomsky 2013)**

The concept of Structure Removal has re-surfaced in recent work by Noam Chomsky. It can be found in published form in Chomsky (2015), as well as in a series lectures given by Chomsky in 2014 at MIT (19th May 2014) and in Olomouc, Czech Republic (*Problems of Projection: Extensions*, 5th June 2014), which were made available online.

I will not recount all the details here, but the relevant proposal of C-deletion pertains to the explanation of *that*-trace effects in the PoP+ (=Problems of Projection) framework (Chomsky 2013). An important background assumption from Chomsky (2013) is that labels on constituents created by Merge are required for the interfaces. Furthermore, T in English is ‘too weak’ to provide a label (42a) (unlike pro-drop languages such as Italian). The consequence here is that a DP must move to Spec-TP in order to provide a label for what would normally be the ‘TP’ constituent.⁷ Simplifying notation somewhat, the $\phi$-features of the subject can provide a label for the $\alpha$ (=traditional TP) constituent (42b).

(42) a. $^*[_{\delta} C [_{\alpha} T [_{\sigma} EA_{\phi} v^* \ldots ]]]$

b. $[_{\delta} C [_{\phi_P} EA_{\phi} T [_{\sigma} v^* \ldots ]]]$

---

⁶ Note that this view entails that Subjacency must be a restriction on movement rules themselves, rather than an output condition as in Freidin (1978) (also see Chomsky 1981:303 for discussion). This is because, in the final output, the intermediate trace will have been removed by Š-deletion, meaning that Subjacency appears to be violated in the output (an instance of counterbleeding opacity).

⁷ This is designed to derive well-known correlation between pro-drop and the EPP, going back to Rizzi (1982).
In prohibited that-trace configurations such as (43), the subject of the embedded clauses moves away, possibly to Spec-CP first for PIC-based reasons.

(43) \[ γ \text{ who do you } v^* [ε \text{ think } [δ C [α T β ]]] \]
\[ = (\text{‘who do you think that read the book’}) \]

(Chomsky 2015:10)

What seems to go wrong here is that the subject is no longer in the canonical subject position (Spec-TP) and thus no label can be provided to the α constituent in (44). The PIC forces the subject to move from Spec-TP, however doing so means that α remains unlabelled (note that copies/traces cannot provide labels). This paradox results in ungrammaticality.

(44)*[δ who C [α T β ]]

However, this analysis faces the familiar problem of why the null complementizer voids this effect. The structure suggested by Chomsky (2015) for these cases is given in (45).

(45) \[ γ \text{ who do you } v^* [ε \text{ think } Ø [α T β ]]] \]
\[ = (\text{‘who do you think read the book’}) \]

(Chomsky 2015:10)

Here, the C head and its associated constituent δ has been ‘eliminated’ by an operation that Chomsky (2015:12) refers to as C-deletion. Concretely, Chomsky (2015:11) proposes the order of operations in (46). First, T inherits the features of C including its phase property (46a). Next, the subject moves to the specifier of T, forming α (46b). α can subsequently be labeled by the EA (46c). Then, the operation of C-deletion deletes the C head and the constituent formed by its merger with α (46d). Finally, Transfer of the lower phase (vP) takes place. However, since the original phase head has ‘disappeared’, the ‘natural assumption is that phasehood is inherited by T’ (Chomsky 2015:11). Importantly, this means that the specifier of T will be accessible for movement in the next phase, thereby avoiding the previous movement paradox we had.

(46) a. Inheritance
b. IM of who in α (EPP)
c. Labeling of α ⟨ϕ, ϕ⟩
d. C → Ø, so that who can remain in situ and still be accessible to IM in the next phase
e. Transfer.
However, one might wonder what kind of operation this ‘C-deletion’ is. As Norbert Hornstein noted on his blog, Chomsky’s approach requires that this is a genuine deletion operation in syntax proper (Hornstein 2014):

This story requires that that is deleted rather than not present at all. Were it never present, C could not transfer its features to T, and T has not features of its own (more below). Thus, to make this work, we need deletion operations in the syntax. A question that arises is how similar the operation deleting that is to more run of the mill ellipsis operations. The latter are generally treated as simply dephoneticization processes. This will not suffice here. […] At any rate, it’s worth observing that C deletion is not simply quieting the phonetics.

This point was also clarified in the Q&A session after a lecture given by Chomsky in 2014 at the Olomouc Linguistics Colloquium. The question posed to him was ‘what kind of deletion operation do we have that can rid of something from the structure? That’s not the sort of thing you can get from Merge.’ Chomsky’s response was printed in the proceedings volume of that conference:

No, that is a deletion operation. It is probably an idiosyncratic operation, which says to take away something. The C has intrinsic properties like Force, like ‘I am a clause’, ‘I am an imperative’, or something, and there has got to be some operation that says ‘lose this property’. I mean it is the kind of operation that we see all the time in phonology. And the question is: can you have an idiosyncratic counterpart to it in syntax? Probably so. I think it is not the only deletion operation, but yes, that is kind of like copy deletion, except in the syntax.

(Chomsky 2014:24)

So, it becomes very clear that what Chomsky has in mind here is something equivalent to Structure Removal in syntax. Again, it is motivated by conflicting structural requirements: The wh-phrase has to be in Spec-TP for the purposes of labeling, but in Spec-CP for PIC reasons. This conflict, which proves fatal for

---

8 From memory, I recall that this question was asked by Thomas McFadden.
that-trace configurations, is alleviated by removing one source of the problem: the CP projection itself.

5. Structure Removal (Müller 2017, et seq)

More recently, a revival of classic syntactic deletion approaches has been pursued by Müller (2017, 2018a,b, to appear) (also see Murphy 2016, Zyman 2018). He argues for an operation parallel to Merge, *Remove*, that triggers deletion of a designated element of a syntactic representation. Müller (2017:28) suggests that Remove mirrors Merge in being feature-driven, able to apply to both heads and phrases, obeying the Strict Cycle Condition, and also having both an internal and external mode of application.

Technically, Remove is assumed to be triggered by a feature [−F−] on a given head that specifies the category of the item to be removed. Müller (2017) also suggests that a diacritic can distinguish between Removal of heads (0) and phrases (2). For example, if a head X bears [−Y−0], it will remove the YP shell of its complement (47a). This is similar to S-Pruning and Ş-deletion that we reviewed in preceding sections. The alternative is that the entire phrase is removed, as in (47b). This seems to be analogous to ellipsis operations that apply to constituents, i.e. VP ellipsis and sluicing (see Murphy & Müller 2019).

(47) a. \[[XP \, X_{[−Y−0]} \, [YP \, Y \, [ZP \, Z \, ... \, ]]] \Rightarrow [XP \, X_{(/Y/)} \, [ZP \, Z \, ... \, ]]\]

b. \[[XP \, X_{[−Y−2]} \, [YP \, Y \, [ZP \, Z \, ... \, ]]] \Rightarrow [XP \, X_{(/Y/)} \, ]\]

The major motivation for Structure Removal comes from evidence for conflicting representations. In other words, a construction has two possible, mutually incompatible analyses, A and B, and there are arguments in favour of each of them. In this situation, Müller (2017) argues that both structures can be accommodated derivationally, by first having A in the derivation and transforming it into B via Structure Removal.

To see a concrete example of this, consider the case of the complex prefield construction in German discussed by Müller (2018b). As a verb-second language, German typically only allows a single constituent to occupy the position proceeding a verb in second position. However, there are examples such as (48) which seem to violate this requirement.
Complex prefield V3 construction (Fanselow 1993:70):

\[
\begin{align*}
\text{DP Einen Brief} & \quad \text{PP nach Hamburg} \\
\text{a letter} & \quad \text{Hamburg}
\end{align*}
\]

‘He has often sent a letter to Anette in Hamburg’

As Müller (2018b) shows, there is evidence for both a single constituency analysis in which the complex prefield forms a single constituent with a VP whose head is silent (49a), and for multiple constituency approach where they fronted elements are independent constituents (49b).

(49) a. Single constituency:

\[
\begin{align*}
\text{CP} & \quad \text{VP XP [V’ e YP ] [C’ C . . . ]}
\end{align*}
\]

b. Multiple constituency:

\[
\begin{align*}
\text{CP} & \quad \text{C’ XP [ C’ YP [ C’ C . . . ] ] ] ]}
\end{align*}
\]

Müller (2018b) shows that there is conflicting evidence in favour of each of these analyses. I will confine myself to briefly summarizing two arguments for each in what follows.

In favour of the single constituency analysis, Müller (2018b:222f.) cites the observation that each of the elements in a complex prefield must come from the same clause (50). If what is fronted is actually a VP with a silent head (49a), then the clausemate restriction follows.

(50) Clausemate condition on complex prefields (Fanselow 1993:66):

a. \text{Ich glaube dem Linguisten nicht [ einen Nobelpreis gewonnen zu haben ]}

\text{I believe the linguist not a ACC nobel.prize won to have}

b. \text{*[CP Dem Linguisten1 einen Nobelpreis2 [C’ glaube ich ____1 the.dat linguist a ACC nobel.prize believe I nicht [ ____2 gewonnen zu haben ]]]}

\text{not won to have}

‘I don’t believe the linguist to have won a nobel prize.’

Furthermore, Müller (2018b) points out that the same ordering restrictions that pertain in the base position hold for the complex prefield. In other words, deviations from default, assumed to be derived by scrambling, are equally
dispreferred in a complex prefield (51b). The preference for base generated argument orders makes sense on the view that what is fronted is a VP constituent.

(51) Ordering restriction in complex prefields (Müller 2018b:224):

a. \[ \text{CP } [\text{DP } \text{den Fahrer }] [\text{PP } \text{zur Dopingkontrolle }] [\text{C'} the.ACC rider to.the.doping.test begleitete ein Chaperon t_{DP} t_{PP}] \]

accompanied a chaperon

b. \[?*[\text{CP } [\text{PP } \text{zur Dopingkontrolle }] [\text{DP } \text{den Fahrer }] [\text{C'} to.the.doping.test the.ACC rider begleitete ein Chaperon t_{DP} t_{PP}] \]

accompanied a chaperon

‘A chaperon accompanied the rider to the doping test.’

There is, however, also evidence that the phrases in a complex prefield do not form a constituent. For example, Müller (2018b) notes an asymmetry in licensing of the NPI auch nur irgendein X by a fronted negative indefinite such as keinen Berg (‘no mountain’). When such a negated expression is fronted as part of a VP, it cannot license an NPI subject (52a). However, as part of a complex prefield, it can (52b). This fact is puzzling if the examples in (52) have the same underlying structure and differ only with regard to the overtness of the verb in the fronted constituent.

(52) NPI licensing from complex prefield (Müller 2018b:236):

a. \[??[\text{CP } [\text{VP } \text{Keinen Berg im Sitzen bewältigt }] [\text{C'} hat auch nur no hill in sitting managed has also only irgendein Fahrer ]]]

some rider

b. \[[\text{CP } [\text{DP } \text{Keinen Berg }] [\text{PP } \text{im Sitzen }] [\text{C'} hat auch nur irgendein no hill in sitting has also only some Fahrer bewältigt ]]]

rider managed

‘No hill was conquered by a rider without getting out of the saddle.’

A similar example comes from the loss of idiomatic interpretation. Fronting of an incomplete idiomatic VP leads to loss of the non-literal interpretation (53a). Heck & Assmann (2014) identify this as part of a more general anti-
reconstruction effect found with remnant movement (i.e. Barss’ Generalization; Barss 1986). In complex prefields, however, idiom reconstruction is possible (53b). Again, this asymmetry is unaccounted for if (53b) also involves fronting of a (remnant) VP.

(53) Idiom reconstruction in complex prefield (Müller 2018b:237f.):

a. #[CP [VP tDP Auf den Kopf getroffen ] [C' hat er [DP den Nagel ] on the head hit has er the nail erneut ]]
   again

b. [CP [DP Den Nagel ] [PP auf den Kopf ] [C' traf er damit ]]
   the nail on the head hit er there.with
   ‘Again, he hit the nail on the head.’

Müller (2018b) argues that conflicting evidence of this kind can be reconciled under an analysis in which we first have single constituency, and later multiple constituency. The analysis first assumes that the VP together with the fronted arguments is moved to Spec-CP to check the [●V●] feature (54a). We can assume that the verb has moved out prior to this to some functional head position. Subsequently, the Remove feature [−V₀−] on the C head is checked by deleting the head of the VP in its specifier, and the associated projection (54b). Finally, the former complement and specifier are ‘reassociated’ into the structure as specifiers of the C head, in both an order preserving fashion and in accordance with the Strict Cycle Condition (54c).

(54) a. VP fronting:

```
    CP
     \   /  
    /   \ 
   VP   C'  
  \   /    \  /  
 XP₁  V'   C    TP
   /   /  \
  YP₂  V   tᵥp  
       /   /  \ 
      e   V₀    T
```

\[ [N] > [−V₀−] \]
b. **Remove:**

\[
\begin{align*}
&\text{CP} \\
&C' \\
&\text{XP}_1 \\
&C \\
&\text{TP} \\
&\text{YP}_2 \\
&[\text{\#N}\text{/}\text{\#N}] > [\text{\#N}\text{/}\text{\#N}] \\
&\text{t}_{\text{VP}} \quad \text{T}
\end{align*}
\]

c. **Reassociation:**

\[
\begin{align*}
&\text{CP} \\
&C' \\
&\text{XP}_1 \\
&\text{C'} \\
&\text{YP}_2 \\
&C \\
&\text{TP} \\
&[\text{\#N}\text{/}\text{\#N}] > [\text{\#N}\text{/}\text{\#N}] \\
&\text{t}_{\text{VP}} \quad \text{T}
\end{align*}
\]

The evidence for single constituency, i.e. the clausemate and ordering restrictions, follow from the fact that the phrases in the complex prefield must be moved as part of a single VP constituent. After Removal and Reassociation as multiple specifiers, the phrases in the complex prefield act as separate constituents, with different c-command relations. Since constraints on NPI licensing and reconstruction access the final output representation, their sensitivity to multiple constituency is accounted for.

Müller’s account provides a solution to the conflicting evidence for single vs. multiple constituency since both representations are available in the derivation, albeit not at the same time. The classic, and arguably only alternative, approach to conflicting representations is the idea that both syntactic structures are present simultaneously, i.e. what is sometimes called *co-analysis* (Haegeman & van Riemsdijk 1986, Di Sciullo & Williams 1987, Saddock 1991, Pesetsky 1995).

To see an example of this, consider *wat voor*-splits in Dutch. In this construction either the full wh-phrase (55a) or just the NP-internal determiner *wat* can be extracted (55b).
wat voor-split in Dutch (Corver 1990, 2017):

a. \[ [\text{NP Wat voor een boeken }]_1 \text{ heb je t}_1 \text{ gelezen ?} \]
   what for a books have you read
b. \[ \text{Wat}_1 \text{ heb je } [\text{NP t}_1 \text{ voor een boeken }] \text{ gelezen ?} \]
   what have you for a books read

‘What kind of books have you read?’

The puzzle here is why this particular construction allows for a certain type of movement (i.e. Left-Branch Extraction) that other NPs do not. A classic approach to this by Bennis (1983) involves parallel representations. Bennis proposes the structure in (56).

(56)

The upper representation is the one in which *wat* is the left-branch of the noun
phrase. The idea is that the lower representation can be generated by a process of restructuring. In this structure, *wat* is reanalyzed as the specifier of *V*. In deriving ordinary wh-movement of whole NP (55a), the upper representation can be accessed, whereas the split construction in (55b) would require the lower representation. This approach therefore accounts for why an apparent violation of the Left-Branch Condition is possible, since *wat* is not NP-internal in the lower representation and therefore freely extractable.

This case of conflicting representations could be equally well captured in a Structure Removal approach in which the two representations are present sequentially, i.e. derived by Remove. Updating our structures somewhat, we could assume that *V* in Dutch can optionally bear a \([-D-]\) feature that can remove the DP shell on its complement (57a). This feature can be checked by removing the DP shell on its complement (57b). Following Müller (2018b), after the former specifier and complement of D, *wat* and *voor een boeken* respectively, are reassociated into the structure in a way that preserves the original c-command relations. Thus, the PP becomes the complement of *V* and *wat* becomes the specifier (57c).

(57) a.
This derives the same result as Bennis' analysis second structure, as *wat* can now be extracted without violating the Left-Branch Condition. We can assume that the feature responsible for removing the DP shell is optional, which would also allow for extraction of the full DP. Thus, Structure Removal provides a strictly derivational alternative to co-analysis approaches with parallel representations.
6. Exfoliation (Pesetsky 2019)

A similar concept of Structure Removal that emerged at the same time as Müller’s is *Exfoliation* (Pesetsky 2019). Exfoliation is a derivational process triggered when an element in a higher clause tries to establish a movement-triggering Agree dependency with another element across a phase boundary. The full definition is given in (58).

(58) *Exfoliation* (Pesetsky 2019:11):
   a. *Structural Description*:
      \[ \ldots \beta \ldots [\gamma P (\text{phase}) \ldots [\alpha P (\text{non-phase}) \ldots \alpha \ldots ]], \]
      where
      (i) \( \gamma P \) is the phase that dominates \( \alpha \) but not \( \beta \),
      (ii) \( \alpha \) occupies the edge of \( \gamma P \), and
      (iii) a movement-triggering probe on \( \beta \) has located \( \alpha \) on its goal.
   b. *Structural Change*:
      Replace \( \gamma P \) with \( \gamma P \), which takes the phasal property of its predecessor.

Pesetsky (2019) focuses in particular on developing Exfoliation analyses of ECM constructions and complementizer-trace effects. For canonical cases of Raising-to-Object (RtO) such as *Sue believes Mary to have solved the problem*, Pesetsky proposes the (here somewhat simplified) analysis in (59).

(59)

```
(59) VP
    V
    [EPP]
    CP
    \ldots
    C'
    C
    TP
    T
    toP
    Mary
    to' to
    \gamma P
```

Exfoliation removes this portion of the embedded clause.
He assumes that all embedded clauses originate as CPs (what he calls the *Full CP Hypothesis*; Pesetsky 2019:9) and that infinitival clauses are derived by means of Exfoliation. In the case of RtO, the matrix V has a movement triggering feature that finds the embedded subject under Agree, however movement is constrained by the PIC. As a response, Exfoliation applies to facilitate movement by deleting the CP and TP projection of the embedded clause (59). This causes the toP projection to become the phase and the abstract to head to be realized overtly. Pesetsky (2019:Ch.4) discusses a number of interesting facts that follow from this analysis.

Pesetsky (2019:36ff.) also proposes Exfoliation of just the CP layer as an explanation for well-known complementizer-trace effects (see section 2.2). Leaving some details aside, the general idea is that movement of the subject to the edge of the Spec-CP phase is ruled out by Anti-Locality or something similar (see Erlewine 2017). Thus, when the probe on matrix v triggering Å-movement finds the embedded subject in Spec-TP, movement cannot apply due to the PIC. As with RtO, the conditions for Exfoliation are met and the CP phase is removed (60). Due to the absence of the CP projection, the impossibility of an overt complementizer is accounted for.

\[(60)\]

Thus, both Pesetsky and Müller's proposals constitute contemporary revivals of...
the old idea of deletion operations in syntax. However, they also differ in some crucial respects (see Pesetsky 2019:12 for discussion). The main one involves the trigger for deletion. While Müller assumes that, like Merge and Agree, Remove is an operation driven by features on heads, Pesetsky’s Exfoliation has more of a Last Resort/repair character, applying when movement would otherwise fail. Furthermore, Müller envisages a much wider application of deletion, for both heads and full projections, as well as in complement and specifier position. In its current form, Exfoliation is primarily triggered by subject movement from an embedded clause.

What both of these approaches have in common, however, is their demonstration of the utility of genuine deletion operations in contemporary syntactic frameworks. Thus, what may seem a rather idiosyncratic, antiquated idea can still provide insightful solutions to long-standing problems such as conflicting representations and variation in clause size. This volume contains a collection of papers that provide new arguments in support of this position.

7. Contributions to this volume

As we have seen, syntactic theories involving some concept of Structure Removal have a long pedigree. However, this idea has manifested itself in different guises. In the present volume, we also see different implementations and conceptions of Structure Removal that shown to have applications in new empirical domains.

The first set of papers discuss feature-driven approaches to Structure Removal in German. The contribution by Johanna Benz explores the phenomenon of ‘status government’ in German. She argues that Structure Removal stands in a counterfeeding relation to status government, accounting for the presence of the second status (zu-infinitive) in restructuring contexts. Gereon Müller contributes two papers developing a Structure Removal analysis of passivization in German. The first, ‘The Short Life Cycle of the External Argument in German Passive Derivations’, argues that there is conflicting evidence for the presence/absence of the external argument in the German passive. This is reconciled by first projecting the external argument and then later removing it from the structure, giving rise to what he calls ‘short life-cycle effects’. The second paper, ‘Long-Distance Passives by Structure Removal’, extends the analysis in the first paper to long-distance passives by additionally adopting a Structure Removal approach to restructuring. Marie-Luise Schwarzer
presents a Structure Removal account of the tough-construction in German. Crucially, she argues that the movement step out of the infinitival clause is not A-movement, as is often assumed, but instead displacement that results from Remove. This avoids the problem of improper movement, for example.

The next group of papers pursue feature-driven Structure Removal approaches from a cross-linguistic perspective. The paper by Imke Driemel and Sören E. Tebay discusses how a Structure Removal approach can derive certain restrictions on subjects in the Balinese object voice construction. Sampson Korsah and Andrew Murphy provide a Structure Removal analysis of clausal determiners in two Kwa languages Akan and Gã that derives their complex distribution. Further cross-linguistic evidence for Structure Removal comes from the paper by Philipp Weisser, who investigates the complementarity effect in Breton, accounting for it with the operation Remove.

The remaining papers all provide different perspectives on Structure Removal. Johannes Englisch presents a computational implementation of a fragment of German grammar that incorporates the core operation of Structure Removal. The paper by Jelena Stojković discusses the intricacies Left-Branch Extraction in Bulgarian and Macedonian (two languages with articles) and provides an analysis based on Pesetsky’s Exfoliation. Furthermore, Anke Himmelreich proposes Structure Removal on the level of features in her account of agreement asymmetries in Modern Standard Arabic. Finally, Gereon Müller provides a unique perspective on Structure Removal as the reduction of syntactic ‘activity’. Couched in the framework of Harmonic Grammar, he argues that Structure Removal can result in Gradient Symbolic Representations, which give rise to weaker violations of grammatical constraints.

References


Abstract
This paper develops an account of German non-finite verbal inflection in terms of (Upward) Agree. I argue for a special treatment of the zu-infinitive, traditionally known as ‘second status’, on the basis of its distribution. Bech (1955) noted that the bare infinitive and the participle have to be licensed by a structurally higher verb within a locality domain, whereas the zu-infinitive does not. Under the present analysis, this follows because the second status is the default value for an inflectional feature otherwise valued by a higher goal, where the default appears if the Agree relationship fails to be established in the relevant locality domain. I also discuss to what extent this morphological dependency informs our understanding of restructuring constructions, which involve the same mechanism to determine verbal inflection. I show that an analysis of restructuring as involving true reanalysis or ‘structure removal’ is compatible with the morphology for optionally/lexically restructuring verbs, but not in most cases of obligatory/functional restructuring.

1. Introduction

Non-finite verbs in German occur in three different morphologically distinct forms. The relevant classification goes back to Bech (1955) and distinguishes the first, second, and third status.

(1) also known as

<table>
<thead>
<tr>
<th>first status</th>
<th>arbeiten</th>
<th>(bare) infinitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>second status</td>
<td>zu arbeiten</td>
<td>zu-infinitive</td>
</tr>
<tr>
<td>third status</td>
<td>gearbeitet</td>
<td>(past) participle</td>
</tr>
</tbody>
</table>

Which of these three forms appear is determined based on the context; specifically, it is determined by a structurally higher element, often the next higher

*Many thanks to Anke Himmelreich, Gereon Müller, Martin Salzmann, and the Morphology/Syntax Colloquium at Universität Leipzig for their very helpful comments! This is a true working paper, and all comments are very welcome, write to me at jbenz@sas.upenn.edu.
verb. Consider the initial examples, each is a finite subordinate clause containing two verbs. In each case, the morphological form of the non-finite verb ’arbeiten’ (to work, in boldface) is dependent on the second verb. In the glosses, I, II and III refer to the first, second, and third status respectively.

(2) a. dass wir Basti arbeiten sahen
   that we Basti work.I saw
‘that we saw Basti work’

b. dass Anna zu arbeiten versprach
   that Anna work.II promised
‘that Anna promised to work’

c. dass Caro gearbeitet hat
   that Caro work.III has
‘that Caro has worked’

This dependency can be thought of as morphological selection: the higher verb selects for a verb with a specific morphological form in its complement. Which form is selected is, broadly speaking, a lexical property of the selecting verb. I argue that these dependencies are established in syntax before they are expressed morphologically: following Adger (2003) and in particular Wurmbrand (2012a,b), this is seen as an implementation of morphological selection rather than an alternative to it.

I propose that in sequences of non-finite verbs in German, first and third status (bare infinitives and participles) are licensed by an Upward Agree operation with the next higher verb. The second status (zu-infinitive) is inserted as a default when this Agree operation fails to take place within the CP. The observation that second status is special goes back to Bech (1955), who notes that first and third status have to be governed by another verb in a locality domain. For Bech, this locality domain is the ’Kohärenzfeld’ (coherence field). Stechow (1984, 1990) shows that this locality domain is in relevant aspects equivalent to the clause. In light of this locality restriction, my goal is to explain why government of first and third status is clause-bounded, while the second status is free to occur without a governing verb as a clause-mate. I propose that this asymmetry follows if German non-finite inflectional morphology arises through Upward Agree for feature values [INFL:I] and [INFL:III] (cf. Wurmbrand 2012a,b), but that second status is inserted as a default if the probe for status on a verb has failed to locate a goal within the (CP-)phase. I relate
this analysis of status government to a recent proposal about restructuring in German, where optionally restructuring control verbs uniformly embed a CP whose CP-projection is later removed (Müller 2017). Optionally restructuring verbs always govern second status, and insertion of second status is contingent on the presence of a CP boundary under the present approach. We learn about the interaction of structure removal in restructuring and status government that structure removal counterfeeds status government. If structure removal were involved in other restructuring contexts such as with modals, verbs of perception and auxiliaries, we would expect second status in all of these constructions, clearly contrary to fact. This is consistent with earlier claims in the literature (especially Wurmbrand 2001 et seq. and other strictly monoclausal approaches to restructuring, but also Müller 2017) that these verbs embed a smaller-than-CP complement from the start. In the absence of a CP, they are then able to govern first and third status. The next section introduces a few additional data points. I present my analysis in section 3. Section 4 discusses the interaction of restructuring and status government and includes a new proposal about raising verbs as involving obligatory restructuring through structure removal for some speakers of German.

2. Some instances of status government

We have seen in the introduction that status government is a process where non-finite verbal inflection is determined by a structurally higher element, typically another verb. This section provides an overview of the different contexts in which status government occurs: in clusters of several verbs, but also with certain complementizers, adjectives and nominals if these categories embed a non-finite verb. First, let us look at status government in periphrastic tense and passive constructions.

\[(3)\]

a. dass der Delfin Zeitung gelesen hat
   that the dolphin newspaper read.III has
   ‘that the dolphin has read the newspaper’

b. dass der Delfin beobachtet wurde
   that the dolphin watch.III was
   ‘that the dolphin was watched’

c. dass jeder Roman gelesen worden sein wird
   that every novel read.III become.III be.I will
   ‘that every novel will have been read’
A comparison of (3a) and (3b) shows that both the perfective auxiliary \textit{haben} ('have') and the passive auxiliary \textit{werden} ('be', 'become') govern the third status (often referred to as the (past) participle). It also becomes apparent that the third status has different morphological realizations. The choice of the correct allomorph depends on verb class and word stress, all in all, there are seven different forms.

I refer to the non-finite verb form that is part of the passive and perfective construction as the ‘third status’ throughout for two reasons. Firstly, because it corresponds to Bech’s original terminology (the term ‘participle’ is reserved for non-verbal forms) and will help us see the crucial contrasts between first, second, and third status. More importantly, it signals the abstract nature of the formal property ‘being third status’ that these forms share, even when they differ both in their selection (belonging, as in (3a) and (3b), to different constructions) and in their surface realization.

We see in (3c) that status government occurs on all non-finite verbs in a complex cluster. For each verb, its status is determined by the next higher verb in the structure.

Verb clusters like this can also include modals, as in (5) (German is different from English in allowing more than one) and other so-called restructuring verbs, as in (6) (more on restructuring below).

(5)  a. dass er \textbf{gescheitert sein} mag
    that he fail.III be.I may
    ‘that he may have failed’

b. dass er \textbf{lesen können} sollte
    that he read.I can.I should
    ‘that he was supposed to be able to read’
In these more complex examples, we again see the inflectional morphology of each non-finite verb determined by the next higher verb. Furthermore, (6c) shows that some infinitival complement clauses can be extraposed, the highest non-finite verb in the cluster appears in second status in both the extraposed and the non-extraposed version of the sentence. In fact, extraposition is only possible for infinitival complements in second status, if we try to do the same with (6a), the result is ungrammatical.

We are starting to uncover an asymmetry between second status on the one hand and first and third status on the other. A further piece of evidence for this asymmetry is that some non-verbal elements such as certain complementizers, adjectives and nominals can also embed non-finite verbs, which always appear in the second status.

Here and elsewhere, this asymmetry is assumed to reflect an important connection between the type of infinitival complement embedded in a given
construction and the possibility to govern status and thereby determine the morphological realization of the non-finite verb in the same construction. This is reflected in the rule of coherence, stated in Bech (1955) and investigated in Stechow (1984, 1999).

(9) Rule of coherence:
   a. Verbs governing the first or the third status are always coherently constructed. (Bech 1955, Stechow 1984)
   b. If a verb actually governs a status, then it is coherently constructed. (Stechow 1990)

In the literature on German syntax, a subordinating construction is ‘coherent’ if it does not embed a sentence, i.e. a CP. Only CPs can be extraposed in German, and only the second status can appear in these extraposed CPs, a generalization the analysis in this paper aims to capture. In the relevant sense, the locality domain we are concerned with can thus be understood to be CP. Government of the first and third status is clause-bound, whereas the second status can appear without a governor in the same clause, as is most obviously the case if it is extraposed. In the next section, we will see a way to implement the notion of status government with the technical machinery of present-day linguistic theory within the Minimalist Program. Specifically, we will characterize status government as a clause-bound Upward Agree relation between verbs.

3. A special status

In line with Wurmbrand (2012a,b, see also Adger 2003, Bjorkman 2011) I analyze morphological selection in the verbal domain as Upward Agree.¹ In her approach to participle licensing, Wurmbrand (2012a,b) posits that participles are verbs which enter the derivation with an unvalued inflectional feature. By Reverse Agree (downward valuation, conceptually closely related to upward probing as in the approach below), the verbs then receive a feature value from a higher head. For example, the participle is valued as [uT:perf] by the

¹A wrinkle: Wurmbrand develops this analysis for participle licensing in Germanic in general and later claims that German actually works differently: in her view, German participles enter the derivation with the feature that will determine their realization as a participle already valued. I will follow the general line of thinking, not this specific point.
higher perfect auxiliary, as illustrated in (10). This feature value is later realized morphologically as the participle.

(10)

I propose a similar analysis that extends to bare infinitives and zu-infinitives, i.e. from the third to the first and second status. The approach below thus differs from Wurmbrand’s in empirical focus. Furthermore, I make different assumptions about the involved feature values. Rather then [perf] or [pass], I will assume entirely abstract feature values [I], [II], and [III]. There are several reasons for this departure from Wurmbrand’s proposal. I propose to follow the Bechian tradition more strictly than previous approaches by treating the alternation under discussion as purely morphological, the features as purely formal and their realizations as non-meaningful (uninterpretable in the literal sense). I represent this choice by naming the feature values [I], [II] and [III], where [I] corresponds to the first status and so on. Somewhat similarly, Sternefeld (2015) characterizes status government with the feature values [INF], [ZU], and [PART]. An important notion in the literature is that status government instantiates the morphological realization of uninterpretable tense/aspect features (eg. [uT:perf] for third status, see above, cf. also Adger 2003). Even if we were to accept the idea that third status should sometimes correspond to ‘being passive”, sometimes to ‘being past” or ‘being perfective”, with the identity in form reduced to accidental syncretism, the bigger problem with this featural representation is that it does not very naturally extend to first and second status, with the result that purely formal feature values like [ZU] for second status or [INF] for first status find their way into this literature as well. It is unclear to me if there is a real difference between uninterpretable feature with and without an interpretable counterpart (Bjorkman 2011:60 raises the same question), but I see no reason to include both types in the same analysis unless this difference can be motivated independently.

The second set of background assumptions concerns German clause structure. In what follows, all verbs that serve as goals or probe for status will be labeled
simply as V. German lexical and 'functional' verbs tend to behave very much alike, contrary to observations for English that have led to the common practice of assuming that auxiliaries and modals are fundamentally different from lexical verbs and merged in fixed functional projections in the clause. Importantly, status government is one of the properties that unifies these classes of verbs, as they all participate in this relation. As an intuitive example, consider the fact that there is no infinitive 'to must' and no participle 'musted' in English, which has been argued to reflect the Merge-site of the modal, namely T, where it can never co-occur with 'to' or receive an inflectional feature of the kind under discussion in this paper (see e.g. Adger 2003). The German modal 'müssen' on the other hand, can occur in all three statuses, and it can be embedded, for example in future tense and perfective aspect and under other modals. In fact, empirically evidence for T in German is scarce more generally. Aside from the missing evidence from distributional effects such as the one described above, there is also no obligatory EPP-movement of the subject to SpecTP in German and it is impossible to see a positional effect of verb movement to T because such movement is string-vacuous due to head-finality in the verbal domain. Considerations like these lead Haider (2010) to claim that German simply does not have a T projection. For mostly theory-internal reasons, I will continue to assume that T is part of the German clause.

One specific assumption about T that I will crucially not adopt here is that second status zu is an instantiation of it (Evers 1990). zu cannot be separated syntactically from the verb and occurs within particle verbs, both facts are easily dealt with if zu expones inflectional features located on or directly above the lexical verb, rather than building zu-infinitives through obligatory verb raising and complex head formation in T. As to the exact location of the inflectional features, Salzmann (2019) provides arguments from displacement phenomena for a separate head F above each verb that is combined with it in the morphology. I take this to be correct but will continue to represent the inflectional features directly on the verbs for simplicity, since no such displacement facts will be considered in this paper.

I propose that verbs probe upward to receive a feature value for inflection.

---

\(^2\)A convincing, but rather indirect piece of empirical evidence can be found in Müller (2017) and concerns unstressed pronoun fronting.
Valued by the next higher verb, these features will be spelled out with the appropriate morphological realization.

The relevant syntactic operation is Upward Agree, defined as follows:

(13)  \textit{Upward Agree} (based on Georgi 2014):
Agree between a probe P and a goal G applies if
\begin{enumerate}
\item G asymmetrically c-commands P
\item P has an unvalued feature $\left[ \ast F: \square \ast \right]$ and G has a matching valued feature $[F]$
\item G is the closest matching goal for P
\item Result: G values P.
\end{enumerate}

I also assume a locality condition on Upward Agree:

(14)  \textit{Locality condition on Upward Agree}
If a probe P fails to locate a valued goal within the phase, a fixed default value is inserted.
The locality condition stops the probe from probing across its own phase head, ie. the head of the phase in which it is contained. Under the assumption that this phase is the CP, the locality condition stops the probe from probing a higher clause domain. Depending on our conception of the phase and upward impenetrability, this may or may not be necessary. Typically and by design, phase impenetrability blocks syntactic operations from outside the phase from interacting with material contained within the phase. It is less clear if it also blocks syntactic operations from within the phase from interacting with material outside it. The questionable period in the derivation consists of the steps between Merge of the phase head and spell-out of the associated phase. There is no reason a priori why a probe should stop to probe upward before it has been sent to the interfaces. I view the locality condition on Upward Agree as a fail-safe to ensure that the probe cannot access material outside the phase even if spell-out is delayed or the phase-head subsequently removed (see next section). Further investigation of the relationship between the phase impenetrability and Upward Agree may obviate the need for such a condition.

Turning to the relevant derivations, in a sequence of verbs, the higher verbs assign inflectional feature values, specifically, they serve as goals and carry the feature values [I] and [III] for first and third status respectively.

(15)
What is different about second status? Recall that its distribution with regard to its syntactic context is different from first and third status. We return to the rule of coherence.

(16) **Rule of coherence:**
   a. Verbs governing the first or the third status are always coherently constructed. (Bech 1955, Stechow 1984)
   b. If a verb actually governs a status, then it is coherently constructed. (Stechow 1990)

We have seen that the second status can appear in incoherent constructions, for example in extraposed CPs. The first clause of the rule of coherence is expected if status government is implemented as a phase-bound Agree operation: only within a coherent construction will it be possible for the Agree relation to be established successfully. Consequently, the second status must have a way of arising that does not involve a successful Agree operation, or successful status government. Indeed, the second clause of the rule of coherence invites the more radical conclusion that the second status never arises through successful probing: in the terminology of Bech and Stechow, the second status is not actually governed at all. Within the present analysis, I propose that [II] is inserted as a default feature value when the verb has probed a completed CP without locating a goal, as required by the locality condition on Upward Agree.
The insertion of this default value happens in syntax and is not merely the realization of an unvalued feature by morphology in the sense of Preminger (2009, 2011), for two reasons: Firstly, there is independent evidence from eg. the IPP (infinitivus pro participio, also Ersatzinfinitiv) that the first status is the morphological default, applied when something goes wrong in morphology (Salzmann 2019), and more importantly, verbs remain in second status even if a goal becomes available at a later stage in the derivation (as in restructuring, which we will examine in the next section), that is, syntactic processes counterbleed the realization of the second status where they might be expected to bleed it if the probe was still unvalued by the time a potential goal becomes available. Under the assumption that the default value is inserted syntactically, it does not only rescue the derivation, but also deactivates the probe permanently.

Before turning to the interactions of restructuring and status government, let us quickly examine the implications of the proposed analysis for the occurrence of non-finite verbs embedded in non-verbal contexts. Recall from section 2 that complementizers, adjectives and nouns, always appear with the non-finite verb in the second status:

(18)  a.  ohne es zu verstehen
    without it understand.II
    ‘without understanding it’

b.  dass der Delfin zu helfen bereit war
    that the dolphin help.II ready was
    ‘that the dolphin was ready to help’

c.  dass mein Plan zu kochen vereitelt wurde
    that my plan cook.II foiled was
    ‘that my plan to cook was foiled’

Under the present analysis, the uniform behavior of these embedding categories follows with the added assumption that non-verbal categories are never lexically specified to govern a specific status, i.e. are not endowed with a status feature that could serve as a goal for a probe on a lower verb. In the absence of such a goal, [II] will always be inserted as soon as a phase-boundary is reached. More generally, the wider distribution of second status follows because it appears in the absence of a successful application of Agree.
4. **Restructuring as structure removal**

I have claimed in the previous section that the second status in sentences with optionally restructuring control verbs such as *versuchen* (‘try’) in (19) depends on the presence of a CP-phase boundary that induces default insertion of a feature value.

(19) dass der Pinguin die Arie zu singen versuchte
    that the penguin the aria sing.I tried
    ‘that the penguin tried to sing the aria’

Whether or not a CP is projected in this construction has been the subject of much debate in the literature; as summarized in Müller (2017, see also references therein), there seems to be evidence both for the presence and the absence of a CP, i.e. for monoclusal and biclausal analyses.

4.1. **Mono- and biclausality in restructuring**

On the one hand, several arguments can be made for the absence of a CP, a monoclausal structure, these arguments include the availability of scrambling and unstressed pronoun fronting, availability of multiple sluicing, and wide scope of embedded negation. As an example for the availability of scrambling out of a restructured infinitival complement, see (20):

(20) dass den Fritz, keiner [ t zu küssen ] versuchte
    that the Fritz no-one kiss.II tried
    ‘that no-one tried to kiss Fritz’

Since scrambling is generally clause-bound in German, (20) should not be possible if the matrix verb embeds a CP. On the other hand, several arguments point to the presence of a CP, a biclausal structure, including the conceptually appealing uniformity of embedding (all of the optionally restructuring control verbs can occur in unambiguously biclausal constructions), licensing of PRO in the embedded clause, the absence of new binding domains, and the availability of extraposition in the so-called Third Construction. To illustrate the relevance of the availability of extraposition in the Third Construction, consider that we just saw that scrambling indicates restructuring, and that we said earlier that only CPs can undergo extraposition in German:
In this example, an infinitival complement has been extraposed, so it must be a CP, and consistent with the analysis in this paper, it appears in the second status. However, the direct object has clearly scrambled out of the embedded clause, indicating a smaller-than-CP structure. Müller concludes that there is convincing evidence for the mono- and biclausal analyses, sometimes within the same construction. For this reason, he proposes a ‘true’ restructuring analysis, in which both types of structures are actually present at different stages in the derivation. He reconciles evidence for monoclausal and biclausal structure in restructuring contexts by innovating the syntactic operation Remove. This operation works to remove structure previously built by Merge. It reconciles conflicting syntactic evidence because it applies after a structure has been built and may have been accessible for operations in need of and in accordance with this bigger structure, which is then shrunk by Remove, yielding a smaller structure that may enable different syntactic processes. Remove is feature-driven, with the features specifying which head or phrase they remove. If a head is removed, its specifier and complement will re-attach within the structure.

In the context of restructuring, this approach means that we have both the biclausal and the monoclausal structure at some stage of the derivation: first, the restructuring verb uniformly takes a CP-Complement. This stage in the derivation accounts for the evidence for a biclausal structure. However, the verb may come with features that require the subsequent removal of the C, T, and possibly v heads. Optionally restructuring verbs can come with or without these features. If the CP- (and TP-)layers are removed, the result is a monoclausal structure. We can easily see that this will have no effect on status government in the embedded clause:

(21) dass sie ihn\textsubscript{2} t\textsubscript{1} versucht [CP\textsubscript{1} t\textsubscript{2} zu küssen ]

that she him\textsuperscript{tries} tries\textsuperscript{kiss.II}

‘that she tries to kiss him’
Even if the CP is subsequently removed and the embedded verb would in principle be accessible for valuation from above once again, it now already has a value and no longer probes:

Once the C head is introduced into the derivation and the CP-layer built, a still unvalued inflectional feature on any of the lower verbal heads will have
to be dealt with and receive a default value. It may well be that the CP-layer will be removed in the very next step of the derivation, but without invoking some kind of look-ahead, the unvalued feature cannot have the information that it will soon be accessible for valuation once again. Therefore, it has to be valued with a default value as soon as the CP is built. In the realm of status government, that default value is second status. It now follows that even if the CP-layer is subsequently removed, the feature is already valued for second status and will no longer probe.

4.2. Obligatory restructuring in raising verbs

The second status morphology governed by optionally restructuring verbs is compatible with a structure removal approach to restructuring, and indeed neatly follows from uniform CP-embedding in this approach, as demonstrated in the previous subsection. I will now discuss the implications of this analysis for a second class of restructuring verbs that governs the second status. This class consists of a small number of raising verbs: \textit{scheinen} (‘to seem’), \textit{pflegen} (‘to usually do,’ ‘to be in a habit of’), \textit{drohen} (‘to threaten’) and \textit{versprechen} (‘to promise’). Müller (2017) does not adopt a structure removal account for these verbs, following Wurmbrand (2001), he assumes that they, along with all other non-optionally restructuring verbs, embed smaller complements from the beginning because they do not display the same evidence for biclausality that we have seen sketched above for optionally restructuring verbs. How then do the verbs in the complement of these raising verbs receive second status? It would be an option to abandon the idea that second status only ever arises as a default value, and then to posit that these verbs simply govern the second status. However, Müller’s structure removal approach actually opens up a different option. Recall that optionally restructuring verbs are optionally endowed with features that trigger subsequent removal of the higher projections of the complement. It is unclear what excludes lexical items that carry these features obligatorily. Obviously such items need not be excluded if they in fact exist, and raising verbs like \textit{scheinen}, \textit{pflegen}, \textit{drohen} and \textit{versprechen} are obvious candidates.

\footnote{All of these verbs have homophonous counterparts with different syntactic properties and a different meaning in other verb classes: \textit{scheinen} and \textit{pflegen} also occur as lexical verbs with respective meanings ‘to shine’ and ‘to take care of’, \textit{drohen} and \textit{versprechen} have subject control counterparts.}
(24) dass der Delfin Zeitung zu lesen scheint
that the dolphin newspaper read.II seems
‘that the dolphin seems to read the newspaper’

Under this analysis, the derivation proceeds in the same way that it did with optionally restructuring verbs, except that *scheinen* carries the removal feature obligatorily.

(25)

Exactly as before, we get the second status on the embedded verb before Remove applies.
For the speakers who allow extraposition in these cases, the analysis of raising verbs as involving an intermediate biclausal stage receives further support. My own judgement, however, aligns with the more traditional view in the literature that this example is ungrammatical. For this second group of speakers, obligatory removal of the C head and the CP projection must be ordered
before extraposition, such that extraposition is bled. Under this analysis of the scheinen-class as involving obligatory restructuring and the derivational presence of a CP-boundary, the stronger clause of the rule of coherence can be maintained: the second status arises exclusively through default valuation and never through the successful application of Agree.

5. Conclusion

In this paper, I have implemented Bech’s (1955) notion of status government as an Upward Agree operation that takes place between verbs in the syntax. Concerning the second status, the zu-infinitive, I have proposed that it arises as a default when the Agree operation fails to apply within the phase. I take this to explain the wider distribution of the second status and its tendency to occur in clausal complements. For occurrences of the second status in constructions that are arguably monoclausal, i.e. in restructuring, we have seen that they remain predicted under an approach that features the intermediate projection of a CP, such an approach is proposed by Müller (2017) for optionally restructuring control verbs and is here extended to the class of obligatorily restructuring raising verbs such as scheinen. In these contexts, the insertion of a default feature value upon completion of the CP permanently blocks the probe - structure removal counterfeeds status government. An open question concerns the containment relationship between the morphological realizations of first and second status, where the first status is contained in the second. It would be in the spirit of the present analysis to take this at face value, given the strong role that the surface forms of non-finite verbs are given within the syntax. Intuitively, it seems appealing to say that zu in the second status is itself a goal in status government and values its complement as [I]. However, this idea is not straightforwardly compatible with the present analysis because this form-internal status government would necessarily be counter-cyclic, given that second status itself does not arise until a fairly late point in the derivation. The alternative is to make the first status a morphological default, realized in the context of the syntactically default zu; the implications of this strategy remain to be investigated.

\[4\]I’m grateful to Andy Murphy for this suggestion.
References


Zeijlstra, Hedde (2012). There is only one way to agree. The Linguistics Review 29. 491–539.
The Short Life Cycle of External Arguments in German Passive Derivations

Gereon Müller

Abstract
In this paper, I pursue two main goals. First, I argue for a new empirical generalization: An external argument in German passive constructions (DP_{ext}) is accessible from positions below it but inaccessible from positions above it. And second, I present a new theory of passivization from which this generalization can be derived: I suggest that an elementary operation Remove should be postulated in phase-based minimalist syntax that is the complete mirror image of Merge in that it triggers structure removal rather than structure building, and that obeys exactly the same restrictions (with respect to triggers, strict cyclicity, etc.). Remove provides a principled approach to conflicting structure assignment and reanalysis in general, with short life cycle effects (derivable from strict cyclicity) at its core.

1. Introduction

In this article, I pursue two main goals. First, I argue for a new empirical generalization: An external argument in German passive constructions is accessible from positions below it but inaccessible from positions above it. And second, I present a new theoretical approach to passivization from which this generalization can be derived: I suggest that an elementary operation Remove should be postulated in phase-based minimalist syntax that is the complete mirror image of Merge in that it triggers structure removal rather than structure building, and that obeys exactly the same restrictions (with respect to triggers, strict cyclicity, etc.). Remove provides a principled approach to conflicting structure assignment and reanalysis in general, with short life cycle effects (derivable from strict cyclicity) at its core.

For helpful comments and discussion, I am grateful to Klaus Abels, Artemis Alexiadou, Karlos Arregi, Katja Barnickel, David Basilico, Hans Bennis, Chris Collins, Norbert Corver, Gisbert Fanselow, Doreen Georgi, Günther Grewendorf, Fabian Heck, Tibor Kiss, Jason Merchant, Andrew Murphy, Dennis Ott, Stefan Müller, Zorica Puškar, Ian Roberts, Uli Sauerland, Florian Schäfer, Barbara Stiebels, Peter Svenonius, Jochen Trommer, and audiences at Universität Frankfurt/Main (Workshop on Generative Linguistics and Philosophy), Universität Stuttgart (Workshop on Exploring Complexity and SFB lecture series), Universität Leipzig (DGfS Conference 37, Workshop ‘What Drives Syntactic Computation?’), and University of Chicago (Comparative Germanic Syntax Workshop 30). Research for this article was supported by a DFG Reinhart Koselleck grant for the project MU 1444/14-1 Structure Removal in Syntax.

Structure Removal, 55–112
Andrew Murphy (ed.)
LINGUISTISCHE ARBEITSBERICHTTE 94, Universität Leipzig 2019
than structure building, and that obeys exactly the same restrictions (with respect to triggers, strict cyclicity, etc.). I proceed as follows. In section 2, I show that there is conflicting evidence concerning the presence of an external argument DP in passive constructions in German, and propose to resolve this conflict by postulating the Accessibility Generalization according to which the external argument is accessible from items that are lower in the structure and inaccessible from items that are higher in the structure. In section 3, I argue that from a conceptual point of view, there is every reason to postulate an operation of structure-removal (\textit{Remove}) as a counterpart to an operation of structure-building (\textit{Merge}). In section 4, I put the two things together and show that an approach to German passive constructions in terms of structure removal accounts for the Accessibility Generalization without further ado, by correctly predicting (via strict cyclicity) a short life cycle of external arguments. After this, section 5 addresses the question of how variation in the area of passivization can be accounted for in the new model. Section 6 draws a conclusion and presents a general outlook. In an appendix, I discuss attempts to maintain strict accessibility or strict inaccessibility in the light of the empirical evidence presented in section 2.

2. Accessibility

2.1. Downward Accessibility

Approaches to passivization differ with respect to the question of whether an external argument DP (DP\textsubscript{ext}) is syntactically accessible or not. Over the last decades, some evidence has been presented that DP\textsubscript{ext} is indeed present in the syntax in passive constructions, and can be accessed by other operations; cf., e.g., Chomsky (1957), Perlmutter & Postal (1983), Roberts (1987), Baker, Johnson & Roberts (1989), Sternefeld (1995), Stechow (1998), Collins (2005), Harley (2013), Merchant (2013), and Georgi (2014b)).\(^1\) A first, well-known argument for the syntactic accessibility of DP\textsubscript{ext} is that DP\textsubscript{ext} can control into adverbal clauses; cf. the case of purpose clauses in (1ab).

\(^1\)For approaches in which DP\textsubscript{ext} is not syntactically represented, see Höhle (1978), Chomsky (1981), Bresnan (1982), Kiss (1992), Wunderlich (1993), Müller, St. (2007) and Kiparsky (2013), among many others; and for approaches where DP\textsubscript{ext} is syntactically represented but not accessible, see Bach (1980), Keenan (1980), Stechow (1987, 1992), Bruening (2013, 2014), Schäfer (2012b), Alexiadou & Doron (2013), Hole (2014), Legate (2014), and Alexiadou, Anagnostopoulou & Schäfer (2015).
The Short Life Cycle of External Arguments

(1) a. Das Schiff wurde \(DP_{ext_1}\) versenkt \([CP \text{ um } PRO_1 \text{ die Versicherung zu kassieren } ]\)
   ‘The ship was sunk in order to collect the insurance.’

b. Der Reifen wurde \(DP_{ext_1}\) aufgepumpt \([CP \text{ um } PRO_1 \text{ die Fahrt fortzusetzen } ]\)
   ‘The tire was inflated in order to continue the journey.’

Second, \(DP_{ext}\) can control into subject-oriented secondary predicates, as in (2abcd).

(2) a. Die Daten wurden \(DP_{ext_1}\) [AP PRO_1 nackt ] analysiert
   ‘The data were analyzed by someone who was naked.’

b. Das Handout wurde \(DP_{ext_1}\) [AP PRO_1 übermüdet ] verfasst
   ‘The handout was written by someone who was tired.’

c. Es wurde [AP PRO_1 absichtlich ] ein Fehler gemacht
   ‘Someone deliberately made a mistake.’

d. Dort wird [AP PRO_1 freiwillig ] gearbeitet
   ‘People work there voluntarily.’

Third, control by \(DP_{ext}\) into a regular complement infinitive is also possible; see (3ab) (with impersonal passives based on a transitive control verb and a ditransitive subject control verb, respectively).²

²As argued by van Urk (2013), examples like (3b) do not violate Visser’s generalization because this generalization should be taken to state that control by an implicit subject in the passive \((DP_{ext})\) into a complement infinitive is impossible if an overt DP agrees with T; this is not the case in the examples in (3), which both exhibit impersonal passives (dative case is not absorbed by the regular passive auxiliary \(werden\), and agreement is only possible with nominative arguments in German). In line with this, an example like (i), where dative case is absorbed by the marked passive auxiliary \(kriegen\) and the remaining overt argument (sie (‘she’)) agrees with T, is correctly predicted to be impossible under Visser’s generalization.
(3)  a. Es wurde $\text{DP}_{ext_1}$ versucht $[\text{CP PRO}_1 \text{ zu schlafen } ]$

   *it was tried to sleep 

   ‘One tried to sleep.’

   b. Gestern wurde $\text{DP}_{ext_1} \ [\text{DP}_2 \text{ ihr }]$ versprochen $[\text{CP PRO}_1/\ast 2$

   *yesterday was her$_{dat}$ promised

   das Zimmer zu reinigen ]

   the room to clean

   ‘Yesterday she was promised by someone that (s)he would clean the room.’

Fourth, reflexive pronouns and reciprocal pronouns in German passive constructions can satisfy Principle A, which suggests the presence of an accessible co-indexed subject DP$_{ext}$.

(4)  a. Hier wurde $\text{DP}_{ext_1}$ sich$_1$ nicht geprügelt

   *here was REFL not hit

   ‘There were no rows here.’

   b. Es wurde $\text{DP}_{ext_1}$ einander$_1$ gedankt

   *it was each other thanked

   ‘People thanked each other.’

Fifth, DP$_{ext}$ cannot easily be interpreted as coreferential with a proper name object, which follows as a Principle C effect if DP$_{ext}$ is syntactically accessible; cf. (5ab).³

³Baker, Johnson & Roberts (1989) analyze this as a strong crossover effect in English; however, since German does not have obligatory fronting to subject position (cf. Grewendorf (1989), Haider (2010), among many others), strong crossover cannot solely be responsible for the illformedness of the sentences in (5).
The Short Life Cycle of External Arguments

Thus, data such as those in (1)-(5) indicate that \( DP_{ext} \) is present in the syntax. Note, however, that all the evidence presented so far concerns material lower in the structure, in domains c-commanded by \( DP_{ext} \). Therefore, we end up with the following generalization.

(6) \textit{Downward Accessibility Generalization:}

The external argument in passive constructions (\( DP_{ext} \)) is accessible for items that it c-commands.

2.2. Upward Inaccessibility

In contrast, a question that does not seem to have been as widely pursued is whether the external argument in passive constructions is also accessible for items higher up in the structure. Closer inspection reveals that this is not the case. Note first that \( DP_{ext} \) can never be bound by a quantified item in a higher clause (see Alexiadou et al. (2015)); this is shown for impersonal passives, where no nominative argument remains after passivization, in (7a).\(^4\) Such binding is unproblematic if the external argument is resumed as part of a \textit{by}-phrase, as in (7b).

(7) a. *Kein Student\textsubscript{1} gibt zu [\( CP \) dass \( DP_{ext1} \) schlecht gearbeitet wurde ]

   no student admits that badly worked was

   ‘No student admits that he did not work well.’

b. Kein Student\textsubscript{1} gibt zu [\( CP \) dass \( DP_{ext1} \) schlecht \[PP von ihm\textsubscript{1} \]

   no student admits that badly by him

   well worked was

   ‘No student admits that he did not work well.’

\(^4\)Here and henceforth, \( DP_{ext} \) in a syntactic representation signals that \( DP_{ext} \) seems to be inaccessible.
In the same way, personal passives disallow binding of \( \text{DP}_{\text{ext}} \) by a quantified item in a higher clause; cf. (8a) vs. (8b).

(8)  

a. *Er hat den meisten Lehrern_1 gesagt [\( \text{CP} \) dass \( \text{DP}_{\text{ext}} \) der he has the most teachers_\text{dat} said that the Maria Bücher geschenkt werden sollen ] Maria_\text{dat} books_\text{nom} given are should ‘He told most teachers that they should give books to Maria.’

b. Er hat den meisten Lehrern_1 gesagt [\( \text{CP} \) dass \( \text{DP}_{\text{ext}} \) der he has the most teachers_\text{dat} said that the Maria [PP von ihnen_1] Bücher geschenkt werden Maria_\text{dat} by themselves books_\text{nom} given are sollen ] should ‘He told most teachers that they should give books to Maria.’

Third, as observed in Stechow & Sternefeld (1988, 447-451), Wunderlich (1989), Stechow (1989), and Haider (2010, 293), control infinitives must have an accessible subject argument. \( \text{DP}_{\text{ext}} \) in passive clauses can evidently not satisfy this condition; this is shown for control into impersonal passives in (9).

(9)  

a. *Er versucht [\( \text{CP} \) \( \text{DP}_{\text{ext}} \) gearbeitet zu werden ] he tries worked to be ‘He tries to ensure that work is being done.’

b. *weil [\( \text{CP} \) bald \( \text{DP}_{\text{ext}} \) geschlafen zu werden ] gewünscht wird because soon slept to be wished is ‘because someone wishes that sleep comes over people.’

Fourth, non-overt material can in principle satisfy criterial movement constraints in German. This is standardly assumed for extraction from embedded verb-second clauses as in (10a) (where the intermediate movement step to SpecC can satisfy a verb-second C projection’s specifier requirement), and it also holds for topic drop constructions as in (10b).

(10)  

a. Wer_1 glaubst du [\( \text{CP} \) – hat Recht ]? who_\text{nom} think you is right ‘Who do you think is right?’
b. – habe ich schon gesehen heute
   (her) have I already seen today
   ‘I have already seen her today.’

However, DP_{ext} can never satisfy a criterial movement constraint in passive constructions; cf. (11a) vs. (11bc) (with a by-phrase PP and an expletive in SpecC, respectively).

(11) a. *Ich denke \[CP \overline{DP_{ext}} \text{ ist gut gearbeitet worden }\]
     I think is well worked been
     ‘I think that people worked well.’

b. Ich denke \[CP [PP von ihr] \text{ ist gut gearbeitet worden }\]
   I think by her is well worked been
   ‘I think that she worked well.’

c. Ich denke \[CP es ist \overline{DP_{ext}} \text{ gut gearbeitet worden }\]
   I think it is well worked been
   ‘I think that people worked well.’

The fifth observation concerns the absence of minimality effects in passive constructions. As noted by Collins (2005), if DP_{ext} is structurally represented in passive constructions, it is a priori unclear why movement of the internal argument to subject position can take place, given that movement obeys minimality: DP_{ext} in SpecV is invariably closer to SpecT than DP_{int} (i.e., an internal argument DP) in VP. The relevant derivation is given in (12).

(12) \[TP \left[ T' \left[ VP \overline{DP_{ext}} \left[ v' v \left[ VP V \text{ DP_{int} } \right] \right] \right] \right]\]

A derivation as in (12) should be expected to be ill formed. However, DP_{int} moves to SpecT in English passive constructions, which suggests that DP_{ext} is in fact not syntactically accessible and can thus not intervene; see (13).

(13) \[TP \text{ John}_{2} \text{ was } \left[ VP \overline{DP_{ext},I} \left[ v' v \left[ VP \text{ killed } t_{2} \right] \right] \right]\]

DP_{int} can also move to SpecT in German passive constructions, which then also indicates that that DP_{ext} is not syntactically accessible. However, German differs from English in that movement to subject position in general is optional (see footnote 3), and that much TP-internal word order variation can be traced back to scrambling. A test for optional movement to SpecT in German is
devised in Müller (2001): First, only a nominative subject argument DP can precede unstressed pronouns and at the same time follow C elements (cf. (14a) vs. (14b)); object DPs cannot do so (cf. (14cd)). This strongly suggests a designated position in which only a nominative DP can show up: SpecT. On this view, unstressed pronouns move to a domain that precedes the landing sites for scrambling (viz., specifiers of vP or VP) but crucially follows SpecT. This rules out both (14c) (where DP_{dat} is scrambled to a position preceding the unstressed pronoun) and (14d) (where DP_{dat} is moved to a domain that cannot be reached by scrambling, viz. TP).

(14) a. dass es_{3} [vP der Fritz_{1} dem Karl_{3} t_{2} gegeben ] hat that it_{acc} the Fritz_{nom} the Karl_{dat} given has 'that Fritz gave it to Karl.'

b. dass der Fritz_{1} es_{3} [vP t_{1} dem Karl_{3} t_{2} gegeben ] hat that the Fritz_{nom} it_{acc} the Karl_{dat} given has 'that Fritz gave it to Karl.'

c. *dass der Fritz_{1} dem Karl_{3} es_{2} [vP t_{1} t_{3} t_{2} gegeben ] hat that the Fritz_{nom} the Karl_{dat} it_{acc} given has 'that Fritz gave it to Karl.'

d. *dass dem Karl_{3} der Fritz_{1} es_{2} [vP t_{1} t_{3} t_{2} gegeben ] hat that the Karl_{dat} the Fritz_{nom} it_{acc} given has 'that Fritz gave it to Karl.'

Given this state of affairs, it is clear that optional movement of a nominative DP_{int} in German passive constructions has taken place in (15a) (but not in (15b)); and such movement is evidently not blocked by an intervening DP_{ext}.

---

5Note that this reasoning requires two further assumptions. First, movement operations like scrambling and unstressed pronoun fronting have some way to circumvent minimality effects, unlike classical A-movement to SpecT; this has long been known and holds true of these operations almost by definition. One possible explanation for this is that the landing sites of scrambling and pronoun fronting show evidence of being A-bar rather than A-positions (see Müller (1995)). Second, A-movement of DP_{ext} in (14b) or of DP_{int} in (15a) to SpecT across a fronted pronoun (or a scrambled DP) also can avoid minimality effects, again in contrast to A-movement that would cross a non-overt DP_{ext} in its base position. Again, it would seem reasonable to account for this by assuming that A-bar specifiers (scrambled DPs, fronted pronouns) can never induce minimality effects with A-movement, in contrast to items in their base positions.
A sixth observation is that $\text{DP}_{\text{ext}}$ does not block anaphoric binding from above in passive constructions, in contrast to other external arguments in German, which act as interveners (see Pitteroff (2014)). This is shown by the active/passive pair in an AcI construction with lassen (‘let’) in (16). In (16a), the subject DP *die Diener* (‘the servants’) blocks a satisfaction of Principle A by *sich* with the matrix subject DP as antecedent; in contrast, in the lassen-passive construction in (16b), $\text{DP}_{\text{ext}}$ does not preclude *sich* from satisfying Principle A with the matrix subject DP as antecedent; in this derivation, the reflexive bears index 1.6 Alternatively, the reflexive can also take the embedded subject as an antecedent, signalled here by index 2. Thus, the two possible readings of (16b) show both inaccessibility of $\text{DP}_{\text{ext}}$ (in the case of index 1) and accessibility of $\text{DP}_{\text{ext}}$ (in the case of index 2, based on the same reasoning as in (4)) in one example.7

---

6Note that German AcI constructions sometimes permit long-distance reflexivization, but this effect only shows up with PPs; cf. Reis (1976), Grewendorf (1983), Gunkel (2003), Barnickel (2014). Also, binding by the matrix subject in (16b) cannot be due to raising of the reflexive pronoun *sich* to the matrix clause because *sich* can participate in VP topicalization. Finally, lassen-passives are special insofar as passivization is not accompanied by any morphological reflex; however, the status of the construction as a regular passive is uncontroversial (among other things, a by-phrase is possible, and lexical restrictions on passivization are identical to those active in standard passives).

7As a matter of fact, there is a third reading of the string in (16b), irrelevant in the present context, where lassen does not have a causative or permissive interpretation (as presupposed so far) and *sich* is not an object of the embedded verb *rasieren*; rather, *sich lassen* acts as a modal passive auxiliary. On this reading, (16b) would mean ‘It is possible to shave the king.’ See Höhle (1978) for a comprehensive description of this construction.
(16) a. Der König lässt [act die Diener sich_{1/2} rasieren ]
   the king lets the servants shave
   'The king lets the servants shave themselves.'

   b. Der König lässt [pass DP_{ext2} sich_{1/2} rasieren ]
   the king lets someone shave
   'The king lets people shave themselves.'

Taken together, we end up with the generalization in (17).

(17) **Upward Accessibility Generalization:**
   The external argument in passive constructions (DP_{ext}) is not accessible for items that it does not c-command.

Combining the two generalizations in (6) and (17), the Accessibility Generalization in (18) emerges.

(18) **Accessibility Generalization** (preliminary version):
   DP_{ext} in passive constructions is accessible for items that it c-commands and inaccessible for items that it does not c-command.

While I take (18) to be descriptively correct for the most part, the following subsection shows that there is a further empirical phenomenon that will lead to a slight modification.

2.3. A Refinement: Quantificational Variability Effects

Consider quantificational variability effects in English, as they arise with indefinites (see Heim (1982), Diesing (1992)) and embedded wh-clauses (see Berman (1991)); cf. (19ab), respectively.

(19) a. A cat is usually smart ≅ Most cats are smart
    b. John partly remembers who cheated

A standard assumption that I will adopt here (though see Hinterwimmer (2005) for a qualification) is that an indefinite DP (which can also be a wh-phrase) denotes an open sentence with a free individual variable that can be unselectively bound by an adverb of quantification; in line with this, existential binding in general comes about as a default operation. As observed in Alexiadou & Müller
(2015), the external argument in German passive constructions is also subject to quantificational variability effects. This is shown by the data in (20ab), with the adverbs of quantification gröstenteils (‘for the most part’) and zum Teil (‘partly’), and the interpretations as specified by the free translations.

\[(20)\]
\begin{align*}
\text{a. } & \text{Es wurde gröstenteils } \text{DP}_\text{ext} \text{ geschlafen beim Vortrag} \\
& \text{it was for the most part slept at the talk} \\
& \text{‘Most people slept during the talk.’}
\end{align*}
\begin{align*}
\text{b. } & \text{ Dann wurde der Sprecher zum Teil } \text{DP}_\text{ext} \text{ ausgebuhlt} \\
& \text{then was the speaker partly booed} \\
& \text{‘Then a proper subset of people booed the speaker.’}
\end{align*}

For binding by the adverb of quantification to be possible in (20ab), \text{DP}_\text{ext} needs to be c-commanded by it, and thus must be accessible for some item outside its c-command domain after all. This requires a modification of the Accessibility Generalization. As it turns out, there is evidence that the adverbs of quantification in question are properly included in the vP; thus, they can participate in vP topicalization, unlike, e.g., sentence adverbials which are merged outside of vP; see (21a) vs. (21b).

\[(21)\]
\begin{align*}
\text{a. } & \text{Gröstenteils } \text{DP}_\text{ext} \text{ geschlafen } \text{wurde beim Vortrag} \\
& \text{for the most part slept was at the talk} \\
& \text{‘Most people slept during the talk.’}
\end{align*}
\begin{align*}
\text{b. } & \text{*[ Wahrscheinlich } \text{DP}_\text{ext} \text{ geschlafen } \text{wurde beim Vortrag} \\
& \text{probably slept was at the talk} \\
& \text{‘People probably slept during the talk.’}
\end{align*}

Note also that \text{DP}_\text{ext} cannot be bound by an adverb of quantification outside the minimal clause; see (22ab) (where only a reading is available where the adverb quantifies over time spans).\footnote{Things are different with overt indefinites, which can be non-locally bound by an adverb of quantification; cf. Heim (1982).}

\[(22)\]
\begin{align*}
\text{a. } & \text{Es war gröstenteils } \text{so dass DP}_\text{ext} \text{ geschlafen wurde} \\
& \text{it was for the most part so that slept was} \\
& \text{beim Vortrag ]} \\
& \text{at the talk} \\
& \text{‘*Most people slept during the talk./’For most of the time people slept during the talk.’}
\end{align*}
b. Es war zum Teil so [CP dass der Sprecher $\text{DP}_{\text{ext}}$ ausgebuh | it was partly so that the speaker booed | wurde] | was  |

*A proper subset of people booed the speaker.* '/At some points the speaker was booed.'

The fact that $\text{DP}_{\text{ext}}$ in German passive constructions is subject to quantificational variability effects can then be accommodated by a minimal refinement of (18), with the concept of c-command replaced by the slightly more liberal notion of m-command.

(23) \textit{Accessibility Generalization} (final version):

$\text{DP}_{\text{ext}}$ in passive constructions is accessible for items that it m-commands and inaccessible for items that it does not m-command.

If the Accessibility Generalization in (23) is correct, it implies that it is not possible to maintain either strict syntactic accessibility or strict syntactic inaccessibility of $\text{DP}_{\text{ext}}$ in German passive constructions. For the time being, I will abstract away from possible attempts to nevertheless maintain existing approaches where $\text{DP}_{\text{ext}}$ either is always, or is never, accessible in the syntax, in view of the empirical evidence presented so far; I will address this issue in the Appendix (and the conclusion will be negative). Beyond that, the Accessibility Generalization in (23) raises the question why it should hold. While there are several ways in which (23) might in principle be derived, the null hypothesis would clearly seem to be that accessibility is simply correlated with existence: Where $\text{DP}_{\text{ext}}$ is accessible in passive clauses in German, it is structurally present; and where it is not accessible, it is gone. The approach I would like to pursue in what follows is designed to reflect this hypothesis.

3. \textbf{Structure Removal}

3.1. The Operation Remove

At this point, the question arises of whether there are existing approaches in current syntactic theory that can reconcile conflicting syntactic representations resulting from evidence both for and against the presence of some item in the structure. As far as I can see, there is only one; it relies on the concept of
multidimensional representations or coanalysis (see Huybregts (1982), Bennis (1983), Haegeman & Riemsdijk (1986), Di Sciullo & Williams (1987), Sadock (1991), and Pesetsky (1995)). On this view, the string of an example like (16b) on the matrix vP level (i.e., before verb-second movement and topicalization) could be taken to be associated with two structures simultaneously, viz., one that includes DP_{ext} and one that does not; see (24).

(24)

Operations that require DP_{ext} to be present (like Principle A satisfaction with the external argument of the embedded verb, which presupposes a c-commanding co-indexed DP) would then access the upper structure, whereas operations that require DP_{ext} to be absent (like Principle A satisfaction with
the external argument of the matrix verb, which must be able to circumvent an intervention effect triggered by an embedded DP\textsubscript{\textit{ext}} would access the lower structure. However, there are severe problems with such a coanalysis approach. First, coanalysis is known to be extremely powerful and insufficiently restricted; in line with this, it is unclear how it could be integrated into a theory of incremental structure-building based on Merge operations. Second, and even more importantly, the coanalysis approach cannot derive the Accessibility Generalization in (23): There is nothing that would determine which processes access which of the two co-existing structures. Therefore, the empirical evidence presented in section 2 could only be accounted for by a number of independent stipulations, with the clear pattern determined by an m-command threshold resulting accidentally.

Against this background, I would like to advance a new, more principled approach to conflicting representations from which the Accessibility Generalization follows in a very direct way. The central claim is that syntactic derivations employ two elementary operations modifying the size of representations: In addition to an operation that \textit{builds} structure – \textit{Merge} (Chomsky (2001, 2008, 2013)) –, there is a complementary operation that \textit{removes} structure: \textit{Remove}.

If \textit{Remove} exists as the mirror image of \textit{Merge}, it is expected to show similar properties and obey identical constraints. I will make the following assumptions about the nature of \textit{Merge}. First, \textit{Merge} is feature-driven. It is triggered by designated [\textbullet{\textit{F}}\textbullet{}] features, which are ordered on lexical items (see Pesetsky & Torrego (2006), Heck & Müller (2007), Abels (2012), Georgi (2014\texttextsubscript{a}), Müller (2014), Stabler (2013), Collins & Stabler (2016), and references cited in these works). Once a feature has triggered an operation, it is discharged (and thereby deleted), and the next feature on the list becomes active. Second, \textit{Merge} may apply to heads or phrases. The difference between heads and phrases needs to be formally expressed in some way; this can be done by attaching appropriate diacritics 0 (which stands for minimal projection) and 2 (which stands for maximal projection, but not necessarily for exactly two X-bar levels in a phrase) to the features triggering \textit{Merge}: [\textbullet{\textit{F}}\textsubscript{0}\textbullet{}], [\textbullet{\textit{F}}\textsubscript{2}\textbullet{}]. Third, \textit{Merge} may be external (by taking some item from the workspace) or internal (by taking some item from the current phrase marker, thereby producing movement). Finally, fourth, \textit{Merge} obeys the Strict Cycle Condition, a core principle of derivational grammar. The version of the Strict Cycle Condition adopted here is given in (25) (see Chomsky (1973, 1995, 2001, 2008)).
The Short Life Cycle of External Arguments

(25) **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 $\alpha$.

The notion of *domain* in (25) is understood as in Chomsky (1995): The domain of a head $X$ is the set of nodes dominated by XP that are distinct from and do not contain $X$. The Strict Cycle Condition ensures that properly embedded structures cannot be exclusively manipulated by syntactic operations; for Merge the restriction is derived that the operation can take place only with a member of the current root projection. Still, it can be noted that (25) is slightly less strict than Chomsky’s (1995) Extension Condition, in that it permits both Merge of a head $Y$ to another head $X$ (as in head movement) and Merge of a phrase $YP$ to (a non-maximal projection of) $X$ (i.e., tucking in in the sense of Richards (2001)), as long as $X$ is the head of the current root.

I assume that Remove obeys identical restrictions. Thus, first, Remove is feature-driven. It is triggered by designated [–F–] features, which are ordered on lexical items; [–F–] features for structure removal are interspersed with [●F●] features for structure building on a head. Second, Remove may apply to heads or phrases; again, this is signalled by a diacritic that accompanies the feature triggering the operation: [–F$_0$–], [–F$_2$–]. Third, Remove can be external or internal. However, the cases I will focus on in what follows involve internal Remove, i.e., removal of items that are part of the syntactic structure that Remove applies to.$^9$ Fourth, Remove obeys the Strict Cycle Condition. By its very nature, it is impossible for a Remove operation to extend the phrase marker created so far; however, the Strict Cycle Condition in (25) ensures that Remove can only apply to heads or phrases in the domain of the head that bears the [–F–] feature, and not to more deeply embedded items.$^{10}$

---

$^9$External Remove affects material that is not actually present in syntactic structure. See Müller (2015) on how this paradox can be resolved. I will briefly come back to external Remove in the appendix.

$^{10}$Note that this presupposes that both with external Merge and with Remove, an operation that is triggered by the head of a projection $\alpha$ and that applies to some item $\delta$ (merging or removing it) does indeed ‘exclusively target’ $\delta$ (in the sense of (25)) in the domain of which $\delta$ is a member. The case is different with internal Merge (i.e., movement), where the operation targets both the embedded domain (the pre-movement position of $\delta$) and the domain of the head triggering the operation (the post-movement position of $\delta$).
In what follows, I present abstract scenarios instantiating Remove, first with phrases, and then with heads.

3.2. Remove and Phrases

Remove applying to phrases is triggered by \([-F_2-]\) features. In (26), a head \(X\) is taken from the workspace (more precisely, from the numeration that is part of the workspace) that is equipped with a feature \([•Y_2•]\) triggering Merge of some \(YP\), and with a feature \([-Y_2-]\) triggering Remove for some \(YP\). The former feature is higher-ranked (indicated by \(\succ\)) and needs to be applied (and thereby discharged) first. The Merge operation is shown in (26a), and the subsequent Remove operation in (26b) (with the target for Remove indicated by a box around it, here and in the following derivations).

(26) Remove and phrases: complements

a. \(\text{Merge}(X_{[•Y_2•]>[-Y_2-]}, YP)\):  
   \[
   \begin{array}{c}
   X' \\
   X_{[-Y_2-]} \\
   \text{YP} \\
   \text{ZP} \\
   Y' \\
   Y \\
   \text{WP} \\
   \end{array}
   \]

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

Three remarks are in order here. First, in (26a), the structural configuration for Remove applying to either \(ZP\) or \(WP\) is not present, because of the the Strict Cycle Condition. Thus, if \(X\) were equipped with a \([-Z_2-]\) or \([-W_2-]\) feature, the derivation would crash: Only \(YP\) can be successfully removed, yielding (26b). Second, the order of features for Merge and Remove is crucial. A head \(X_{[-Y_2-]>[•Y_2•]}\) could not produce the derivation in (26).\(^{11}\) And third, (26)

\(^{11}\) Depending on further assumptions which are not relevant to the issues under consideration in this article, the output based on \(X_{[-Y_2-]>[•Y_2•]}\) would either be external removal of some phrase of type \(Y\) from the workspace (cf. footnote 9) followed by Merge of some (other) phrase of type \(Y\); or a crash of the derivation. Note that I do not assume that a frustrated \([-F-]\) feature
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 – as we will see below, the intermediate representation can have an influence on the applicability of other processes before it is undone again.

Consider next a scenario where some head selects its specifier for a Remove operation, as in (27). First, X merges with YP in (27a); subsequently, YP is removed again in (27b). As before, it is possible that some other operation intervenes between the two processes, and for such an operation the temporary presence of YP will make a difference, even if YP is not part of the final output representation after having undergone removal.

(27)  

Remove and phrases: specifiers

\[
\begin{align*}
\text{a.} & \quad \text{Merge}(X'[\bullet Y_2\bullet]_{-Y_2-}, YP): \\
& \quad \begin{array}{c}
\text{XP} \\
\text{YP} \\
\text{ZP} \\
\text{Y'} \\
\text{Y} \\
\text{WP} \\
\text{UP}
\end{array} \\
\text{b.} & \quad \text{Remove}(X'_{-Y_2-}, YP): \\
& \quad \begin{array}{c}
\text{XP} \\
\text{X} \\
\text{UP}
\end{array}
\end{align*}
\]

As with complement removal as in (26), phrases that are more deeply embedded in YP cannot be removed by X because of the Strict Cycle Condition. Thus, even if X in (27) were to be equipped with the appropriate categorial features for Remove (like [-Z_2-] or [-W_2-]), ZP and WP could not be removed in (27). However, in principle (i.e., if it were equipped with an appropriate categorial

---

that cannot trigger structure removal can simply undergo deletion, in contrast to what has been argued for probe features (see footnote 15 below).
feature \([-U_2-]\), X might also remove UP in this configuration after YP has been merged (in analogy to tucking in scenarios with Merge); see Murphy & Müller (2016) (and literature cited there) on empirical motivation for such an operation (based on sluicing constructions).

Remove applying to a phrase is what will be relevant for the analysis of the variable accessibility of external arguments in German passive derivations. Still, for the sake of completeness, let me also briefly consider the case of Remove applying to a head.

3.3. Remove and Heads

If Remove applies to a head rather than a phrase, this is due to the presence of \([-F_0-]\) rather than \([-F_2-]\) on a head. (28) illustrates a case where the head Y of a complement YP is removed.

(28) \textit{Remove and heads: complements}

\begin{itemize}
  \item a. \text{Merge}(X_{[\bullet Y_2 \bullet]} > [-Y_0-], YP):
  \begin{center}
  \begin{tikzpicture}
  \node (X) at (0,0) {$X'$};
  \node (X_y) at (-1,-1) {$X_{[-Y_0-]}$};
  \node (yp) at (-1,-2) {YP};
  \node (y) at (0,-2) {$Y$};
  \node (zp) at (2,-2) {ZP};
  \node (x) at (1,-1) {X};
  \node (z) at (2,-1) {ZP};
  \draw (X) -- (x);
  \draw (x) -- (y);
  \draw (x) -- (z);
  \end{tikzpicture}
  \end{center}

  \item b. \text{Remove}(X_{[-Y_0-]}, Y):
  \begin{center}
  \begin{tikzpicture}
  \node (X) at (0,0) {$X'$};
  \node (x) at (1,-1) {X};
  \node (zp) at (2,-1) {ZP};
  \end{tikzpicture}
  \end{center}
\end{itemize}

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 (28)) 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, and that is temporarily split off from the current tree after removal of the head and its 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 process works like Tree Pruning; see Ross (1967, ch. 3).\textsuperscript{12}

\textsuperscript{12}Also, cf. Stepanov’s (2012) approach to head movement and, in particular, Pesetsky’s (2016) Exfoliation operation for embedded clauses for related concepts.
In the same way, Remove applying to heads can also affect a specifier. The operation is shown in (29), where X has first merged with a UP complement; again, an XP included in the specifier (here: ZP) cannot be targeted by the operation, due to the Strict Cycle Condition. ZP reassociates with the X projection as a specifier, in a maximally order-preserving way.

(29) Remove and heads: specifiers

a. Merge($X'_{[\bullet Y_2 \bullet]} > [-Y_0-], YP$):

```
XP
  / \       /  \\
YP   X'   ZP   X'[-Y_0-]
```

b. Remove($X'_{[-Y_0-]}, Y$):

```
XP
  / \       /  \\
ZP   X'   UP
```

In the cases discussed so far, the head Y that is subject to Remove takes a complement but is not accompanied by a specifier. If there are two or more items in YP (e.g., ZP and WP, as in (26) and (27)), the null hypothesis is that they reassemble in their original linear and hierarchical order in the XP domain, preserving the original c-command relations of items in the YP domain, so that structural changes induced by the operation are minimized; see Müller (2018) for a much more comprehensive discussion of Remove applying to heads, and evidence for order preservation with reassociation.

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 an order-preserving way.

3.4. Short Life Cycle Effects

The Remove-based approach to conflicting structure assignments in syntax has been applied to a number of recalcitrant phenomena exhibiting evidence for
conflicting representations. A characteristic property of Remove operations is that removed material is expected to what can be called short life cycle effects; i.e., once some item is merged that is subject to removal, it can only survive in the structure for as long as it takes the derivation to finish the phrase in which the item was merged. Suppose that some item \( \alpha \) (YP or Y) has undergone Merge with X, triggered by an appropriate feature \([\bullet Y_{0/2} \bullet]\) on X. Some other operation \( \Gamma \) (e.g., Agree) can then take place that requires the presence of \( \alpha \). After that, \( \text{Remove}(X, \alpha) \) applies, so that \( \alpha \) is not part of the structure anymore. Thus, due to the Strict Cycle Condition in (25), an \( \alpha \) merged with X cannot be removed by \([-F_{0/2-}]\) on another head Z merged later in the derivation. Therefore, \( \alpha \) is predicted to have a short life cycle: It is only accessible for other operations for a small part of the derivation. Given incremental, bottom-up derivations, this implies that an \( \alpha \) merged with X is accessible from within XP (downward accessibility) and inaccessible from positions outside of XP (upward inaccessibility), as stated in the Accessibility Generalization in (23). Thus, Remove interacts with other operations in such a way that it counter-bleeds \( \Gamma \) operations applying in XP (because Remove comes too late to block application of \( \Gamma \)) but bleeds subsequent operations requiring \( \alpha \) to be present (because it removes \( \alpha \) from the structure); see Chomsky (1951), Kiparsky (1973).

There is empirical evidence for short life cycle effects in various domains, which can thus be viewed as confirmations of a Remove-based approach (see the references given above). As we will see in the next section, there is also evidence for short life cycle effects with DP_{ext} arguments that are subject to Remove in passive clauses in German.\(^{14}\)

\(^{13}\)For instance, removal of phrases is argued to underlie variable accessibility of theme arguments in applicative constructions in German in Müller (2017), and variable syntactic accessibility of deleted material in sluicing constructions in English, German and Serbo-Croatian in Murphy & Müller (2016). In contrast, the approach to restructuring in German sketched in Müller (2017) relies on removal of heads, with restructuring verbs embedding a CP throughout that can then be reduced (recursively) to a TP, or even to a vP (so as to permit operations like scrambling, pronoun fronting, and negation scope assignment to target the matrix domain), or, most radically and with only a subset of basic restructuring verbs, to a bare VP (so as to permit long-distance passives). In Müller (2018), the complex prefield construction in German, which exhibits conflicting evidence as to whether it involves a single topicalized VP with an empty V head or separate XPs occupying multiple specifiers of C in what is otherwise a strict verb-second language, is analyzed in terms of Remove applying to the VP shell.

\(^{14}\)In principle, there is one potential loophole that makes it possible to extend the life cycle of an item that is subject to removal, viz., movement. If an item is moved to a higher domain, it
4. Life, Death, and Resurrection

4.1. Proposal

I would like to suggest that the core operation in passive constructions is the removal of a DP$_{ext}$ merged in Spec$_v$. More specifically, suppose that 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 DP$_{ext}$). [-D$_2$] on v will remove an existing DP specifier of v. The generation of a passive v head is illustrated in (30). First, a standard (‘active’) v head is selected from the lexicon, equipped with two structure-building features that induce Merge operations with a VP and DP$_{ext}$, in that order; see (30a). Second, a feature for DP removal is added to such a v below the Merge-inducing features of the list in the numeration; see (30b).

(30) Passive v:

a. Lexicon: v[●V$\_2$●]>[●D$\_2$●]
b. Numeration: v[●V$\_2$●]>[●D$\_2$●]>[-D$_2$-]

In languages like German, passivization goes hand in hand with structural case absorption. This is a consequence of whatever derives Burzio’s generalization. There are various ways to implement the effect in the present approach. Let me briefly consider three options. The first two possibilities rely on the assumption that structural accusative case is assigned by v under Agree, triggered by a probe feature [∗acc∗] on v (where [∗F∗] designates a probe feature for Agree); [∗acc∗] must be low on the list of features that induce syntactic operations. It can then simply be stipulated as a pre-syntactic restriction on v heads that [∗acc∗] cannot occur on v in the numeration if [-D$_2$-] shows up. A version of this approach relocates feature deletion from the numeration to the syntax. On this view, the presence of [-D$_2$-] on v in the syntax leads to deletion of a [∗case∗] feature on v.\footnote{This implies that probes can be deleted locally when the need arises; see Béjar \& Řezáč (2009), Preminger (2014), and Georgi (2014a), among others. Depending on what one assumes about active v with unergative intransitive verbs, such an option might be independently required.} Case feature deletion in both of these analyses can be given a rationale based on count invariants (see Stabler (1996)), in the sense that v assumes that the number of DPs and case features is balanced; undoing can be targeted there by a head with a [-F$_{0/2}$] feature, in accordance with the Strict Cycle Condition in (25). See Murphy (2016), Müller (2018) for evidence and discussion.
the effect of a \([\bullet D_2 \bullet]\) feature by discharging a \([-D_2-]\) feature may therefore be taken to imply the automatic deletion of a case feature as a repair operation re-establishing an equal number of DPs and case features for them. A third, slightly different option relies on a dependent case approach to argument encoding (see Marantz (1991), Bittner & Hale (1996), Wunderlich (1997), Stiebels (2000), McFadden (2004), Schäfer (2012a), Preminger (2014), Baker (2015), and Bobaljik (2015), among others). On this view, accusative case is assigned not by \(v\), but by a higher DP in the same phase, and if there is no such higher DP (as a consequence of Remove), accusative assignment will not be possible. Like the first two approaches, this approach requires case assignment to follow argument removal. A choice among these different approaches to case absorption, while ultimately far from trivial, is not required for present purposes; these issues are strictly speaking orthogonal to the main issues addressed in the present article, and all three versions are equally compatible with the overall analysis. For the sake of concreteness, I will adopt the second kind of approach in the following derivations; i.e., passive \(v\) is equipped with an \([\ast acc\ast]\) case feature in the numeration which is deleted in the syntax after \([-D_2-]\) is discharged.

On the basis of these assumptions, consider a simple German passive construction based on a transitive verb, as in (31).

\[
\text{(31) dass DP}_{ext} \text{ das Buch}_2 \text{ gelesen wurde } \\
\text{that the book}_{nom} \text{ read was } \\
\text{‘that the book was read.’}
\]

According to present assumptions, the Remove-based derivation of (31) looks as in (32).

\[
\text{(32) a. Selection of } v \text{ and } VP \\
\v'[\bullet D_2 \bullet] \text{ [-}D_2-\text{] -} \text{ [\ast acc\ast], [VP das Buch gelesen ]} \\
b. \text{Merge}(v, VP) \\
[v' \v' [\bullet D_2 \bullet] \text{ [-}D_2-\text{] -} \text{ [\ast acc\ast] [VP das Buch gelesen ]}] \\
c. \text{Merge}(DP_{ext}, v') \\
[\v' \v' \v' [\bullet D_2 \bullet] \text{ [-}D_2-\text{] -} \text{ [\ast acc\ast] [VP das Buch gelesen ]}]] \\
d. \text{Syntactic activity of } DP_{ext} \\
[\ldots] \\
e. \text{Remove}(DP_{ext}, v') \\
[\v' [\ast acc\ast] [VP das Buch gelesen ]]
\]
f.  **Case feature deletion**

\[ v_P \nu [v_P \text{das Buch gelesen}] \]

The output in (32f) has a DP with an unvalued case feature. This DP will be assigned nominative by T later in the derivation. Crucially, between Merge(DP\textsubscript{ext}, v') in (32c) and Remove(DP\textsubscript{ext}, v') in (32e), there is an option of carrying out other operations, including in particular those that require the presence of DP\textsubscript{ext}. This narrow window, defined by the m-command domain of DP\textsubscript{ext} (i.e., vP) determines the short life cycle of external arguments in passive derivations – once the derivation has moved on beyond the vP domain, there is no DP\textsubscript{ext} anymore that could be accessed by syntactic operations.

Thus, all the pieces of evidence discussed in section 2, and the Accessibility Generalization in (23) that captures them, can be accounted for in a very direct, simple way. In the next two subsections, I will first illustrate the accessibility of DP\textsubscript{ext} for vP-internal material, and then turn to the inaccessibility of DP\textsubscript{ext} for vP-external material.

4.2.  **Life**

Recall that DP\textsubscript{ext} can carry out control into adjunct clauses (cf. (1)), control into secondary predicates (cf. (2)), control into complement clauses (cf. (3)), binding of anaphors (cf. (4)), and binding of proper names (cf. (5)). Suppose that control and reflexivization both instantiate Agree operations (cf., e.g., Landau (2013) and Reuland (2011), respectively), such that non-overt PRO arguments and overt anaphors derive a binding index from an antecedent that they undergo Agree with.\(^{16}\) These Agree operations can then successfully be carried out, establishing control and binding relations with items c-commanded by DP\textsubscript{ext} in Spec\(v\textsubscript{P}\), in the part of the derivation in (32d) that lies between Merge(DP\textsubscript{ext}, v') and Remove(DP\textsubscript{ext}, v').

As a first illustration of DP\textsubscript{ext} accessibility, consider the case of control into

\(^{16}\)It should be noted, though, that these assumptions are not crucial. If control involves movement (cf., e.g., Boeckx, Hornstein & Nunes (2010)), or if reflexivization involves movement (cf., e.g., Fischer (2006)), the consequences will not be radically different from the present perspective. – That said, it has to be assumed that if the Agree relations are mediated by functional heads rather than established directly between two DPs, it has to be \(v\) rather than T (or some other vP-external head) that is involved – when T is merged, DP\textsubscript{ext} has already been removed from Spec\(v\textsubscript{P}\). (For reasons of simplicity and overall coherence, I will assume that Agree is possible directly between two XPs in what follows.)
non-complements (adjuncts or secondary predicates); relevant examples are repeated in (33) (see (1a), (2a)).

(33)  
a. Das Schiff wurde DP_{ext1} versenkt [CP PRO\textsubscript{1} um die ship was sunk in order the Versicherung zu kassieren ]  
‘The ship was sunk in order to collect the insurance.’

b. Die Daten wurden DP_{ext1} [AP PRO\textsubscript{1} nackt ] analysiert  
‘The data were analyzed by someone who was naked.’

In (34a), v has undergone Merge with a VP that contains an adjunct (a purpose clause or a secondary predicate, in the cases discussed above), which in turn has a PRO subject that needs to get its binding index valued (signalled here by an empty box). Next, in (34b), DP\textsubscript{ext} is merged with v′, triggered by [•D\textsubscript{2}•] on v, which is discharged as a consequence of the operation. In (34c), DP\textsubscript{ext} undergoes Agree with the embedded PRO subject, and thereby establishes its index on it; i.e., control takes place. Possibly, other operations that are not triggered by v’s features may then also take place, but it is clear that the next operation induced by v will have to take place soon; and given v’s features, this is Remove. The precise DP target of Remove does not have to be stipulated. As shown in (34d), Remove targets DP\textsubscript{ext} in Spec\textsubscript{v} rather than, say, DP\textsubscript{int} in VP (or some other DP in a more deeply embedded position, like PRO in the adjunct). This is so because of the Strict Cycle Condition (cf. (25)): In (34c), DP\textsubscript{ext} is the only DP that can be affected by v’s [−D\textsubscript{2}−] feature. Finally, in (34e), case probe deletion takes place. As a consequence, DP\textsubscript{int} cannot be assigned accusative case; it is assigned nominative case via Agree with T later in the derivation. The representations in (34de) illustrate counter-bleeding of control of PRO by DP\textsubscript{ext}: Remove would bleed control (because it removes the context in which control can apply) but comes too late to actually do so because control has already applied, and has instantiated an index on PRO (viz., the index of DP\textsubscript{ext}). More generally, in the present approach, instances of accessibility of DP\textsubscript{ext} for other operations always involve cases of opacity, in Kiparsky’s (1973) sense: The output representation is opaque because it is not clear how control could have applied successfully – there is no controller left at this point.
Control into non-complements in passive derivations

a. Merge(\(v, VP\))

\[
\begin{array}{c}
\text{\(v'\)} \\
\text{\(v\{\bullet D\} \rightarrow [-D_2] \rightarrow [\ast acc\ast]\)} \\
\text{\(\text{CP/AP} \quad \text{VP}\)} \\
\text{\(\text{PRO} \quad \text{...} \quad \text{DP}_{int} \quad \text{V}\)}
\end{array}
\]

b. Merge(\(DP_{ext}, v'\))

\[
\begin{array}{c}
\text{\(vP\)} \\
\text{\(DP_{ext_1} \quad \text{\(v'\)} \quad \text{\(VP\)} \quad \text{\(\text{CP/AP} \quad \text{VP}\)} \quad \text{\(\text{PRO} \quad \text{...} \quad \text{DP}_{int} \quad \text{V}\)}
\end{array}
\]

c. Agree(\(DP_{ext}, PRO\))

\[
\begin{array}{c}
\text{\(vP\)} \\
\text{\(DP_{ext_1} \quad \text{\(v'\)} \quad \text{\(VP\)} \quad \text{\(\text{CP/AP} \quad \text{VP}\)} \quad \text{\(\text{PRO} \quad \text{...} \quad \text{DP}_{int} \quad \text{V}\)}
\end{array}
\]
d. \( \text{Remove}(\text{DP}_{\text{ext}}, v') \)

\[
\begin{array}{c}
vP \\
\v_{[\ast \text{acc}\ast]} \\
\text{VP} \\
\text{CP/ AP} \\
\text{PRO}_1 \\
\ldots \\
\text{DP}_{\text{int}} \\
\text{V}
\end{array}
\]

e. \text{Case probe deletion}

\[
\begin{array}{c}
vP \\
v \\
\text{VP} \\
\text{CP/ AP} \\
\text{PRO}_1 \\
\ldots \\
\text{DP}_{\text{int}} \\
\text{V}
\end{array}
\]

As a second illustration of temporal accessibility of \( \text{DP}_{\text{ext}} \), consider binding. Relevant examples (licensing of anaphors and Principle C effects) are repeated in (35) (see (4a), (5a)).

(35)  

a. Hier wurde \( \text{DP}_{\text{ext}} \) sich \( _1 \) nicht geprügelt
   here was \( \text{REFL} \) not \( \text{hit} \)
   'There were no rows here.'

b. *Gestern wurde \( \text{DP}_{\text{ext}} \) Fritz \( _1 \) eingeladen
   yesterday was \( \text{Fritz} \) invited
   'Yesterday, Fritz invited himself.'

Consider the abstract derivation in (36). In (36a), a passive \( v \) (i.e., a \( v \) that has a \([\ast \text{D}_2\ast]\) added below its structure-building features) has combined with a \( \text{VP} \) that contains a reflexive pronoun; the latter does not have a binding index yet. In (36b), \( \text{DP}_{\text{ext}} \) is merged with \( v' \). Next, \( \text{DP}_{\text{ext}} \) can undergo Agree with the object reflexive, instantiating its index on it and thereby triggering reflexivization; see (36c). After that, \([\ast \text{D}_2\ast]\) on \( v \) induces Remove of a \( \text{DP} \) in (36d); the Strict Cycle Condition ensures that it is \( \text{DP}_{\text{ext}} \) (rather than \( \text{DP}_{\text{refl}} \))
that undergoes removal. Finally, the accusative case probe on v is deleted in (36e). As before, the output representations after DP\textsubscript{ext} removal involve counter-bleeding: Remove of DP\textsubscript{ext} would bleed reflexivization but comes too late to have this effect.\footnote{This presupposes that Principle A (or, under current assumptions, the requirement for reflexives to derive a binding index via Agree with another DP in the same phase) is not a representational constraint (a ‘filter’) that is checked on output representations. Indeed, there is general consensus that Principle A is an Anywhere Principle that can be satisfied at any step of the derivation (see Belletti & Rizzi (1988), Epstein et al. (1998), among many others).}

(36) \textit{Binding in passive derivations}

a. $\text{Merge}(v, VP)$

```
     v'
   ┌───┐
   │   │
   │ v │ VP
   └───┘
       ┌───┐
       │   │
       │ DP\textsubscript{refl} V │
               └───┘
```

b. $\text{Merge}(\text{DP\textsubscript{ext}}, v')$

```
     vP
   ┌───┐
   │   │
   │ vP │ VP
   └───┘
       ┌───┐
       │   │
       │ v' │ V
       └───┘
         ┌───┐
         │   │
         │ v' │ V
         └───┘
```

c. $\text{Agree}(\text{DP\textsubscript{ext}}, \text{DP\textsubscript{refl}})$

```
     vP
   ┌───┐
   │   │
   │ vP │ VP
   └───┘
       ┌───┐
       │   │
       │ vP │ V
       └───┘
         ┌───┐
         │   │
         │ v' │ V
         └───┘
d. \( \text{Remove}(\text{DP}_{\text{ext}}, v') \)

\[
\begin{array}{c}
\text{vP} \\
/\ \\
v_{[*\text{acc}*]} \\
\text{VP} \\
/\ \\
\text{DP}_{\text{ref}l_1} \quad V
\end{array}
\]

e. \textit{Case probe deletion}

\[
\begin{array}{c}
\text{vP} \\
/\ \\
v \\
\text{VP} \\
/\ \\
\text{DP}_{\text{ref}l_1} \quad V
\end{array}
\]

The other cases of control and binding by \( \text{DP}_{\text{ext}} \) discussed above are derived in essentially the same way. Similarly, quantificational variability effects that indicate unselective binding of \( \text{DP}_{\text{ext}} \) by an adverb of quantification (see (20)) can be accounted for: Given that the adverb of quantification is merged in an outer specifier of \( v \), it can successfully bind \( \text{DP}_{\text{ext}} \) in an inner specifier of \( v \) before the latter is removed from the syntactic representation. Finally, note that the counter-bleeding effect with \( \text{Remove}(\text{DP}_{\text{ext}}, v') \) as it arises both with index valuation via Agree (as in control and binding configurations) and with unselective binding presupposes that effects established by binding/control relations persist for semantic interpretation after \( \text{DP}_{\text{ext}} \) has been removed from the structure (I will come back to this issue in the final section of the paper).

4.3. \textit{Death}

Recall next the evidence illustrating inaccessibility of \( \text{DP}_{\text{ext}} \): \( \text{DP}_{\text{ext}} \) cannot be bound by a quantified DP from a position outside of \( \text{vP} \), e.g., by an argument in a matrix clause (cf. (7), (8)); \( \text{DP}_{\text{ext}} \) does not license a control infinitive by providing a target for a controller in a matrix clause (cf. (9)); \( \text{DP}_{\text{ext}} \) cannot undergo criterial movement to the SpecC position of verb-second clauses (cf. (11)); and \( \text{DP}_{\text{ext}} \) does not act as an intervener for A-movement to SpecT (cf. (13), (15)). All these operations involve items outside of \( \text{vP} \), and when they have a chance to take place, \( \text{DP}_{\text{ext}} \) has long been removed from the structure.
A first example illustrating this effect is the impossibility of binding of $\text{DP}_{\text{ext}}$ by a DP in the matrix clause, as in (37) (cf. (7a), (8a)).

(37) Binding from above:

a. *Kein Student$_1$ gibt zu [CP dass $\text{DP}_{\text{ext}_1}$ schlecht gearbeitet]
   no student admits that badly worked
   was
   'No student admits that he did not work well.'

b. *Er hat den meisten Lehrern$_1$ gesagt [CP dass $\text{DP}_{\text{ext}_1}$ der Maria Bücher geschenkt werden sollen]
   he has the most teachers$_{\text{dat}}$ said that the Maria$_{\text{dat}}$ books$_{\text{nom}}$ given are should
   'He told most teachers that they should give books to Maria.'

The abstract derivation in (38) (underlying (37a)) shows why $\text{DP}_{\text{ext}}$ cannot be interpreted as a variable bound by a quantified DP in the matrix clause. In (38a), $\text{DP}_{\text{ext}}$ is merged with $v'$ headed by a passive $v$. In the next step in (38b), Remove applies; this concludes the short life cycle of $\text{DP}_{\text{ext}}$. After this (in (38c)), there is case probe deletion. The derivation then finishes the basic $vP$ projection and moves on to the TP, CP, and matrix VP cycles (these steps are left out here; see (38d)). Finally, a quantified DP is merged in the matrix Spec$v$ position; see (38e). However, as shown here, this DP cannot bind the embedded $\text{DP}_{\text{ext}}$ for the simple reason that there is no $\text{DP}_{\text{ext}}$ present anymore at this point. Consequently, in this case, the interaction of operations is transparent rather than opaque: Remove in the embedded clause applies before Merge in the matrix clause and therefore bleeds binding of the embedded $\text{DP}_{\text{ext}}$ by a matrix DP.

(38) Bound variable interpretation in passive derivations

a. Merge($\text{DP}_{\text{ext}}, v'$)

$$
\begin{array}{c}
\text{vP} \\
\text{v'} \\
\text{v'[-D$_2$-][*acc*]} \\
\text{VP} \\
\ldots \text{V}
\end{array}
$$
b. \[ \text{Remove}(DP_{ext}, v') \]

\[
\begin{array}{c}

vP \\
\downarrow \\
 VP \\
\downarrow \\
\vdots V
\end{array}
\]

\[ v_{[*\text{acc}*]} \]

c. \[ \text{Case probe deletion} \]

\[
\begin{array}{c}

vP \\
\downarrow \\
 VP \\
\downarrow \\
\vdots V
\end{array}
\]

d. ...

e. \[ \text{Merge}(DP_{ext}, v') \text{ in the matrix clause} \]

\[
\begin{array}{c}

vP \\
\downarrow \\
 DP_{ext} \\
\downarrow \\
 v' \\
\downarrow \\
 VP \\
\downarrow \\
 CP \\
\downarrow \\
\vdots vP \ldots \\
\downarrow \\
 vP \\
\downarrow \\
\vdots V
\end{array}
\]

As a second and final illustration of how inaccessibility of $DP_{ext}$ is derived, consider the lack of minimality effects with local movement of an object DP to subject position (i.e., to SpecT); see (39) (= (13)).

\[(39) \quad [\text{TP John}_2 \text{ was } [vP \overrightarrow{DP_{ext,1}} [v' v [vP \text{ killed t}_2 ]]]] \]
In (40a), DP\textsubscript{ext} has been merged with v'; in (40b), DP\textsubscript{ext} is subject to Remove, and leaves the structure again. After this, in (40c), the structural case probe is deleted. No more operations take place within vP, and vP is next merged with T; see (40d). English T has a [●D\textsubscript{2}●] (i.e., EPP) feature that requires DP\textsubscript{int} to raise to its specifier; by doing so, it now does not have to move across an intervening DP\textsubscript{ext} in Spec\textsubscript{v} anymore; see (40e). This accounts for the absence of a minimality effect: There is no intervention because there is no potential intervener at the stage of the derivation where the movement step takes place.

(40) \textit{Minimality in passive derivations}

\begin{enumerate}
\item \textit{Merge}(DP\textsubscript{ext}, v')

\begin{tikzpicture}
\Tree [.
  \node {vP}
  [.\node {v'}]
  [.\node {v}[\node {v[−D\textsubscript{2}−]>[∗acc∗]}]
    [.\node {VP}
      [.\node {DP\textsubscript{int}}]
      [.\node {V}]
    ]
  ]
]
\end{tikzpicture}

\item \textit{Remove}(DP\textsubscript{ext}, v')

\begin{tikzpicture}
\Tree [.
  \node {vP}
  [.\node {v[∗acc∗]}]
  [.\node {VP}
    [.\node {DP\textsubscript{int}}]
    [.\node {V}]
  ]
]
\end{tikzpicture}

\item \textit{Case probe deletion}

\begin{tikzpicture}
\Tree [.
  \node {vP}
  [.\node {v}
    [.\node {VP}
      [.\node {DP\textsubscript{int}}]
      [.\node {V}]
    ]
  ]
]
\end{tikzpicture}
\end{enumerate}
d. \( \text{Merge}(T,vP) \)

\[
\begin{array}{c}
T' \\
\text{T[●D}_2\text{●]} \\
vP \\
\text{v} \\
\text{VP} \\
\text{DP}_{\text{int}} \\
\text{V}
\end{array}
\]

e. \( \text{Move}(\text{DP}_{\text{int}}, T'): \text{Minimality respected} \)

\[
\begin{array}{c}
\text{TP} \\
\text{DP}_{\text{int}} \\
\text{T'} \\
\text{T} \\
vP \\
\text{v} \\
\text{VP} \\
\text{V}
\end{array}
\]

Other cases exhibiting inaccessibility of \( \text{DP}_{\text{ext}} \) can be derived in the same way.

4.4. Intransitive Constructions and Strict Cyclicity

The analysis makes a clear prediction concerning unaccusative verbs. As noted above, structure removal exhibits short life cycle effects: Because of the Strict Cycle Condition, an item can only be targeted by Remove on the same XP cycle on which it has been merged; in addition, the XP needs to be the current root projection. Given that passive is identified with the addition of \([-\text{D}_2-]\) to \(v\) in the numeration, \(\text{DP}_{\text{int}}\) arguments of unaccusative intransitive verbs, which are merged within VP, are expected not to give rise to impersonal passives, in contrast to \(\text{DP}_{\text{ext}}\) arguments of unergative intransitive verbs, which are merged within \(vP\). This is so for the very same reason that \([-\text{D}_2-]\) on \(v\) does not intrinsically have to be associated with information as to which DP in a transitive clause it is that is affected by Remove (\(\text{DP}_{\text{ext}}\) or \(\text{DP}_{\text{int}}\); see above).
As first observed by Perlmutter (1978) (for Dutch), this prediction is correct; compare the well-formed cases of passivization with unergative verbs in (41) with the ill-formed cases of passivization with unaccusative verbs in (42).

(41)  
\begin{align*}
a. & \text{Hier wird jetzt gearbeitet} \\
& \text{here is now worked} \\
& \text{‘People are working here now.’} \\
\end{align*}

\begin{align*}
b. & \text{Getanzt wurde nicht} \\
& \text{danced was not} \\
& \text{‘There was no dancing.’} \\
\end{align*}

(42)  
\begin{align*}
a. & \text{*Hier wird jetzt gefallen} \\
& \text{here is now fallen} \\
& \text{‘People fall here now.’} \\
\end{align*}

\begin{align*}
b. & \text{*Es wurde angekommen} \\
& \text{it was arrived} \\
& \text{‘People arrived.’} \\
\end{align*}

As noted in footnote 14, there is a potential loophole for circumventing short life cycle effects: Movement of some item can extend its accessibility, with subsequent removal in a derived specifier position. To close this loophole for passivization in German, it must be ensured that DP_{int} cannot move to Spec_{cv} and be subject to removal in this position after all. There are various possibilities that can be pursued here. For present purposes, it may suffice to assume that a v that does not introduce a DP_{ext} is a defective phase head (see Chomsky (2001) vs. Legate (2003)), in the sense that it cannot be equipped with any structure-building feature (i.e., it cannot bear features that might trigger intermediate or criterial movement to Spec{cv}).

---

18 Primus (2010, 2011) and Kiparsky (2013) claim that there are well-formed cases of passivization of unaccusative verbs in German, but, to the extent that the relevant grammaticality judgements can be substantiated, it would seem that they involve a meta-grammatical use of an otherwise illegitimate construction or a re-interpretation of unaccusative verbs as unergative (by assigning an agent-like interpretation to the sole DP argument).

19 See Heck & Müller (2016) for independent evidence for the defective nature of unaccusative v in German, based on extraction options in ECM environments. However, ultimately a bit more will have to be said since the same scenario (with movement feeding removal) will also have to be excluded in the case of a transitive construction, where it must be ensured that \([-D_2-]\) removes DP_{ext}, not DP_{int} (cf. subsection 4.2), which would give rise to an antipassive-like structure. The relevant configuration after movement of DP_{int} to Spec_{cv} would look as in (i).
4.5. Voice and v

The present approach postulates a dual role for v: This functional head is responsible both for manipulation of $DP_{ext}$ and for accusative case assignment. There does indeed seem to be quite a bit of evidence for this clustering of features on a single functional head cross-linguistically; see, e.g., Coon & Preminger (2011) on Chol. However, it has also sometimes been argued that $DP_{ext}$ and case manipulation should be distributed across two separate heads, Voice and v. Let us see what consequences would arise under the present approach. Suppose first that $DP_{ext}$ is introduced and removed by Voice, whereas v is the locus of accusative case assignment. This would leave the gist of the Remove-based approach to passivization in German intact; however, the resulting analysis would cease to be compatible with two of the three approaches to case absorption discussed as options above: An $[^{\text{acc}}\ast]$ feature on v could not be deleted pre-syntactically as the relevant information (viz., $[-D_{2}-]$) would be located on a different head (viz., Voice); and $[^{\text{acc}}\ast]$ could also not be deleted in the syntax because this would require either look-ahead capacity (when case feature deletion needs to be decided on, Voice is not yet part of the structure) or a massive violation of strict cyclicity (by going back to an embedded domain and deleting the case feature both on v and on $DP_{int}$ that has already undergone Agree with it). So, this would leave the dependent case approach as the sole remaining option.

Suppose next that $DP_{ext}$ is introduced by v, and v is also the locus of accusative case, but argument removal is handled by Voice (i.e., Voice would basically be a designated Passive head); cf., e.g., Collins (2005), Merchant (2013). On this view, $DP_{ext}$ would always have to move from Specv to SpecVoice to be accessible to the $[-D_{2}-]$ feature on Voice (because of the Strict Cycle Condition); and whereas this might technically work (and be compatible with the ban on such movement of $DP_{int}$ in regular transitive and unaccusative contexts discussed at

(i) $[vP\; DP_{int}\; [v^{\prime}\; DP_{ext}\; [v^{\prime}\; [vP\ldots]\; v]]]]$

A removal of $DP_{int}$ by v in (i) is excluded if Remove is subject to a minimality requirement, in the same way that Merge is subject to minimality in tucking-in contexts. On this view, the inability of internal Merge to feed Remove of $DP_{int}$ has two slightly different sources in unaccusative and transitive environments. However, both accounts are compatible with the assumption that Remove can be fed by movement in the different syntactic environments identified in Murphy (2016) for double passivization in Turkish and in Müller (2018) for complex prefields in German.
the end of the previous subject), it seems clear that the obligatory movement step preceding removal would not be independently motivated.

In the absence of strong arguments for separate Voice and v heads in German, I would therefore like to conclude that there is every reason to maintain v as the sole locus of DP\textsubscript{ext} introduction, DP\textsubscript{ext} removal, accusative case assignment, and accusative case deletion. Given that the approach to passivization developed here can be taken to make cross-linguistic predictions, this implies that arguments for a simultaneous presence of VoiceP and vP in passive constructions as they have been advanced for some languages need to be re-evaluated. For reasons of space and coherence, I cannot attempt to do this here in any detail; I will confine myself to pointing out two relevant cases that will eventually have to be looked at in detail. First, there is morphological evidence based on affix order (and the Mirror Principle) in Hiaki that suggests separating Voice and v; see Harley (2013). Sundaresan & McFadden (2014) develop a similar argument on the basis of Tamil. Second, Merchant (2013) presents syntactic evidence based on ellipsis of verbal categories under identity in English.\footnote{To wit, with VP ellipsis constructions in English, the elided constituent looks as though it can have a different voice value than its antecedent; see, e.g., (i). Merchant observes that there is no actual voice mismatch between antecedent and elided constituent here if active/passive information is located on a separate head Voice, and deletion affects the lower constituent vP. This analysis also correctly predicts voice mismatches to be impossible if the elided constituent must include VoiceP (as with sluicing, which affects TP).}

Assuming that these (and other) arguments for separating Voice

\begin{enumerate}
\item[(i)] This problem was to have been looked into, but obviously nobody did look into this problem.
\item[(ii)] *They sell Hyundais in Greece because Hondas don’t sell
\end{enumerate}

However, as noted by Merchant (2013, 89–90), an alternative analysis where v contains voice information, and VP is in fact deleted, would work just as well for the core data. Furthermore, as also noted by Merchant, while the analysis in terms of Voice and v does without a voice feature clash between antecedent and elided constituent, it cannot actually postulate complete identity of the deleted item – in (i), e.g., Specv in the antecedent is filled by the external argument that Merchant takes to be syntactically represented in passives, whereas Specv in the deleted constituent is filled by a trace of nobody. This problem disappears if it is VP rather than vP that undergoes deletion. The core argument for postulating both Voice and v comes from the behaviour of anticausative and middle alternations. In (ii), e.g., there is no active/passive voice contrast, but v is different (transitive vs. middle) according to Merchant’s assumptions, so deletion is not licensed.
and \( v \) can be successfully addressed on the basis of the present approach, I will next turn to the question of what happens with a removed DP\(_{ext}\).

### 4.6. Resurrection

There are three basic issues that still need to be clarified in the Remove-based approach to passive constructions: What is the nature of DP\(_{ext}\), where does DP\(_{ext}\) go once it is removed from a structure, and where do by-phrases come from? Let me address the second question first. 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. By complete analogy, Remove can be expected to put a (possibly complex) item back into the workspace. I assume (also based on evidence from Remove operations outside of passivization) that removed material can in principle remain in the workspace for the rest of the derivation without giving rise to ungrammaticality.\(^{21}\) However, it is clear that material that was once part of the syntactic structure of a sentence and is now contained in the abstract workspace of the derivation can in principle give rise to recoverability problems; this leads directly to the first question, concerning the nature of DP\(_{ext}\).

So far, nothing has been said about what DP\(_{ext}\) looks like in passive constructions. Indeed, I would like to contend that DP\(_{ext}\) can be virtually anything – a pronoun, a proper name, a full DP of any type; and it can contain any amount of structure (relative clauses, argument clauses, etc.). However, in all cases but one, leaving a removed DP\(_{ext}\) in the workspace for the remainder of the derivation will give rise to a fatal recoverability problem. The sole exception to this is the maximally uninformative, unmarked DP type, viz., an indefinite pronoun, given that bare indefinites are interpreted as variables

---

\(^{21}\)Still, this evidence can in principle also be accommodated in the approach advocated here if it is assumed that the difference between transitive verbs and their middle or anticausative versions is not syntactic, but lexical. On this view, V (rather than \( v \)) is featurally different in (ii), and this accounts for the impossibility of VP deletion.

---

\(^{21}\)Thus, this approach 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 postulated that there are two separate domains of the workspace reserved for the two different types of linguistic expressions.
The Short Life Cycle of External Arguments

(plus, possibly, contextually enriched restrictions); see Heim (1982). These DP\textsubscript{ext} arguments, and only these, can stay in the workspace, and trigger default existential quantification – unless, that is, they have already been bound by an adverb of quantification in the syntax; see subsection 2.3 above (and recall that quantificational variability effects cannot occur in the passive with an adverb outside the minimal clause, which follows directly under the present approach). All other kinds of DP\textsubscript{ext} cannot permanently stay in the workspace without generating a recoverability violation; so they are remerged into the structure in the only way that is available in the absence of structure-building features, viz., as an adjunct. This answers the third question posed at the outset: By-phrases are resurrected DP\textsubscript{ext} arguments that re-enter the syntactic tree from the workspace of the derivation; the accompanying preposition is selected via a last-resort access to the lexicon. It follows that the by-phrase is accessible for subsequent syntactic operations.

Three further remarks are due here. First, consider again the issue of non-intervention of DP\textsubscript{ext} in passive clauses. DP\textsubscript{ext} does not block object movement to Spec\textsubscript{T} via minimality because it is not present anymore in the structure when object movement takes place; but what about the by-phrase? One possibility is that the by-phrase does not intervene because it is a PP (not a DP) after all. Alternatively, the reason for non-intervention might be that the by-phrase is merged after movement of DP\textsubscript{int} to either an intermediate or a final position has taken place (cf. Epstein et al. (1998); note that this latter option would presuppose counter-cyclic Merge of adjuncts; possibly this would account for why the target position can be quite low, next to the verb).

Second, the morphological realization of DP\textsubscript{ext} depends on the properties of the prepositional head. In German, von (‘by’) assigns dative case to DP\textsubscript{ext}, which the latter would not otherwise have (but note that the case feature of DP\textsubscript{ext} has not yet been valued syntactically prior to Remove).

And third, there is the issue of locality and timing of DP\textsubscript{ext} resurrection via by-phrase integration. I suggest that a DP\textsubscript{ext} that is removed from Spec\textsubscript{v} and then subsequently remerged into the structure as an adjunct must do so before the derivation moves on to the next phase. This excludes cases as in (43a), where DP\textsubscript{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 (43b)) (or (7a)), where DP\textsubscript{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 $\bullet D_2 \bullet$ features based on $\theta$-roles).

(43)  

\begin{enumerate}
\item a. *Karl gibt \_[PP von keinem Studenten] zu \_[CP dass gut Karl admits of no student to that well gearbeitet wurde] \[wurde] \[\text{\textquoteleft Karl admits that no student worked well.\textquoteright} \]
\item b. *[\_[DP Die zwei Leute] glauben \_[CP dass einander] gedankt the two people believe that each other thanked \[wurde] \[\text{\textquoteleft The two people believe that each one of them thanked the other.\textquoteright} \]
\end{enumerate}

5. Remarks on Variation

Under present assumptions, the core property of passivization in German is the presence of $[-D_2-]$ on $v$ in the numeration. Other properties of passive constructions are either secondary (structural case absorption is a consequence of argument removal); or they do not show a uniform behaviour and permit exceptions (there are different morphological reflexes with different types of passivization in German, and in the case of the lassen passive, there is no morphological reflex at all; cf. (16)); or they are orthogonal (obligatory object promotion to subject position shows up in languages where $T$ always has an EPP feature, which is not the case in German). In line with this, I would like to suggest that $[-D_2-]$-driven removal of $DP_{ext}$ is the sole cross-linguistically invariant property of passivization. On this view, a main locus of variation in passive constructions concerns the interaction of argument removal via $[-D_2-]$ with case absorption: If case is determined after removal of $DP_{ext}$, structural case cannot be assigned anymore by $v$; but if the order is reversed, passive will not be accompanied by case absorption (as an instance of counter-bleeding).

For concreteness, suppose, as before, that case absorption is handled in terms of case feature deletion on $v$. Then, a language without obligatory case absorption (like Ukrainian, Northern Russian varieties, and Czech, among many others) will result if $[-D_2-]$ can show up below $[\ast case\ast]$ on $v$. Further variation in this area arises in double object constructions. Focussing just
on Germanic languages, the following picture emerges: If, in a double object construction, v has two structural case probes, \([-D_2-]\) may rank above both \([*\text{case}*]\) features, as in Danish and English, where the higher object case is absorbed; see Vikner (1990). Alternatively, \([-D_2-]\) may show up between the two case probes, as in Dutch, where only the lower case is absorbed; see Zwart (1993) (but also cf. Haegeman (2016) on qualifications based on varieties of Dutch). Finally, it may use either option, as in Norwegian, with a single passive auxiliary à bli; see again Vikner (1990). German is like Norwegian, but with the two options accompanied by a different choice of passive auxiliary: There are two verbal passives in German, one with the passive auxiliary \(\text{werden}\) and one with the passive auxiliary \(\text{bekommen (kriegen)}\); the second type of passive is sometimes called ‘recipient passive’, and it mainly shows up in double object constructions (see Höhle (1978), Reis (1985), Müller (1995, ch. 4), Fenselow (2001), Haider (2010), and Alexiadou, Anagnostopoulou & Sevdali (2014), among others). The standard passive with \(\text{werden}\) absorbs accusative case and leaves dative case intact; cf. (44ab). In contrast, the recipient passive absorbs dative case and leaves accusative case intact; cf. (44cd).

\[
\begin{align*}
(44) & \quad \text{a.} & \text{dass der Maria das Buch geschenkt wird.} \\
& \quad \text{that the Maria}_{\text{dat}} \text{ the book}_{\text{nom}} \text{ given is} \\
& \quad \text{b.} & \text{*dass die Maria das Buch geschenkt wird.} \\
& \quad \text{that the Maria}_{\text{nom}} \text{ the book}_{\text{acc}} \text{ given is} \\
& \quad \text{c.} & \text{dass die Maria das Buch geschenkt bekommt (kriegt).} \\
& \quad \text{that the Maria}_{\text{nom}} \text{ the book}_{\text{acc}} \text{ given gets} \\
& \quad \text{d.} & \text{*dass der Maria das Buch geschenkt bekommt (kriegt).} \\
& \quad \text{that the Maria}_{\text{dat}} \text{ the book}_{\text{nom}} \text{ given gets}
\end{align*}
\]

‘that Mary is given the book.’

In typical double object constructions, v has two structural cases to assign to VP-internal DPs: dative and accusative. Thus, the features for Merge and Agree that v needs to bear in double object constructions in German look as in (45).

\[
\begin{align*}
(45) & \quad V[\bullet V_2 \bullet]>[\bullet D_2 \bullet]>[*\text{dat}*]>[*\text{acc}*]
\end{align*}
\]

The data in (44) can then be accounted for by assuming that in German, \([-D_2-]\) may either be inserted directly above \([*\text{dat}*]\), or it may be inserted directly above \([*\text{acc}*]\). In the first case, a recipient passive construction results – dative cannot be assigned anymore, and the DP that would bear dative in an active
clause gets nominative case from T; cf. (46a). In the second case, a standard passive construction results – accusative case cannot be assigned anymore, and the DP that would bear accusative in an active clause gets nominative case from T; cf. (46b).

(46)  
\[
\begin{align*}
\text{a. } & \mathbf{V}[^{\bullet V_2 \bullet}] \rightarrow \bullet D_2 \bullet \rightarrow [\neg D_2 -] \rightarrow [* \text{dat} \ast] \rightarrow [* \text{acc} \ast] \\
\text{b. } & \mathbf{V}[^{\bullet V_2 \bullet}] \rightarrow \bullet D_2 \bullet \rightarrow ([* \text{dat} \ast] \rightarrow) \rightarrow [\neg D_2 -] \rightarrow [* \text{acc} \ast]
\end{align*}
\]

Finally, as far as the morphological realization of passivization is concerned, the present approach suggests that this issue is also tied to the presence of \([- D_2 -]\) on v. Depending on the presence or absence of this feature and its position on the feature stack, v itself may receive a different realization. In principle, v may also receive the same realization in active and passive environments, as is the case in Achenese (cf. Perlmutter & Postal (1983)), or in German lassen-passives. Alternatively, different kinds of passive v heads may be selected by different kinds of passive auxiliaries. In German, the passive auxiliary bekommen selects a vP headed by (46a); and the passive auxiliary werden selects a vP headed by (46b).

6. Conclusion and Outlook

To sum up, I have argued that modelling passive by Remove operations in a local derivational approach accounts for the variable syntactic accessibility of external arguments in passive derivations: Removal of the external argument DP\text{ext} triggered by \([- D_2 -]\) gives rise to counter-bleeding with operations confined to the m-command domain of v; but removal of DP\text{ext} triggered by \([- D_2 -]\) gives rise to bleeding with operations involving positions outside the

\[22\text{There is a technical issue here related to the interaction of morphological realization and feature discharge which shows up more generally in derivational approaches to syntax; see, e.g., Adger (2003). The problem is that strictly speaking, the features on v in (46) that need to be accessed by an embedding auxiliary V should be discharged and deleted by the time when V combines with the vP. There are two standard ways to make the relevant information available after all: One possibility is to postulate diacritics for the different kinds of v heads that morphological realization can then be sensitive to; another one (that is arguably preferable under present assumptions) is to assume that discharged features become inactive, but are still visible from outside (also cf. Chomsky (1995) on the difference between deletion and erasure). – Note incidentally that basically the same situation shows up with the approach to case absorption developed in section 4.2: Case feature deletion needs information (viz., \([- D_2 -]\) on v) that is not available anymore when the issue becomes relevant.}
m-command domain of v. On this approach, external arguments are indeed present in German passive constructions, but they have a short life cycle in which they can be syntactically active: the period between discharge of \([\bullet D_2 \bullet]\) and discharge of \([-D_2-]\) on one and the same head. As far as I am aware, no alternative approach to the passive exists that derives the Accessibility Generalization in \((23)\) in a simple, non-conspiratorial way; so, to the extent that this generalization is correct, I take the Remove-based approach to be corroborated by the empirical evidence. In addition, I have argued that an operation like Remove, as the complete mirror image of Merge, is to be expected from a minimalist perspective (see Müller (2017) for further evidence and discussion of this point); and Remove can be shown to be independently motivated on the basis of a number of other constructions (among them applicatives, antipassive, sluicing, restructuring, DP/NP oscillation in Slavic languages, and complex prefields; see Müller (2015) for an overview)).

Still, it goes without saying that the postulation of an operation Remove raises a number of non-trivial issues, and will ultimately require a rethinking of several core assumptions that are often taken for granted in Principles and Parameters approaches to syntax. Here I will only briefly mention one issue that would seem particularly obvious, viz., semantic interpretation. Structure removal in general, and the removal of DP\text{ext} in particular, may indeed lead to incompatibilities with the standard concept of transparent logical forms as laid out, e.g., in Heim & Kratzer (1998). However, 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 (see Chomsky (2001, 2013)). For concreteness, let me name two 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 they are invariantly assigned during the syntactic derivation; variable binding relations established in the derivation persist throughout the derivation. And second, if an argument remains in the workspace after the phase from which it has been removed is completed (i.e., if it is an indefinite DP\text{ext}, as discussed above, or if it is some other item that does not give rise to recoverability problems, as is the case with bare XP shells with restructuring and complex prefields as discussed in Müller (2017, 2018)), and has not found a binder in that phase, it is interpreted in the clause in which it does not structurally show up anymore as being bound via default existential quantification. These two requirements can be met 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).

Appendix: Denying the Accessibility Generalization in Standard Approaches

As noted at the end of section 2, the Accessibility Generalization in (23) poses problems both for approaches to the passive that envisage an unqualified syntactic accessibility of $\text{DP}_{\text{ext}}$ (see references on page 56), and for approaches in which $\text{DP}_{\text{ext}}$ is systematically inaccessible in the syntax (see references in footnote 1). In this appendix, I look at possible strategies to maintain either strict syntactic accessibility or strict syntactic inaccessibility in the light of the evidence discussed in the present article. The conclusion will be that neither strategy is successful.

Maintaining Accessibility

The first option is to strictly maintain an approach where $\text{DP}_{\text{ext}}$ is generally accessible in German passive constructions, and account for the instances of inaccessibility (which the generalization in (23) captures in a uniform way) separately, on a case-by-case basis. Thus, Collins (2005) derives the absence of minimality effects as in (13), (15a) by postulating a smuggling operation: A constituent including $\text{DP}_{\text{int}}$ and $V$ (the PartP, alternatively: VP) moves to a higher position (SpecVoice), across $\text{DP}_{\text{ext}}$, and $\text{DP}_{\text{int}}$ then undergoes extraction from the moved VP (PartP). However, this analysis is far from unproblematic. For instance, a smuggling derivation would normally be expected to incur a freezing effect: Extraction from a moved VP (or PartP) otherwise leads to ungrammaticality in German (cf. Müller (2014, ch. 3) and references cited there). Furthermore, as noted by Collins, smuggling is in fact incompatible with several constituency tests (given that a by-phrase is assumed to have by in Voice, and $\text{DP}_{\text{ext}}$ in Specv, such that it is not a proper phrase after all); e.g., it requires movement of non-constituents in cases like By whom was the book given to Mary?

Next, Pitteroff (2014) accounts for non-intervention of $\text{DP}_{\text{ext}}$ in (16b) by assuming that lassen-passive constructions differ from other passive constructions in German in that $\text{DP}_{\text{ext}}$ does not in fact show up here, the reason being
that passive and active complements of lassen differ in size. However, on the one hand, there is no independent evidence for this alleged difference in complement size (verbal morphology is identical, sentential adverbs have identical distributions, etc.); and on the other hand, all the tests that point to accessibility of DP_{ext} from below in German carry over to lassen-passives: Thus, as noted above, the very option of Principle A satisfaction with binding by the *embedded* subject in (16b) strongly suggests that DP_{ext} is structurally present in (16b) as it is in (16a). Similarly, DP_{ext} can control PRO in purpose clauses and secondary predicates in these contexts; cf. (47ab).

(47) a. Der König lässt [vP DP_{ext2} die Schule besuchen [CP um PRO_{1/2} mehr zu lernen ]] in order more to learn
   b. Der König lässt [vP DP_{ext2} das Handout übermüdet ] verfassen ]
   tired write

As for the evidence showing that DP_{ext} in a passive construction cannot be bound by items in a higher clause (see (7), (8)), a possible approach might be to assume that DP_{ext} is always locally bound by default existential quantification. However, to the extent that this is empirically correct (recall the quantificational variability effects discussed above), locality of default existential closure for DP_{ext} variables should be derived rather than stipulated: Since other locally unbound variables (including pronouns in *by*-phrases) can easily pick up a binder in a matrix clause, assuming obligatory clause-bound existential

---

23 Höhle (1978, 71-72) and Gunkel (2003, 188) observe that DP_{ext} in a lassen-passive cannot control a PRO subject in an *ohne zu* adjunct infinitive; see (i-a). However, as shown by (i-b), control by DP_{ext} in a standard passive sentence is also impossible in this environment, so there clearly is an independent factor at work.

(i) a. Ich ließ DP_{ext2} ihm auspeitschen [CP ohne PRO_{1/2} einen Laut zu machen ]
   b. Er wurde DP_{ext2} ausgepeitscht [CP ohne PRO_{1/2} einen Laut zu machen ]
quantification for $\text{DP}_{ext}$ amounts to nothing more than a restatement of the facts.

Thus, I would like to conclude that none of the existing analyses of upward inaccessibility of $\text{DP}_{ext}$ in approaches that maintain a basic accessibility is unproblematic. Even more importantly, all the individual accounts fail to recognize the systematic pattern expressed in (17): The individual cases of inaccessibility must be traced back to diverse sources, and the emergence of a uniform behaviour can only be addressed in terms of a conspiracy of independent factors.

Maintaining Inaccessibility

The second basic option is to assume that $\text{DP}_{ext}$ is always inaccessible in German passive constructions – either because passive is not syntactic and $\text{DP}_{ext}$ is never syntactically represented, or because $\text{DP}_{ext}$ is syntactically represented somehow but inaccessible for syntactic operations throughout, for principled reasons; see the references given in footnote 1. On this view, cases where it looks as though $\text{DP}_{ext}$ is in fact accessible – viz., for operations in its m-command domain, as expressed in (23) – are only apparent.

Thus, Schäfer (2012b) and Alexiadou et al. (2015) claim that examples like those in (4), which I have taken to show that Principle A can be satisfied with reflexives and reciprocals by $\text{DP}_{ext}$, are cross-linguistically rare, and show unexpected restrictions and properties. More specifically, among the Germanic languages such examples would seem to be confined to German and Icelandic; and it can be observed that that there seems to be a verb type restriction: Reflexive passives can be found much more often with inherently reflexive and naturally reflexive predicates than with naturally disjoint predicates in corpora. However, independently of whether or not these two claims can ultimately be substantiated by further typological and corpus studies, it seems clear that they do not call into question the existence of the phenomenon as such, which then needs to be accounted for.\textsuperscript{24}

\textsuperscript{24} Also note that the examples in (4) do indeed involve naturally disjoint predicates; also see Schäfer (2012b, 220) on well-formed examples with the verb \text{schneiden} (‘cut’), which also belongs in this class. Furthermore, the example \textit{weil sich gehasst wird} (‘because refl. hated is’), which is starred in Alexiadou et al. (2015, 130) becomes well formed for most speakers if additional (linguistic and non-linguistic) context is provided.

Another example presented in Schäfer (2012b) and Alexiadou et al. (2015) as a challenge for analyses that postulate regular reflexivization via Principle A in (4) takes the form \textit{dass}
Next, one strategy to maintain inaccessibility of $\text{DP}_{\text{ext}}$ with control into adjunct clauses and secondary predicates (cf. (1) and (2), respectively) is to simply deny the reality of the phenomenon. Thus, Williams (2015, ch. 12) concludes for English analogues of examples as in (1) that the syntactic presence of a controller $\text{DP}_{\text{ext}}$ does not have to be postulated. A core argument is that instances of remote, inter-sentential control as in (48a) cannot possibly be accounted for by postulating a local $\text{DP}_{\text{ext}}$ as a controller; on this view, whatever accounts for remote control might the perhaps be extended to local control as in (analogues of) (1).

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

b. Zwei Innenverteidiger wurden verkauft. Das Ziel war, dafür zwei centre-backs were sold the goal was for it einen besseren Linksaußen zu bekommen.

As shown in (48b), remote control also works in German. 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. In addition, it does not seem to be a priori clear that there could not be a DP-internal non-overt controller in the second clause in (48ab). In the same vein, Landau (2013, ch. 6) argues, based on data from English, that control into purpose clauses as in (1) does not require the presence of a controller. The main observations are the following. First, purpose clause infinitives are possible in contexts where the matrix clause looks like it cannot include a $\text{DP}_{\text{ext}}$ controller to begin with since the matrix predicate is an adjective; second, $\text{DP}_{\text{int}}$ (theme) arguments of unaccusative verbs can sometimes license control into purpose clauses; and third, sometimes a matrix $\text{DP}_{\text{int}}$ argument can effect control whereas a matrix $\text{DP}_{\text{ext}}$ argument cannot. However, notwithstanding the issue of whether the evidence from English does in fact justify such far-

*uns/sich von uns gewaschen wird* (‘that us/refl by us washed is’). Here the reflexive $\text{DP}_{\text{int}}$ does not take the first-person form (as it does in active contexts), but rather the third-person form. Assuming for the sake of the argument that such sentences are acceptable in principle (many speakers would seem to reject them), I do not take this to pose a particular problem for an approach envisaging syntactic accessibility of $\text{DP}_{\text{ext}}$; the only thing that it might show is that $\phi$-agreement is a relatively late process, taking place after reflexivization but before eventual PF realization of the reflexive (on which see, e.g., Fischer (2006)).
reaching conclusions, it is worth pointing out that pertinent constructions from German reveal a somewhat different picture. First, $DP_{int}$ of an adjective cannot bring about control into a purpose clause; see (49a). The pattern is identical to that with adjectival passives (i.e., Zustandspassiv constructions), for which there is ample evidence that there is no $DP_{ext}$ present syntactically (and what may at first sight look like marginal cases of optional by-phrases behaves very differently from regular by-phrases; cf. Maienborn (2007, 2011)); see (49b) (vs. (1b)). This suggests that German purpose clause infinitives of the type in (49ab) do indeed require a syntactically present $DP_{ext}$ argument.  

(49) a. ?*Reifen sind rund \[ CP \text{ um } PRO_1 \text{ auf die Felge zu passen } \] 

\text{tires are round in order on the rim to fit}

b. ?*Der Reifen ist aufgepumpt \[ CP \text{ PRO}_1 \text{ um die Fahrt fortzusetzen } \] 

\text{the tire is inflated in order the journey to continue}

Second, in contrast to what may be the case in English, it is impossible to render examples with $DP_{int}$ as the sole possible controller for a purpose clause of the type in (1) felicitous by an appropriate choice of context; thus, (50) (a translation of Landau’s English example) is still not acceptable.

(50) ?*Das Schiff sank \[ CP \text{ um } PRO_1 \text{ die Königin im zweiten Akt zum Mord zu bewegen } \] 

\text{the ship sank in order the queen in the second act to the murder to persuade}

\footnote{Incidentally, Zustandspassiv constructions might lend themselves to an analysis in terms of external Remove (see footnote 9 above). External Remove directly targets XPs in the workspace before they have a chance to enter syntactic structures. An externally removed XPc will therefore never be accessible for operations that are properly syntactic, but it will be accessible for semantic interpretation (by default existential closure, like $DP_{ext}$ arguments in passive derivations that enter the workspace – and stay there permanently – as a consequence of internal Remove). External Remove of $DP_{ext}$ in Zustandspassiv derivations would also imply moving material from the active part of the workspace into the inactive part; see footnote 21. An independent argument for treating Zustandspassiv by external Remove might be that the set of contexts in which it is possible seems to be a proper subset of the contexts in which regular verbal passives are possible, indicating that Zustandspassiv does not have an independent source.}

\footnote{Note also that both examples in (49) become well formed if the infinitival purpose clause is replaced with a finite purpose clause introduced by the complementizer \textit{damit} (‘so that’).}
Finally, as in English, there are cases where DP\textsubscript{int} obligatorily controls into an adjunct clause and control by DP\textsubscript{ext} is in fact ungrammatical; see (51a). As shown in (51b), this effect is contingent on the form of the purpose clause infinitive, subject to a thematic identity requirement. Also, as indicated by (51c) vs. (51d), thematic identity of controller and controllee can only override DP\textsubscript{ext} control if passivization takes place in the matrix, and there is a choice between a deep-structure ‘subject’ (DP\textsubscript{ext} base-generated in Spec\textsubscript{v}) and a surface-structure ‘subject’ (DP\textsubscript{int} bearing nominative); see Růžička (1983), Stechow & Sternefeld (1988).\footnote{Of course, the grammatically legitimate reading in (51d) is possible only to the extent that landlords can be demolished.}

(51)  
\hspace{0.5cm} a. Das Haus\textsubscript{2} wurde DP\textsubscript{ext\textsubscript{1}} geleert [CP um PRO\textsubscript{1/2}  
the house\textsubscript{nom} was emptied in order  
abgerissen zu werden ]  
demolished to be  

\hspace{0.5cm} b. Das Haus\textsubscript{2} wurde DP\textsubscript{ext\textsubscript{1}} geleert [CP um PRO\textsubscript{1/2} die  
the house\textsubscript{nom} was emptied in order the  
Prämie zu kassieren ]  
bonus to collect  

\hspace{0.5cm} c. Die Vermieter\textsubscript{1} leerten das Haus\textsubscript{2} [CP um PRO\textsubscript{1/2}  
the landlords\textsubscript{nom} emptied the house\textsubscript{acc} in order  
die Prämie zu kassieren ]  
the bonus to collect  

\hspace{0.5cm} d. Die Vermieter\textsubscript{1} leerten das Haus\textsubscript{2} [CP um PRO\textsubscript{1/2}  
the landlords\textsubscript{nom} emptied the house\textsubscript{acc} in order  
abgerissen zu werden ]  
demolish to be  

More generally, then, not only is the assumption that examples as in (1) involve control by DP\textsubscript{ext} not called into question; in fact, the ill-formed examples in (49), (50) and (51) provide further evidence for the presence of a syntactically encoded control relation: It is subject to general structure-dependent constraints – control into adjuncts is licensed only by deep or surface subjects, and thematic identity is known to play a role in the latter case.

Given this state of affairs, the remaining strategy to accomodate control
into adjunct clauses and into secondary predicates as in (1) and (2) under an approach that maintains strict syntactic inaccessibility of DP\textsubscript{ext} consists in postulating that control can be brought about in some other way that does not require accessibility of DP\textsubscript{ext}. To evaluate the mechanics and consequences of such an approach, it needs to be clarified first what approaches to passivization look like that envisage strict inaccessibility of the DP\textsubscript{ext}. There are two general options, a lexical one and a syntactic one (cf. the references in footnote 1), but they share a common core: An abstract operator Pass is postulated that applies to a predicate and reduces its arity by one, by existentially binding the highest argument of the predicate; a simple version of Pass is given in (52) (where P is a predicate).

\[(52) \quad \text{Pass: } \lambda P \exists x P(x)\]

Pass can either be viewed as an operator triggering passivization in the lexicon, or as a functional morpheme triggering passivization in the syntax; see, e.g., Bach (1980, 314) and Bruening (2013, 23). According to the latter view, the external argument is represented in the syntax, via Pass (see Alexiadou & Doron (2013)). However, since the working of Pass presupposes that the predicate P has not yet been merged with a regular DP\textsubscript{ext}, and since the external argument variable is existentially bound throughout, it can never be syntactically accessible for operations like c-command. This kind of approach, while well established, is not without inherent problems. One problem that arises under at least the lexical version of the approach is that P in (52) can in principle be an intransitive V, a transitive V, a ditransitive V, a V taking a DP and a PP, and so on. For each of these contexts, a separate entry must then be specified for Pass.\(^{28}\) Second, it is unclear how quantificational variability effects can be accounted for. A third problem that arises throughout concerns by-phrases; in this case the external argument must not be existentially quantified over (the by-phrase can of course itself contain a different quantifier). The standard solution here is to postulate another version of Pass that does not involve existential quantification but essentially amounts to an identity function. However, this raises the question why the two Pass morphemes, with their radically different semantic contributions, never seem to be realized differently.

\(^{28}\)See Büring (2005, 44) for the same problem with lexical operators for reflexivization. Note also that the propagation of separate entries for Pass is not reduced by the purely notational conventions introduced in work like Bruening (2014) or Wunderlich (2015).
in the world’s languages (e.g., by two separate passive morphemes, or two separate passive auxiliaries); it also requires an additional stipulation to ensure that the by-phrase Pass (i.e., the identity function) can never be combined with an active predicate.\textsuperscript{29}

These inherent problems notwithstanding, consider now the options for the evidence suggesting control of DP\textsubscript{ext} into adjunct clauses and secondary predicates (as in (1), (2)) under such an approach. What is required is an approach to control that does not involve empty categories (like PRO) but rather relies on an identification of two variables. For control into complements, this kind of approach is quite standard.\textsuperscript{30} The two variables that need to be identified are (i) an argument of the control verb (typically the object if the verb takes two DP arguments, otherwise the subject), and (ii) the subject of the control verb’s complement. This can be accomplished by postulating a lexical entry for a control verb like \textit{versuchen} (‘try’) as in (53a); here P stands for a property. If the dependent verb is intransitive, like \textit{zu schlafen} (‘to sleep’) in (53b), the two verbs can be combined directly into one complex predicate, where functional application ensures that the sole argument slot of the embedded verb is identified with the external argument of the control verb; see (53c). If, however, the embedded verb is transitive or ditransitive, an additional operation of function composition must take place, which requires long-distance λ-abstraction over the remaining argument slot(s) of the embedded predicate (so as to turn it into a property); cf. (53de), based on transitive \textit{zu lesen} (‘to read’).

(53) a. \textit{versuchen}: λP λx \textsc{try}(x, P(x))

\textsuperscript{29}Stechow (1987, 1992) shows that it is in principle technically possible to maintain a single Pass operator along the lines of (52) if the head of the by-phrase is itself a higher-level operator that combines first with an individual and then a predicate, and that states the identity of the external argument variable that is eventually existentially quantified over via Pass (z in (i)), and the new variable introduced by the preposition that can be quantified over (y in (i)):

(i) \textit{von\textsubscript{pass}}: λyλQ<\textsubscript{ext}>λz . Q(z) & z = y

This approach (a version of which is also developed in Legate (2014, 41)) has a brute force quality; it stipulates the upward accessibility of the external argument in by-phrase environments which should ideally be derivable from the analysis, and it requires construction-specific assumptions (there has to be a special passive \textit{von} that is different from other occurrences of the preposition).

\textsuperscript{30}By now, there is substantial evidence in support of the existence of a structurally encoded non-overt subject (like PRO) in control constructions; see, e.g., Landau (2013). However, at least for the sake of the argument, I will assume that this evidence is not decisive.
b. *zu schlafen: $\lambda y \text{sleep}(y)$

c. *zu schlafen versuchen: $\lambda P \lambda x \text{try}(x,P(x)) \ (\lambda y \text{sleep}(y)) \Rightarrow \lambda x \text{try}(x,\text{sleep}(x))$

d. *zu lesen: $\lambda v \lambda u \text{read}(u,v)$

e. *zu lesen versuchen:

$$\lambda z \ [ \lambda P \lambda x \text{try}(x,P(x)) \ [ \ (\lambda v \lambda u \ (\text{read}(u,v))) \ ] \ (z) \ ] \Rightarrow \lambda z \ [ \lambda x \text{try}(x,[\text{read}(x,z)])]$$

An account of control along these lines can in principle either take place in the lexicon or in the syntax; see Müller, St. (2002), Wurmbrand (2002), Haider (2010), Stiebels (2010) for various proposals for German. In both versions, it can feed passivization, triggered by a Pass operator as in (52). This way, control by a DP_{ext} of a passivized control verb into an infinitival complement of the control verb can be derived in approaches where the external argument of the matrix verb is not syntactically accessible; cf. (3).

Crucially, things are not so straightforward with control into non-complements, as in (1) and (2). In both environments, a lexical approach is excluded: On the one hand, adjunct clauses of the type in (1) clearly do not form complex predicates with a verb; and on the other hand, depictive secondary predicates of the type in (2) cannot be assumed to be combined with the main predicate pre-syntactically either: As shown in some detail in Müller, St. (2002), such secondary predicates can be separated from the main predicate in German in ways that parts of complex predicates formed in the lexicon cannot be (see Haider (2010)). This leaves, as the only remaining possibility, a syntactic approach to both passivization and control, where the argument identification must take place in a lower position than passivization via Pass in (52). However, at this point the problem arises that it is not quite clear how the effect of argument identification that is locally encoded on the control predicate in a lexical entry such as (53a) can be brought about if the two items (the main predicate, and the adjunct clause or secondary predicate) are separate in the syntax. As far as I am aware, all existing solutions to the problem require additional mechanisms without independent justification in order to bring about an identification of the two variables involved (for secondary predicates, cf., e.g., the lexical rule stipulating identity of the argument of the secondary predicate with some member of the argument list of the main predicate in Müller, St. (2002), or the complex semantics of the abstract Dep
operator identifying argument slots of the main and secondary predicates in Bruening (2014)).

More generally, I conclude that maintaining strict inaccessibility of $\text{DP}_{\text{ext}}$ in German passive constructions is not a convincing option in view of the evidence in (1)–(5).

References


Long-Distance Passives by Structure Removal

Gereon Müller

Abstract
In this paper, an analysis of long-distance passives in German is developed according to which these constructions basically emerge from the co-occurrence of (i) passivization and (ii) restructuring in the language. Based on Müller (this volume) and Müller (2017a), I assume that passivization and restructuring both involve an operation of structure removal in the course of the derivation – of an external argument DP in the first case, and of CP and TP layers of an infinitive in the second case. The null hypothesis that I would like to pursue against this background is that a combination of the two structure removal operations gives rise to the intricate properties of long-distance passives in German. A core feature of the analysis is that it does not involve any long-distance relation at any point; argument demotion, case assignment, and morphological realization as passive all take place extremely locally. Another basic property of the structure removal approach, which sets it apart from most other analyses, is that all DP arguments selected by the verbs involved (including in particular external arguments in the embedded and matrix domains) can be assumed to be structurally represented at some point of the derivation.

1. Introduction

Infinitival complements selected by control predicates in German can either show up as complete sentential objects (i.e., CPs), in which case the construction is unequivocally biclausal, or they can lack some properties indicative of CPs, in which case the construction behaves mono-clausally, at least to some extent; this latter type of infinitive is also known as a restructuring infinitive. In the former case, with a full CP level present in the infinitive, passivization can apply to either the matrix verb or the embedded verb, and the resulting sentences are then completely well-behaved; see the examples in (1a) (with passivization of the matrix verb, and the CP infinitive in situ or extraposed) and in (1b) (with passivization of the embedded verb, and again in-situ and extraposed versions of the object infinitive).

Structure Removal, 113–166
Andrew Murphy (ed.)
LINGUISTISCHE ARBEITSBERICHTE 94, Universität Leipzig 2019
(1)  a. (i) dass gestern \( [\text{CP PRO den Traktor zu reparieren}] \) that yesterday \( \text{the tractor}_{\text{acc}} \) to repair
versucht wurde tried was
(ii) dass versucht wurde \( [\text{CP PRO den Traktor zu reparieren}] \) that tried was \( \text{the tractor}_{\text{acc}} \) to repair
b. (i) dass Karl\(_1\) \( [\text{CP PRO\(_1\)} geliebt zu werden] \) versucht that Karl\(_{\text{nom}}\) loved to be tries
(ii) dass Karl\(_1\) versucht \( [\text{CP PRO\(_1\)} geliebt zu werden] \) that Karl\(_{\text{nom}}\) tries loved to be

However, if passivization takes place under restructuring, a long-distance passive construction arises. German long-distance passivization has first been described from the perspective of grammatical theory by Höhle (1978). Since then, it has figured prominently in a number of analyses; cf. Fanselow (1985), Stechow (1992), Haider (1993, 2010), Bayer & Kornfilt (1994), Sabel (1996), Kornfilt (1999), Wöllstein-Leisten (2001), Wurmbrand (2001, 2015a,b), Müller, St. (2002), Sternefeld (2006), and Keine & Bhatt (2016), among others. In long-distance passives in German, an object of the embedded verb is assigned matrix clause nominative and agrees with the matrix verb; passive morphology only shows up on the matrix verb. See (2a) (with the restructuring infinitive in situ) and (2b) (with the restructuring infinitive extraposed, an option that is also known as the third construction; see below).

(2)  a. dass der Traktor zu reparieren versucht wurde that the tractor\(_{\text{nom}}\) to repair tried was
b. dass die Traktoren versucht wurden zu reparieren that the tractors\(_{\text{nom}}\) tried were to repair

(2ab) involve subject control: The external argument of the matrix verb *versuchen* (‘try’) binds the external argument of the embedded infinitival verb. However, as noted by Sabel (1996), Wurmbrand (2001), and Müller, St. (2002), long-distance passives can also show up with matrix verbs that trigger object control. This is shown for the object control restructuring predicate *empfehlen*
Long-Distance Passives by Structure Removal

(‘recommend’) in (3ab), and for the object control restructuring predicate erlauben (‘permit’) in (3c).

(3)   a. dass ihm der Artikel zu lesen empfohlen wurde
      that him\textsubscript{dat} the article\textsubscript{nom} to read recommended was
      
      b. dass ihr der Traktor zu reparieren empfohlen wurde
      that her\textsubscript{dat} the tractor\textsubscript{nom} to repair recommended was
      
      c. dass ihr keine Zeitung zu lesen erlaubt wird
      that her\textsubscript{dat} no journal\textsubscript{nom} to read permitted is

Since the external argument DP of the matrix verb and the external argument DP of the embedded verb (henceforth DP\textsubscript{ext}) remain without overt realization in (2) (subject control) and (3) (object control), the question arises which of the two DP\textsubscript{ext} arguments is subject to passivization (i.e., which is affected by external argument demotion). Note that the object control case clearly signals that this issue is not negligible since it shows that the two DP\textsubscript{ext} arguments do not necessarily have to be co-indexed: For instance, as regards (3a), the one who does the recommendation and the one who is supposed to read the article are evidently different individuals. As a matter of fact, the empirical evidence would seem to be fairly uncontroversial in this respect: There is no genuine long-distance passivization in the sense that passive morphology associated with the matrix verb would indicate demotion of a DP\textsubscript{ext} with the embedded verb; rather, it is exclusively the DP\textsubscript{ext} of the matrix verb associated with passive morphology that is subject to argument demotion. To see this, consider first the sentences in (4). Here, the matrix verb (versuchen) is a control verb permitting restructuring in all three cases, and the embedded predicate is a psych verb in (4ab) (gefallen (‘please’), beeindrucken (‘impress’)), and an unaccusative intransitive verb in (4cd) (sterben (‘die’), einschlafen (‘fall asleep’)).

(4)   a. dass sie [ PRO ihm zu gefallen ] versucht
      that she him to please tries
      
      b. dass sie [ PRO ihn zu beeindrucken ] versucht
      that she him to impress tries
      
      c. dass sie [ PRO zu sterben ] versucht
      that she to die tries
d. dass sie [ PRO einzuschlafen ] versucht
   that she to fall asleep tries

As shown in (5), the embedded verbs in (4) do not easily permit passivization in simple contexts.¹

(5) a. *dass DP_{ext} ihm (von ihr) gefallen wurde
   that him by her pleased was
b. ?*dass DP_{ext} er (von ihr) beeindruckt wird
   that he by her impressed is
c. ?*dass DP_{ext} (von ihr) gestorben wurde
   that by her died was
d. ?*dass DP_{ext} (von ihr) eingeschlafen wird
   that by her died is

Crucially, however, all of the sentences in (4) can be affected by long-distance passivization; see (6).²

¹In the current context, it does not matter what exactly the source of these restrictions on passivization is. On passivization problems with psych verbs in German, as in (5a) and (5b), see Grewendorf (1989), among others. As for unaccusative verbs (cf. (5c) and (5d)), it is worth pointing out that there is in fact some controversy in the literature as to whether verbs like sterben (‘die’) and einschlafen (‘fall asleep’) can marginally undergo passivization in German. Thus, in Růžička (1989), Sternefeld (1995), Kiss (1995), Fanselow (1992), Müller, St. (1999; 2002), and Müller (2018a) it is argued that unaccusative verbs in German cannot undergo passivization (and it is then sometimes postulated that the few cases where it looks as though they can do so after all can and should be analyzed as reinterpretations of unaccusative verbs as unergative verbs, with the original theta-role of the argument affected by passivization acquiring agent-like properties). On the other hand, in Haider (1991), Eisenberg (1999), Primus (2010), and Kiparsky (2013) it is assumed that unaccusative verbs can in principle be passivized. For present purposes, this issue need not be resolved, as long as there is a general consensus that passivization in (5c) and (5d) is at least somewhat marked, and impossible for at least some speakers in the particular environments given here. I will briefly come back to this issue in section 3, though.

²The verb in (4b) takes a direct (accusative) object; so the case change to nominative in (6b) that is indicative of long-distance passives clearly makes this point. In the other three cases ((4a), (4c), and (4d)), there is no accusative object in active environments, so the presence of long-distance passivization can only be detected indirectly here, e.g., via joint topicalization of the two verbs (as in (6a), (6c), and (6d)), which presupposes restructuring (see Haider (1993), among others).
This shows that long-distance passivization affects a DP<sub>ext</sub> of the matrix verb (which needs to be able to undergo passivization for the construction to be legitimate), and not a DP<sub>ext</sub> of the embedded verb (which does not independently need to be able to undergo passivization for long-distance passivization to succeed). More generally, then, it can be concluded that there is no passivization in the embedded clause of a long-distance passive construction; and this directly explains the absence of case morphology on the embedded verb.

Furthermore, long-distance passives in German behave fully regularly as far as by-phrases are concerned. This is already evident from (6); it also emerges from the examples in (7). Note that with an object control verb like empfehlen (‘recommend’), it is the DP<sub>ext</sub> argument of the matrix clause (which is first person in (7b)), and not the DP<sub>ext</sub> argument of the embedded clause (which is third person in (7b)) that is resumed by the by-phrase. Again, this

---

3In contrast, it has sometimes been proposed that a DP<sub>ext</sub> in the embedded clause needs to be able to be directly affected by passivization in the long-distance passive construction. Thus, Pitteroff (2014) claims that German long-distance passives are only possible if the embedded verb as such allows passivization. The argument is based on the contrast between versenken (transitive ‘sink’) in (i-a) and versinken (unaccusative ‘sink’) in (i-b).

(i)  
    a. dass das Schiff zu versenken versucht wurde  
       that the ship<sup.nom</sup> to sink tried was  
    b. *dass das Schiff zu versinken versucht wurde  
       that the ship<sup.nom</sup> to sink tried was

However, independently of the question of marginal passivizability of unaccusative verbs (see footnote 1), (i-b) is straightforwardly excluded as a control failure: A control infinitive must provide an accessible argument controller for a controller of the matrix clause (see, e.g., Stechow & Sternefeld (1988), and below), and there simply is no such DP<sub>ext</sub> in (i-b), where the only argument of the embedded verb is das Schiff (‘the ship’). In line with this, leaving out this overt DP in (i-b) immediately gives rise to well-formedness.
clearly suggests that passivization in long-distance passives exclusively affects the matrix domain, and not the embedded domain.

(7) a. dass der Traktor \text{von mir zu reparieren} versucht wurde
   that the tractor\text{nom} by \text{me} to repair \text{tried} was

   b. dass ihm \text{der Artikel} \text{von mir zu lesen} empfohlen wurde
   that \text{him}_{\text{dat}} \text{the article} by \text{me} to \text{read} \text{recommended} \text{was}

As already noted above, long-distance passivization in German presupposes restructuring, i.e., the lexically determined capacity of certain kinds of matrix verbs to induce transparency of its clausal complement. However, the two concepts are not co-extensive: As observed by Höhle (1978), Wöllstein-Leisten (2001), Sternefeld (2006), and Haider (2010), there are predicates that allow restructuring but do not permit long-distance passivization. Thus, for many speakers, restructuring control verbs like \text{versuchen} (‘try’), \text{vergessen} (‘forget’), and \text{empfehlen} (‘recommend’) permit long-distance passivization, whereas restructuring control verbs like \text{beabsichtigen} (‘intend’) and \text{wünschen} (‘wish’) do not. This is shown for \text{beabsichtigen} (‘intend’) in (8a) (where scrambling of the embedded object \text{den Traktor} (‘the tractor’) to a position in front of the matrix subject \text{keiner} (‘no-one’) signals restructuring; see the next section) vs. (8b) (where long-distance passivization fails).

(8) a. dass den Traktor\textsubscript{1} \text{keiner} \text{zu reparieren} beabsichtigt hat
   that the \text{tractor}_{\text{acc}} \text{no-one}_{\text{nom}} \text{to repair} \text{intended} \text{has}

   b. *dass der Traktor \text{zu reparieren} beabsichtigt wurde
   that the \text{tractor}_{\text{nom}} \text{to repair} \text{intended} \text{was}

There may also be speakers who permit both of these sentences, and speak-

\footnote{As a matter of fact, Höhle (1978) postulates that only \text{versuchen} (‘try’) can trigger long-distance passivization in German, whereas all other control predicates permitting restructuring cannot do so. Höhle already concedes that other speakers may be somewhat more liberal as regards the class of verbs that permit long-distance passivization (even though he ultimately attributes more liberal judgements in this area to autosuggestion); however, it would seem fair to conclude that the vast majority of researchers concerned with long-distance passives in German since his groundbreaking work in the seventies do not follow Höhle in this respect and rather assume that there are several control verbs that permit both restructuring and long-distance passivization in German. All that said, it seems clear that there is substantial variation among speakers concerning which restructuring predicates also permit long-distance passivization; and this variation needs to follow from the theoretical analysis.}
ers who permit none of them, but it seems safe to conclude that there are no speakers with the two grammaticality judgements reversed. In the same vein, Wurmbrand (2015a,b) has argued that cross-linguistically, long-distance passivization and restructuring may occur as independent phenomena, but as a tendency the former presupposes the latter, not vice versa.

In the remainder of this paper, I will come up with a new approach designed to capture these and other core properties of long-distance passivization in German. The primary goal will be to show that the new analysis basically follows from the independently motivated analyses of (i) passivization and (ii) restructuring developed in Müller (this volume) and Müller (2017a), respectively; these analyses both rely on a concept of structure removal in syntax that complements standard structure-building operations via Merge in a minimalist approach to grammar (see Chomsky (2001, 2013)). I will proceed as follows. Section 2 lays out the general principles and mechanisms underlying a principled approach to structure removal in syntax. Sections 3 and 4 then show how passivization and restructuring work under a structure removal approach, based on evidence from German. Finally, in section 5 I return to long-distance passives in German, and I illustrate how the main properties of this phenomenon are accounted for in a structure removal approach.

2. Structure Removal

The basic premise underlyng the approach to structure removal developed in Müller (this volume, 2017a,b, 2018b) is 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. Clearly, the optimal assumption is that if Remove exists as the mirror image of Merge, it is expected to show similar properties and obey identical constraints. The assumptions made about Merge are the following: First, it is a feature-driven operation; Merge is triggered by designated [●F●] features, which are ordered on lexical items, and which are discharged by the operation they trigger (see Heck & Müller (2007), Abels (2012), Stabler (2013), Georgi (2014), Müller (2014), and references cited there). Second, Merge may apply to heads or phrases. This distinction requires a diacritic on the structure-building feature: [●F0●] for heads (0=min), [●F2●] for XPs (2=max). Third, Merge obeys the Strict Cycle
Condition in (9) (see Chomsky (1973, 1995, 2001, 2008)), according to which an operation cannot apply at a given stage of the derivation that exclusively affects an embedded structure. And finally, Merge can be external or internal; it brings about basic structure building in the former case, and movement in the second.

(9) **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 \( \alpha \).

Exactly the same assumptions are made about Remove. Thus, first, Remove is feature-driven. It is triggered by designated \([-F-] \) features, which are ordered on lexical items. Second, Remove may apply to heads or phrases: \([-F_0-] \), \([-F_2-]\). Third, Remove obeys the Strict Cycle Condition. Fourth, Remove can be external or internal; but only internal Remove (where the operation affects material of the current tree) will be relevant in what follows.

Turning to removal of phrases first, a head \( X \) that is equipped with a feature \([-Y_2-] \) can remove a full YP subtree that is either its complement or its specifier (more deeply embedded phrases are not accessible, because of the Strict Cycle Condition). Focussing on removal of specifiers here (since this will turn out to be the scenario that is relevant for passivization), the abstract derivation in (10) illustrates how \( X_{[\bullet Y_2 \bullet]} > [\bullet Y_2 \bullet] \) first triggers Merge with a YP as its specifier in (10a) (via the structure-building feature \( [\bullet Y_2 \bullet] \) which is at the top of the list of ordered features triggering operations after \( X \) has discharged the initially top-most structure-building feature \( [\bullet U_2 \bullet] \) for its complement), and subsequently removes that specifier YP again in (10b) (via the structure-removal feature \( [-Y_2-] \) that is at the top of the list after discharge of \( [\bullet Y_2 \bullet] \)).

---

5 The concept of domain of a head \( X \) is taken over from Chomsky (1995) here; it is the set of nodes dominated by XP that are distinct from and do not contain \( X \).

6 External Remove affects material that is not present in the current tree. In Müller (this volume, 2017b), I suggest that external Remove affects an XP in the workspace of the derivation; the external argument of adjectival passives (‘Zustandspassiv’) is a candidate for this.
(10) \textit{Remove and phrases: specifiers}

\textbf{a. Merge}(X'[●Y₂●]>[-Y₂−], YP):

```
  XP
 /\       /
 YP ┌─ ZP ┌─ Y' ┌─ X' ┌─ UP
  \  \    \  \    \  \  
   ↓   ↓    ↓   ↓    ↓   ↓
      Y       WP          
```

\textbf{b. Remove}(X'[−Y₂−], YP):

```
  XP
   /\     /
  X   UP
```

The sequence of structure building followed by structure removal applying to one and the same item (YP in (10)) qualifies as a Duke-of-York derivation, where an operation is first carried out and subsequently undone again (see Pullum (1976), McCarthy (2003)): Even though it might at first sight seem to be entirely redundant, it may in fact have non-trivial consequences, given that other operations (triggered by other features not listed here, on X and possibly elsewhere) can in principle be interspersed with the structure-building and structure-removal operations.

Next, \textit{Remove} can also apply to heads, based on a feature [−F₀−]. Focussing on removal of heads of complements here (since this is the context that will be relevant for restructuring), (11) shows how a head X may first trigger \textit{Merge} with a YP complement (via its structure-building feature [●Y₂●]), and then induce \textit{Remove} of the head Y of its complement (via its structure-removal feature [−Y₀−]). [−Y₀−] on X removes the head Y of YP, and consequently, it takes away the YP projection as well, but only this: More deeply embedded material is attached to the head responsible for removal and replaces the original item (YP) in a way that maximally preserves the previous c-command and linearization relations within YP. As a consequence, an original complement of Y (WP in (11a)) becomes the new complement of X (see (11b)), and if Y
originally had a specifier in addition (ZP in (11a)), this specifier is reassociated as a new specifier of X (see (11b)). This latter effect is remarkable insofar as it opens up the possibility of dislocation (here: of ZP) without actual movement taking place; as it turns out, this mechanism will be shown to play a vital role in the analysis of long-distance passivization to be developed in section 5 below. For now, it can be concluded that removal of heads has the effect of cutting out a top-most layer of structure and reintegrating lower material in an order-preserving way.\footnote{For predecessors and alternative versions of essentially this concept of removal of heads, see Ross (1967, ch. 3), Chomsky (1981, 2015), and Pesetsky (2016), among several others.}

(11) Remove and heads: complements with specifiers

a. Merge\( (X_{[●\⇒\{−Y_0−\}]} \bowtie YP) \):

\[
\begin{array}{c}
X' \\
X_{[−Y_0−]} \\
\quad ZP \\
\quad \quad Y' \\
\quad \quad \quad WP
\end{array}
\]

b. Remove\( (X_{[−Y_0−]} \bowtie Y) \):

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

Like removal of phrases, removal of heads gives rise to Duke-of-York derivations where the successive application of structure-building and structure-removal operations makes the derivation revert to a previous stage; however, as before, these undoings of earlier structure building are by no means innocuous since other operations can be interspersed with, say, introduction of the
YP layer in a step preceding (11a), and subsequent removal of the YP layer in (11b).

A restriction on Remove operations is that they must not violate the general recoverability requirement for derivations. Thus, in principle, feature-driven structure removal may affect any XP subtree (via [–X₂–]), and any XP layer (via [–X₀–]); however, if such a removal has the effect that important information gets lost that is not recoverable from the syntactic context in some way, the derivation will invariably fail.

From a more general point of view, the postulation of an elementary Remove operation makes it possible to account for phenomena where it looks as though conflicting structure assignments are necessary in the syntax – i.e., where the applicability or inapplicability of some syntactic operations indicates the presence of some item α, and the applicability or inapplicability of other syntactic operations indicates that α is in fact not present in the structure. These kinds of phenomena are numerous in natural language syntax, and they have often been treated by recourse to some unrestricted concept of reanalysis. In this sense, Remove-based structure removal can be viewed as the core of a principled approach to reanalysis.

In the following two sections, I will show how structure removal affects the external argument DP in German passive derivations and the CP and TP layers of infinitival clauses in German restructuring derivations.

3. Passivization

A central assumption of the approach to passivization in German developed in Müller (this volume) is that the early stages of derivations of active sentences and corresponding passive sentences are fully identical. In particular, after v has undergone Merge with a VP (which may or may not contain an internal argument (DP_{int})), triggered by a [•V₂•] feature, an external argument (DP_{ext}) is introduced as a specifier of v in both cases, triggered by a [•D₂•] feature which comes next on the list. The basic difference between an active v and a passive v is that the latter functional head is then equipped with a [–D₂–] feature in addition, which brings about a Remove operation that takes DP_{ext} out of the structure again. This accounts for the argument demotion effect of the passive. Absorption of structural accusative case can then be triggered as a side effect, either by assuming a structural case assignment feature on v
not to show up in the presence of $[-D_2-]$ on the list, or by assuming a dependent case analysis (see Stiebels (2000), McFadden (2004), Preminger (2014), Baker (2015), Bobaljik (2015), among others) according to which dependent accusative is assigned in the presence of $DP_{ext}$, and such case assignment is determined after removal of $DP_{ext}$. In either approach, $DP_{int}$ will be assigned structural nominative case by $T$, which is freed up because there is no $DP_{ext}$ in the structure anymore that would depend on it.\(^8\) Next, the specific morphological reflex of passivization (a combination of the passive auxiliary \textit{werden} and the past participle form of $V$, in the regular, primary passive construction in German) results from $[-D_2-]$ on $v$, which determines morphological realization even though it is discharged from the list of features on $v$ that trigger morphological realization.\(^9\)

Against this background, consider a simple German passive construction based on a transitive verb, as in (12).

(12) dass das Buch gelesen wurde
that the book\textit{nom} read was

According to present assumptions, the relevant part of the Remove-based derivation of (12) looks as in (13). First, $v$ is merged with the VP \textit{das Buch gelesen} (triggered by $[\bullet V_2 \bullet]$ at the top of the feature list of $v$), as in (13a). Next, $v$ is merged with $DP_{ext}$; see (13b). And ultimately, $DP_{ext}$ is removed again in (13d).

(13) a. $\textit{Merge}(v[\bullet V_2 \bullet] \rightarrow [\bullet D_2 \bullet] \rightarrow [-D_2-], VP)$
$[v' \textit{v}[\bullet D_2 \bullet] \rightarrow [-D_2-] [\textit{VP das Buch gelesen }])$

b. $\textit{Merge}(v[\bullet D_2 \bullet] \rightarrow [-D_2-], DP_{ext})$
$[\textit{VP DP_{ext} v}[-D_2-] [\textit{VP das Buch gelesen }])$

c. $[\ldots]$

\(^8\)In German, structural nominative case assignment by $T$ is not accompanied by an \textit{obligatory} EPP feature on $T$ (see below on \textit{optional} EPP features on $T$), so $DP_{int}$ can in principle stay in its in situ position in the VP (or undergo local scrambling within $vP$); see den Besten (1981), Grewendorf (1989), Haider (2010).

\(^9\)If morphological realization ist post-syntactic, this implies that discharged features, although inactive, are still accessible in some form for morphological realization; see Adger (2003) for discussion of this much more general issue, and Chomsky (1995) for a distinction between erasure and deletion which addresses a similar challenge. Alternatively, if morphological realization is pre-syntactic, no such issues arise in the first place.
d. \( \text{Remove}(v[\neg D_2], DP_{\text{ext}}) \) \\
\[ \begin{array}{c}
vP \quad vP \quad \text{das Buch gelesen}\end{array} \]

An immediate consequence of the Remove-based approach to passivization is that passivization is always an extremely local operation: Due to the Strict Cycle Condition in (9), a \([-D_2] \) feature on \( v \) can target a \( DP_{\text{ext}} \) in Spec\( v \), but it can never target a \( DP_{\text{int}} \) that is a daughter of VP; thus, a general prohibition against passivization of unaccusative verbs can be derived (and for reasons laid out in Müller (2018a), I take this prediction to be empirically correct for German; see footnote 1 above). Clearly, a \([-D_2] \) feature on a matrix \( v \) can then also never affect a \( DP \) that is contained in an embedded \( vP \) or VP; this straightforwardly derives an impossibility of true long-distance passivization; recall from section 1 that this consequence is confirmed by the empirical evidence.

The question arises of how this approach models resuming \( DP_{\text{ext}} \) with a \textit{by}-phrase. Here the issue becomes relevant what actually happens to \( DP_{\text{ext}} \) once it is removed from the structure. The suggestion in Müller (this volume) is that structure removal via \([-D_2] \) places \( DP_{\text{ext}} \) in the workspace of the derivation. At this point, two options arise. The first one is that \( DP_{\text{ext}} \) simply stays in the workspace. In accordance with the general recoverability restriction on structure removal (see section 2), it then triggers default existential quantification; for this is the only failsafe interpretation procedure that be carried out with non-overt arguments. Still, there is also a second option, and this consists in remerging \( DP_{\text{ext}} \) into the structure in the only way that is available without structure-building features, viz., as an adjunct. This is the source of the \textit{by}-phrase in sentences like (14).

(14) dass [PP von Karl] das Buch gelesen wurde  
that by Karl the book read was

Finally, and most importantly, this approach makes clear predictions with respect to the accessibility of \( DP_{\text{ext}} \) for syntactic operations. Depending on which phenomena one looks at, it seems that there is both evidence for the syntactic presence of \( DP_{\text{ext}} \) in German passive constructions, and evidence against it. Standardly, however, approaches to passivization have to either postulate the presence of a (non-overt) \( DP_{\text{ext}} \) in the structure (see, e.g., Baker, Johnson & Roberts (1989), Sternefeld (1995), Harley (2013), Merchant (2013), Collins (2005)), or they have to postulate that no such \( DP_{\text{ext}} \) shows up in the
syntax (see Chomsky (1981), Müller, St. (2007), Kiparsky (2013), Bruening (2013), Schäfer (2012), Alexiadou & Doron (2013), Legate (2014), and Alexiadou, Anagnostopoulou & Schäfer (2015), among others). In contrast, an approach to passivization in terms of structure removal makes it possible to accommodate empirical evidence both for and against $\text{DP}_{\text{ext}}$ in syntactic representations in the passive. More specifically, the prediction is that $\text{DP}_{\text{ext}}$ should be accessible for syntactic operations that involve other material contained within the same vP (provided general locality requirements are met), because $\text{DP}_{\text{ext}}$ is still part of the structure at this point (these operations are the ones that can take place at the stage (13c) of passive derivations, i.e., in the narrow window when $\text{DP}_{\text{ext}}$ has been merged and has not yet been removed); and $\text{DP}_{\text{ext}}$ should be inaccessible for syntactic operations that involve other material outside of the same vP, because $\text{DP}_{\text{ext}}$ has been removed from the structure once the derivation has proceeded beyond vP (this latter consequence is due to the Strict Cycle Condition; see above).

In Müller (this volume), it is argued that this prediction is correct for passivization in German. On the one hand, $\text{DP}_{\text{ext}}$ exhibits downward accessibility, i.e., it can participate in operations like control into adjunct clauses, complement clauses, and secondary predicates, binding of reflexive and reciprocal pronouns (via Principle A of the binding theory), disjoint reference enforcement with non-pronominal DPs (via Principle C of the binding theory), and unselective binding by adverbs of quantification. This is shown for control into non-finite purpose clauses in (15a) (see Roberts (1987)), and for control into non-finite complement clauses in (15b) (see van Urk (2013)); in both cases, $\text{DP}_{\text{ext}}$ can undergo Agree with the PRO$_1$ subject of the embedded infinitive (and thus establish a binding relation, assuming an approach to control along the lines of Landau (2013)).

(15) a. Der Reifen wurde $\text{DP}_{\text{ext1}}$ aufgepumpt [CP PRO$_1$ um die Fahrt fortzusetzen ]

b. Es wurde $\text{DP}_{\text{ext1}}$ versucht [CP PRO$_1$ zu schlafen ]

The same effect obtains with control into subject-oriented secondary predicates; see, e.g., (16ab) (cf. Müller, St. (2002)).
Long-Distance Passives by Structure Removal

Next, the data in (17) signal downward accessibility of DP\textsubscript{ext} for binding of reflexive pronouns (see (17a)) and binding of reciprocal pronouns (see (17b)) in the passive; again, assuming Principle A satisfaction to involve an Agree operation in the syntax (see, e.g., Reuland (2001, 2011), Fischer (2004), and Hicks (2009)), this follows straightforwardly from the presence of DP\textsubscript{ext} at the relevant stage of the derivation (i.e., before the vP that DP\textsubscript{ext} is the specifier of is completed, and the derivation moves on to the next cyclic domain).

On the other hand, the Remove-based approach to passivization in German predicts upward inaccessibility of DP\textsubscript{ext}, i.e., inaccessibility of DP\textsubscript{ext} for operations that involve other material in vP-external positions. As argued in Müller (this volume), this is corroborated by the non-availability of variable binding from above, failed attempts at control of DP\textsubscript{ext} from outside, inability of DP\textsubscript{ext} to satisfy movement criteria, lack of minimality effects, and transparency for reflexivization. Let me just illustrate upward inaccessibility of DP\textsubscript{ext} on the basis of evidence from variable binding and control here. As shown in (18a), DP\textsubscript{ext} cannot be bound by a quantified DP in the matrix clause (see Alexiadou et al. (2015)), and this follows directly under the assumption that DP\textsubscript{ext} is not present anymore in the structure at the point where the matrix DP is merged. In contrast, if DP\textsubscript{ext} returns into the structure as a by-phrase after it has undergone removal, such binding is unproblematic; see (18b).

\begin{enumerate}
  \item[(16)]
    \begin{enumerate}
      \item Die Daten wurden DP\textsubscript{ext} \textsubscript{1} [SC PRO\textsubscript{1} nackt ] analysiert \hfill the data were naked analyzed
      \item Das Handout wurde DP\textsubscript{ext} \textsubscript{1} [SC PRO\textsubscript{1} übermüdet ] verfasst \hfill the handout was tired written
    \end{enumerate}
  \end{enumerate}

\begin{enumerate}
  \item[(17)]
    \begin{enumerate}
      \item Hier wurde DP\textsubscript{ext} \textsubscript{1} sich\textsubscript{1} nicht geprügelt \hfill here was \textbf{REFL not} hit
      \item Es wurde DP\textsubscript{ext} \textsubscript{1} einander\textsubscript{1} gedankt \hfill it was \textbf{RECIPI} thanked
    \end{enumerate}
  \end{enumerate}

\begin{enumerate}
  \item[(18)]
    \begin{enumerate}
      \item *Kein Student\textsubscript{1} gibt zu [\textsubscript{CP} dass DP\textsubscript{ext} \textsubscript{1} schlecht gearbeitet \hfill no student admits that badly worked wurde ] \hfill was
    \end{enumerate}
  \end{enumerate}
Interestingly, $\text{DP}_{\text{ext}}$ in a passive construction can be bound by an adverb of quantification, giving rise to the quantificational variability effect (see Heim (1982)). As observed in Alexiadou & Müller (2018), an adverb of quantification like \textit{größtenteils} (‘for the most part’) can bind $\text{DP}_{\text{ext}}$ in the passive in German; see (19a). This is expected if the adverb of quantification can indeed be assumed to be part of the minimal vP projection that also contains $\text{DP}_{\text{ext}}$ – in this case, Remove can apply after the binding relation is established; i.e., Remove counter-bleeds variable binding. In contrast, the same adverb cannot effect binding of a $\text{DP}_{\text{ext}}$ variable if it shows up in the matrix clause; see (19b). Here removal of $\text{DP}_{\text{ext}}$ in the embedded clause must have taken place before variable binding by a matrix clause quantifier can be established; i.e., Remove transparently bleeds variable binding.\(^\text{10}\)

\begin{enumerate}
\item \textit{Es wurde größtenteils} \textit{DP}_{\text{ext}} \textit{geschlafen beim Vortrag}
\item \textit{Es war größtenteils} \textit{so dass} \textit{DP}_{\text{ext}} \textit{geschlafen wurde}
\end{enumerate}

(19) a. \textit{Most people slept through the talk.}

b. \textit{*Most of the time, people slept through the talk.}

\(^{10}\text{As indicated in the glosses, (19b) can only have a reading where quantification is over time spans or situations, not over individuals, as it is possible in (19a) (where the other reading is available, too). Note also that there is independent evidence for a very low, vP-internal position of the adverb of quantification \textit{größtenteils} (‘for the most part’); as shown in (i-a), it can undergo topicalization together with the verb, which higher adverbs in German cannot do (see (i-b)).}\)

\begin{enumerate}
\item \textit{[vP\text{Größtenteils} DP}_{\text{ext}} \textit{geschlafen}] wurde beim Vortrag \text{t\textsubscript{1}}.
\item \textit{*[TP Wahrscheinlich DP}_{\text{ext}} \textit{geschlafen}] wurde beim Vortrag.
\end{enumerate}

\begin{enumerate}
\item \textit{‘Most people slept through the talk.’}
\item \textit{‘People probably slept through the talk.’}
\end{enumerate}
A second argument for the inaccessibility of $\text{DP}_{\text{ext}}$ for items merged in higher projections comes from control infinitives. As noted by Stechow & Sternefeld (1988), control infinitives must have an accessible subject argument. The empirical evidence shows that $\text{DP}_{\text{ext}}$ in a passivized control infinitive does not qualify as an accessible argument; see (20ab). This is accounted for if $\text{DP}_{\text{ext}}$ has undergone removal before the legitimacy of the control infinitive (i.e., the accessibility of an external argument) is checked.

(20)  

\begin{align*} 
&\text{a.} \quad \text{*Er versucht } [\text{CP } \text{DP}_{\text{ext}} \text{ gearbeitet zu werden } ] \\
&\text{he tries } \text{worked } \text{ to be} \\
&\text{b.} \quad \text{*weil } [\text{CP bald } \text{DP}_{\text{ext}} \text{ geschlafen zu werden } ] \text{ gewünscht wird} \\
&\text{because } \text{soon } \text{slept } \text{ to be } \text{ wished } \text{is}
\end{align*}

To sum up, a Remove-based approach to passivization in German provides a principled approach to the conundrum created by conflicting evidence for the presence of $\text{DP}_{\text{ext}}$ in the structure: At first, $\text{DP}_{\text{ext}}$ is present (it is merged as Spec$v$ like any other external argument in active clauses), and can accordingly be involved in various syntactic operations, but subsequently, it is removed, and can therefore not be accessed anymore by syntactic operations applying at a later stage.

4. Restructuring

According to the approach to restructuring infinitives in German developed in Müller (2017a), control verbs uniformly take CP complements.\textsuperscript{11} Against this background, the special property of restructuring predicates is assumed to be the capacity to bring about removal of the CP and TP layers of the sentential complement, which gives rise to derived vP complements. Such a removal is triggered by $[-C_0-]$ and $[-T_0-]$ features that show up on the matrix V. These $[-C_0-]$ and $[-T_0-]$ features for structure removal applying to heads are taken to be present on a restructuring control V only optionally. The reason is that all control verbs that permit restructuring also permit a full clausal infinitival complement, i.e., they are all compatible with restructuring not taking place.\textsuperscript{12}

\textsuperscript{11}The complement of a control verb is typically non-finite; but see Stiebels (2010) on control into finite clauses in German.

\textsuperscript{12}It is a priori unclear why CP and TP layers both have to be removed with control predicates that induce restructuring; but one may speculate that the concept of a phase (cf. Chomsky
Unlike restructuring control verbs, other verbs that trigger restructuring obligatorily are assumed to take smaller complements from the start; i.e., there is no structure removal involved here. This holds, e.g., for modal and raising verbs; more generally, for so-called functional restructuring predicates (see Wurmbrand (2001, 2015b)).

To see how structure removal in restructuring infinitives works, consider (21), where the matrix verb *versuchen* (‘try’) is equipped with the optional features [–C] and [–T].

(21) dass sie_1 [ PRO_1 ihn zu küssen ] versucht
    that she nom him acc to kiss tries

Suppose that the embedded T head has undergone [•v_2•]-driven Merge with the embedded vP, projecting TP; and that subsequently C has triggered Merge (via [•T_2•]) with the TP, projecting CP. In the next step in (22a), matrix V (*versucht*) is merged with CP, triggered by [•C_2•] on V. In (22b), the CP layer is removed again, as a consequence of [–C_0–] on V. Thus, TP becomes the new complement of V, and while TP was not accessible for removal of its head as long as it was embedded in a CP (due to the Strict Cycle Condition), it is accessible now: Head removal can apply recursively. Accordingly, given the remaining [–T_0–] feature on V, the TP layer is removed at this point, and vP becomes the next (and final) complement of V; see (22c).  

(22) a. \( \text{Merge}(V,CP): \)
    \[
    [V \text{ [CP C [TP [vP PRO ihn zu küssen ] T ]]} [V \text{ versucht }][–C_0–]>[–T_0–]]
    \]

(2001, 2008, 2014), which postulates an intimate relation between C and T, might play a role. That said, it is not in fact entirely straightforward in the first place to detect differences between a theory where restructuring verbs embed a TP, and a theory where restructuring verbs embed a vP. Haider (1993, 2010), e.g., holds that there is no distinction between these projections to begin with. However, see Wurmbrand (2001, 2007, 2015b) for relevant empirical evidence; in addition, I will rely on an argument for the presence of a separate TP in German later in the present section, and then also in the following section, when I turn to long-distance passivization again.

\(^{13}\)Note that the Strict Cycle Condition thus ensures that the order of features for removal of functional layers need not be stipulated; if V exhibits an initial order [–T_0–]>[–C_0–] instead of [–C_0–]>[–T_0–], the derivation will crash because [–T_0–] cannot be discharged (removing the TP layer) while CP is still present; but [–T_0–] at the top of the list will block the discharge of [–C_0–].
b. \( \text{Remove}(V, CP) \):
\[
[\text{VP} \ [\text{TP} \ [\text{VP} \ PRO \ ihn \ zu \ küssen] \ T]] \ [v \ versucht \ [-T_{0} -]]
\]
c. \( \text{Remove}(V, TP) \):
\[
[\text{VP} \ [\text{VP} \ PRO \ ihn \ zu \ küssen] \ [v \ versucht]]
\]

As with passivization, the Remove-based approach to restructuring in German is able to reconcile conflicting evidence regarding the accessibility of CP and TP layers in restructuring configurations. Thus, there is evidence both for a biclausal approach to restructuring constructions (with CP status of the complement), and for a monoclausal approach (where the complement is not a full CP but a smaller structure). However, standard approaches have to either postulate a CP for restructuring complements throughout (see, e.g., Baker (1988), Sternefeld (1990), Müller & Sternefeld (1995), Sabel (1996), Koopman & Szabolcsi (2000)), which makes it difficult to account for the evidence for monoclausality, or they have to postulate that there is always less structure in restructuring complements (see, e.g., Haider (1993, 2010), Kiss (1995), Wurmbrand (2001, 2007, 2015b), Sternefeld (2006)), among many others), which then poses problems with regard to evidence for biclausality. In contrast, an approach in terms of structure removal makes it possible to have one’s cake and eat it, since it can in principle account both for evidence in support of CP and TP layers of restructuring infinitives, and for conflicting evidence that suggests a smaller (vP) structure.\(^{14}\) For concreteness, CP and TP are predicted to be accessible for syntactic operations that involve lower parts of the structure (i.e., material within the infinitive), and inaccessible once matrix V has brought about structure removal, as in (22bc).

In Müller (2017a), it is argued that this prediction is correct for restructuring infinitives in German. On the one hand, there are many phenomena where the simplest account would seem to clearly favour a monoclausal approach where there is no CP (or TP) layer present. Most of these phenomena are well known and have been widely discussed in the literature (see Stechow

\(^{14}\)There are predecessors which postulate that in restructuring configurations, a structure that is initially biclausal is reduced to a structure that is monoclausal, via some form of structure removal; see Ross (1967, ch. 3), Rizzi (1982), Aissen & Perlmutter (1983), and Stechow & Sternefeld (1988). The problem with all these approaches is that they rely on transformations that are (i) ad hoc, (ii) not constrained in interesting ways, and (iii) not embedded in a general system of elementary, primitive operations manipulating syntactic structure. Thus, the present analysis based on the elementary operation Remove can be viewed as an attempt to come up with a principled version of these earlier approaches to restructuring.
& Sternefeld (1988), Grewendorf (1988), Fanselow (1991), Bayer & Kornfilt (1994), Wurmbrand (2001), and Haider (2010), among others). Among them are the clause-boundedness of scrambling and unstressed pronoun fronting, the clause-boundedness of extraposition, the clause-boundedness of multiple sluicing, the compactness of verb clusters, intonational breaks, and wide scope of negation. Let me just address two of these phenomena here.

First, it is an old observation going back to Ross (1967) that (optional) scrambling is strictly clause-bound in German. In a minimalist approach to syntax, this means that scrambling cannot leave a CP; see (23a). The same goes for (obligatory) fronting of unstressed pronouns; see (23b).

(23) a. *dass den Fritz_{1} keiner gesagt hat [CP dass wir_{1} t_{1} that the Fritz_{acc} no-one_{nom} said has that we_{nom} einladen sollen ] invite should
b. *dass die Maria es_{1} meinte [CP solle man_{1} t_{1} lesen ] that the Maria_{nom} it_{acc} said should one_{nom} read

If a control verb embedding an infinitival clause does not permit restructuring (as a lexical property), scrambling and unstressed pronoun fronting from the infinitive to a position which is clearly part of the matrix clause are impossible throughout. With verbs like auffordern (‘request’) and leugnen (‘deny’), this is the case for many speakers of German (though not for all – there is quite a bit of low-level variation between individual speakers in this domain). The effect is shown for the two movement operations in (24a) and (24b); the illformedness of these examples can be derived in exactly the same way as with finite clauses as in (23), as a result of the presence of a CP layer.

(24) a. *dass den Fritz_{1} keiner die Maria [CP t_{1} zu küssen ] that the Fritz_{acc} no-one_{nom} the Maria_{acc} to kiss aufforderte requested

\[^{15}\text{For present purposes, it does not matter why exactly CP is a barrier for movement operations targeting a low position in the matrix clause (like scrambling and unstressed pronoun fronting, which can both be assumed to end up in a specifier position of v) but not for movement operations targeting a higher position in the matrix clause (like wh-movement and topicalization). See Müller (2014, ch. 2) for a recent overview of some of the options.}\]
Long-Distance Passives by Structure Removal

b. *dass die Maria es₁ gestern \([\text{CP } t₁ zu kennen }]\) geleugnet
that the Maria\(_{nom}\) it\(_{acc}\) yesterday to know denied
hat
has

In contrast, if a control verb permits restructuring, as with *versuchen* (‘try’) and *empfehlen* (‘recommend’), both scrambling (as in (25a)) and unstressed pronoun fronting (as in (25b)) are possible. This follows, given that structure removal can have removed the CP and TP layers of the infinitival complement before extraction from the complement takes place (which is a vP at the relevant stage of the derivation).

(25) a. dass den Fritz\(_{1}\) keiner \([\text{vP } t₁ zu küssen }]\) versuchte
that the Fritz\(_{acc}\) no-one\(_{nom}\) to kiss tried
b. dass die Maria es₁ ihm gestern \([\text{vP } t₁ zu lesen }]\)
that the Maria\(_{nom}\) it\(_{acc}\) him\(_{dat}\) yesterday to read
empfohlen hat
has

As a second illustration of monoclausal properties of restructuring, consider the scope of clausal negation. If negation shows up in an embedded finite CP, it can normally not take wide scope over the matrix clause, and it seems plausible to trace this restriction to the presence of a CP barrier for scope taking. However, with predicates like *empfehlen* (‘recommend’), a negation *nicht* takes wide scope over the matrix clause if restructuring applies; see (26a). In contrast, if the matrix control verb does not permit restructuring (as with *auffordern* (‘request’), for many speakers), wide scope of negation is impossible; see (26b).

(26) a. dass Maria ihm \([ \text{das Buch nicht zu lesen } ]\)
that Maria\(_{nom}\) him\(_{dat}\) the book\(_{acc}\) not to read
empfiehlt
recommends

\(\textit{recommend} \gg \textit{not}, \textit{not} \gg \textit{recommend}\)

b. dass Maria ihn \([ \text{CP das Buch nicht zu lesen } ]\)
that Maria\(_{nom}\) him\(_{acc}\) the book\(_{acc}\) not to read
auffordert
requests
Given that scope of negation is determined at a late stage of the derivation, when an intervening CP barrier for scope taking has long been removed in restructuring environments, these data are also straightforwardly accounted for.

On the other hand, there are a number of phenomena whose account crucially involves the presence of a CP. Among these are uniformity of embedding,\footnote{Thus, as noted above, all control verbs embedding non-finite complements can be assumed to uniformly subcategorize for CPs. Also, assuming, as before, that \([-C_0-]\) and \([-T_0-]\) can only ever be \emph{optionally} present on V, the non-existence of control verbs that would obligatorily require restructuring, such as the fictive verb *entsuchen (which might also mean ‘try’, like \emph{versuchen}, or something else) in (i), can be derived; see Koster (1987), Stechow & Sternefeld (1988).} licensing of PRO by C (see Adger (2003)), absence of new binding domains, unstressed pronoun fronting, and the very existence of what is known as the third construction. Focussing on the just the latter two phenomena here, consider first the evidence from unstressed pronoun fronting. I have just invoked this movement type in order to develop an argument for a small, CP-less structure of an embedded infinitival complement. Interestingly, the same movement type can also be used to provide an argument for a larger structure involving a CP.

Unstressed pronouns must undergo fronting to a position that can only be preceded by a subject DP (but does not in fact have to be preceded by it), not by non-pronominal object DPs or PPs. This is illustrated in (27). In (27a), the unstressed direct object pronoun \textit{es} (‘it’) precedes the subject DP (\textit{die Maria}) and the indirect object DP (\textit{dem Fritz}). In (27b), the subject DP precedes the unstressed direct object pronoun, which in turn precedes the indirect object DP (deviating from what would be the basic word order with two non-pronominal object DPs with verbs like \textit{geben} (‘give’)). In (27c) and (27d), two failed attempts at having other material (the indirect object \textit{dem Fritz} and an adverb \textit{wahrscheinlich} (‘probably’), respectively) precede the un-
stressed pronoun are documented; the illformedness of these examples shows that unstressed pronoun fronting is an obligatory operation in German.

(27) a. dass es₁ die Maria dem Fritz t₁ gegeben hat
     that it\textsubscript{acc} the Maria\textsubscript{nom} the Fritz\textsubscript{dat} given has

b. dass die Maria es₁ dem Fritz t₁ gegeben hat
     that the Maria\textsubscript{nom} it\textsubscript{acc} the Fritz\textsubscript{dat} given has

c. *dass die Maria dem Fritz es₁ gegeben hat
     that the Maria\textsubscript{nom} the Fritz\textsubscript{dat} it\textsubscript{acc} given has

d. *dass die Maria wahrscheinlich es₁ dem Fritz t₁ gegeben
     that the Maria\textsubscript{nom} probably it\textsubscript{acc} the Fritz\textsubscript{dat} given
     has

The effect in (27ab) can be taken to indicate that subject DPs have an option to move to a vP-external position that object DPs and PPs do not have. A candidate that clearly suggests itself here is optional EPP-driven movement to Spec\textsubscript{T}, and this is indeed the conclusion drawn in Müller (2001) and Fanselow (2004), which I take to be correct.\textsuperscript{17} The position for fronted unstressed pronouns can therefore be assumed to be a specifier position at the left edge of vP; here the unstressed pronoun precedes other object arguments (including scrambled ones), adverbials, and the base position of subjects. Next, even though the landing site of unstressed pronoun fronting is the left edge of vP, the empirical evidence suggests that a CP is required for unstressed pronoun fronting in German – it acts as a barrier blocking the movement as such if it intervenes in the movement chain (see above), but it is necessary to license the target position of the movement from above. Thus, complements of non-control (i.e., obligatory) restructuring verbs, for which there is no reason whatsoever to postulate a CP layer, do not have sufficient space for unstressed pronoun fronting, and ungrammaticality arises if the unstressed pronoun does not show up in the matrix vP domain (which is licensed by a C). This is illustrated for the perfect auxiliary hat (‘has’) in (28a), for the raising verb schien (‘seemed’) in (28b), and for the causative exceptional case marking verb ließ (‘let’) in (28c).

\textsuperscript{17}This, then, also provides an argument for the existence of a TP projection in German; see above.
(28)  a. *dass sie mir₁ schon letzte Woche [ t₁ es₂ gegeben ] hat 
that she₊nom me₊dat already last week it₊acc given has 
b. *dass sie mir schon letzte Woche [ es₂ zu lesen ] schien 
that she₊nom me₊dat already last week it₊acc to read seemed 
c. *dass sie mich schon letzte Woche [ es₁ lesen ] ließ 
that she₊nom me₊acc already last week it₊acc read let 

However, things are different with restructuring infinitives under control predicates. As shown in (29), unstressed pronoun fronting to the left edge of the embedded vP is legitimate. This suggests that there is a CP layer in the embedded domain, such that C can license embedded unstressed pronoun fronting at an earlier stage of the derivation before this C head is subsequently subject to structure removal – and we can be sure that restructuring (i.e., removal of the CP and TP layers) has eventually taken place in the embedded domain because another unstressed pronoun can undergo long movement from this domain.

(29)  dass sie mir₁ schon letzte Woche [ t₁ es₂ t₂ zu geben ] 
that she₊nom me₊dat already last week it₊acc to give 
versucht hat 
tried has

A second piece of evidence supporting a CP status of restructuring infinitives embedded under control verbs involves the third construction. Note first that the correct generalization about which categories of the extended projection of V can undergo extraposition in Standard German is that CP can undergo extraposition in German, whereas vP, VP, and TP cannot do so.¹⁸ The option of extraposition is illustrated in (30ab) for finite clauses and non-restructuring infinitives, respectively.

(30)  a. dass er t₁ gesagt hat [ CP₁ dass es regnet ] 
that he₊nom said has that it₊nom rains

¹⁸The only obvious case where this generalization is not true is the Ersatzinfinitiv (‘infinitivus pro participio’) construction, but this would seem to be the exception that proves the rule – by its very nature, the Ersatzinfinitiv (‘substitute infinitive’) suggests a repair operation, i.e., the violation of grammatical principles (more specifically, of morphological selection (status government), and of the ban on non-CP extraposition of projections in the clausal spine that is currently at issue) in order to prevent greater damage; see Schmid (2005).
b. dass sie \( t_1 \) versucht hat \( [\text{CP}_1 \ PRO \ zu \ schlafen] \)
    that she\(_{nom}\) tried has to sleep

In contrast, it is not possible to extrapose a TP (as in (31a), with an exceptional case marking verb *sehen* ('see')) or a vP/VP (as in (31b)).

\[(31)\]
\[\text{a. } *\text{dass ich } t_1 \text{ gesehen habe } [\text{TP}_1 \ den \ Mann \ das \ Buch \ lesen] \]
\[\text{that I}_{nom} \ seen \ have \ \text{the} \ man_{acc} \ \text{the} \ book_{acc} \ \text{read}\]
\[\text{b. } *\text{dass er } t_1 \text{ hat } [\text{VP}_1 \ das \ Buch \ gelesen] \]
\[\text{that he}_{nom} \ has \ \text{the} \ book_{acc} \ \text{read}\]

However, extraposition is possible in the third construction (see Besten & Rutten (1989)), i.e., in cases where there is scrambling or unstressed pronoun fronting from a control infinitive, which thus shows that restructuring must have taken place; see (32).

\[(32)\]
\[\text{a. dass sie } \text{ihn}_2 \ t_1 \text{ versucht } [\text{CP}_1 \ PRO \ t_2 \ zu \ küssen] \]
\[\text{that she}_{nom} \ \text{him}_{acc} \ \text{tries} \ \text{to kiss}\]
\[\text{b. dass es}_2 \ Fritz \ ihr_3 \ t_1 \text{ empfohlen hat } [\text{CP}_1 \ PRO \ t_1 \ zu \ lesen] \]
\[\text{that it}_{acc} \ \text{Fritz}_{nom} \ \text{her}_{dat} \ \text{recommended has} \ \text{to read}\]

This strongly suggests that the extraposed item is a CP at the point where extraposition applies. If the third construction were to involve extraposition of a VP (see Wöllstein-Leisten (2001) and Haider (2010)), or of a vP or TP, ungrammaticality should be expected to result in (32). After extraposition of the CP, matrix V successively removes the CP and TP layers, thereby creating a restructuring environment, and scrambling and unstressed pronoun fronting can then take place from the extraposed vP in accordance with locality constraints.\(^{19}\)

Thus, as in the case of passivization, a Remove-based approach to restruc-
turing ensures that conflicting pieces of evidence concerning the presence of CP and TP layers in infinitives embedded under certain control verbs are accounted for: At the beginning of the derivation, CP and TP are present, but later they are removed by the matrix V bearing [–C₀–] and [–T₀–] features.

With these structure removal analyses of passivization and restructuring as background, let me now return to long-distance passives.

5. Long-Distance Passive

5.1. Analysis

As noted in section 1, there is an implicational generalization emerging if one considers the capacity of control verbs to trigger restructuring on the one hand, and the capacity of control verbs to participate in long-distance passive on the other: Control verbs that permit restructuring basically form a proper superset of control verbs that permit long-distance passivization. This generalization holds exceptionless if one further assumes (as I will do here) that all (subject or object) control predicates can in principle undergo passivization.²⁰ Thus, consider the examples in (33) vs. the examples in (34). In (33), versuchen (‘try’) participates both in restructuring, as evidenced by long-distance scrambling in (33a), and in long-distance passivization, as in (33b) (= (2a)). In (34) (= (8)), beabsichtigen (‘intend’) can trigger restructuring, as in (34a), but not long-distance passivization, as shown by the ungrammaticality of (34b).

(33)  a. dass den Traktor₁ keiner t₁ zu reparieren versucht hat
      that the tractor_{acc} no-one_{nom} to repair tried has
     b. dass der Traktor zu reparieren versucht wurde
      that the tractor_{nom} to repair tried was

(34)  a. dass den Traktor₁ keiner t₁ zu reparieren beabsichtigt hat
      that the tractor_{acc} no-one_{nom} to repair intended has

²⁰Passivization is not possible with wollen (‘want’), which induces restructuring. However, with wollen, restructuring is in fact obligatory; and this verb also differs from typical control verbs in German in not governing the second status (see Bech (1955/1957)), i.e., a zu infinitive, but the first status, i.e., a bare infinitive. In line with this, there is good independent evidence that this verb is not a regular control verb to begin with (see Geilfuß (1992) and Gergel & Hartmann (2009)), even though it has typically been analyzed as one.
b. ?*dass der Traktor zu reparieren beabsichtigt wurde
   that the tractor\textsubscript{nom} to repair intended was

I would like to suggest that this state of affairs can be captured if the long-distance passive in German involves (i) removal of DP\textsubscript{ext} by matrix v (like the regular passive), (ii) removal of CP and TP layers by matrix V (as with standard cases of restructuring), and (iii) one additional operation which ensures that not all control verbs that permit passivization and restructuring can show up in this construction. More specifically, I would like to propose that this third operation is also one involving Remove: It is the removal of the embedded vP layer, triggered by an additional feature [–v\textsubscript{0}–] on matrix V. I will argue that an immediate consequence of removal of the embedded vP in long-distance passives is that a direct object in the embedded vP cannot receive accusative case anymore, and thus needs to rely on nominative case assignment by matrix T.

To illustrate the Remove-based approach to long-distance passives, let me consider a step-by-step derivation of (2a) (= (33b)), which is repeated here once more for convenience.

$$(35) \quad \text{dass der Traktor zu reparieren versucht wurde}$$
$$\quad \text{that the tractor\textsubscript{nom} to repair tried was}$$

In the first part of the derivation, a complete infinitival CP is generated. First, DP\textsubscript{int} (\textit{der Traktor}) is merged with V (\textit{zu reparieren}), triggered by [•D\textsubscript{2}•] on V; see (36).

$$$(36) \quad \text{Merge (V[•D\textsubscript{2}•], DP):}$$
$$\quad \text{VP:}$$
$$\quad \text{DP} \quad \text{V}$$
$$\quad \text{der Traktor} \quad \text{zu reparieren}$$

Next, v is merged first with VP (triggered by [•V\textsubscript{2}•]), and then with DP\textsubscript{ext} (PRO) (triggered by [•D\textsubscript{2}•]). The result of applying these two operations in shown in (37).
At this point PRO's referential index is not yet determined (i.e., control has not yet been carried out), which is signalled here by [□]. In addition, I follow Adger (2003), among others, in assuming that PRO must be formally licensed by a designated head that ensures that it is spelled out as zero (also see Chomsky (1981) on the PRO theorem, and Rizzi (1986) on the basic idea that empty pronominal elements must both be formally licensed and referentially identified). Generalizing Adger's proposal, I assume that PRO needs to be licensed by a c-commanding phase head, in accordance with Chomsky's (2001) Phase Impenetrability Condition (PIC), as in (38). I also assume that the requirement in (38) must be satisfied at the end of the derivation.\(^{21}\)

\[\text{(38) \quad PRO Licensing:} \]

\[\text{PRO must be c-commanded by a PIC-accessible phase head.} \]

After the completion of vP, T is merged with it (driven by [●v₂●] on T), and then C is merged with TP (via [●T₂●] on C); as noted above, there is no obligatory EPP requirement in German, so that PRO can stay in situ, within vP. The outcome of these two steps is shown in (39).

---

\(^{21}\)I.e., licensing of PRO by a c-commanding accessible phase head X at some early stage cannot be preserved if X at some later stage fails to be in a configuration in which licensing obtains; this will become relevant below.
Thus, like other restructuring infinitives, long-distance passive infinitivals are at first full CPs. Next, the matrix control verb is merged with CP (triggered by $[\bullet C_2 \bullet]$ on V); see (40), which is exactly like (22a). In fact, the first steps are virtually identical to what has been shown for ordinary restructuring with control infinitives in (22) above.\(^\text{22}\)

\(^{22}\)I assume here that *versucht wurde*, although morphologically complex, is actually a single V item in the syntax which can in principle be split up by verb-second movement, but nothing depends on this particular assumption in the present context. Note also that if morphological realization is not post-syntactic, the proper choice of passive auxiliary and verb form can be read off the set of features on V that will subsequently trigger removal of layers; see footnote 9.
It is at this point that recursive structure removal starts. First, matrix V removes the CP layer again, as shown in (41).

In the next step, the TP layer is removed by matrix V, via discharge of $[-T_0-]$; see (42). Recall that TP is accessible to matrix V at this point, so the Strict
Cycle Condition is not violated (but it would be if TP removal were to apply in (40), when CP is still present).

(42) \emph{Remove} \((V_{[-\text{T}_0-]>[-\text{v}_0-]}, \text{TP})\):

\[
\begin{array}{c}
\text{VP} \\
\text{vP} \\
\text{PRO}_{[\Box]} \\
\text{V} \\
\text{DP}
\end{array}
\]

\[
\begin{array}{c}
\text{versucht wurde} \\
\text{VP} \\
\text{v} \\
\text{V}
\end{array}
\]

\[
\begin{array}{c}
\text{der Traktor} \\
\text{zu reparieren}
\end{array}
\]

So far, everything is exactly as shown above for standard restructuring environments; see (22). However, next the crucial step in generating long-distance passives takes place, and it is this step that is mainly responsible for the special properties of the construction: The matrix verb also removes the vP shell, due to \([-\text{v}_0-]\) on V. Unlike the CP and TP layers, the embedded vP has a specifier at this point, viz., \(\text{DP}_{\text{ext}} = \text{PRO}\). As laid out in section 2 (in the relevant part of an abstract derivation that is documented in (11)), the specifier is reassociated with the projection of the head responsible for the removal; and since reassociation must preserve the original c-command and linearization relations within the removed layer (such that \(\text{DP}_{\text{ext}} = \text{PRO}\) c-commands and precedes VP), PRO is integrated into the VP as a specifier. All of this is illustrated in (43).
As a consequence of removal of the vP layer and reassociation of PRO as a specifier of VP in (43), DP_{int} (der Traktor) finds itself in a situation where it cannot receive structural accusative case – a dilemma that naturally arises if case assignment (or case checking) is viewed as an output-oriented, late operation. Depending on whether an approach to structural case assignment in terms of functional heads or a dependent case approach is adopted (see above), this dilemma may be traced back to one or the other of two different sources: First, if v assigns accusative case, DP_{int} cannot get accusative case simply because v is gone. Second, if PRO is normally responsible for assigning dependent accusative case in infinitival complements, DP_{int} cannot get accusative case because PRO is not in Specv anymore (and daughters of VP, e.g., dative objects, never license accusative case on a co-argument). Assuming every DP to be in need of some case assignment, this means that DP_{int} will need to be assigned case by finite matrix T; it will therefore invariably be marked as nominative in the long-distance passive construction (as already indicated in all the previous derivational steps by rendering the article in the nominative form: der Traktor), and agree with matrix T (hence, the matrix verb) in person and number. For now, however, V has discharged all of its features for triggering operations, and the derivation moves on to a new cycle.

The next operation is a concatenation of matrix v with the VP generated in (43); v is a regular passive head that first triggers Merge with a VP, then with a DP_{ext}, and finally removes this DP again, as illustrated in (13) above. The first of these steps is shown in (44).
At this point, control of PRO\textsuperscript{[□]} can take place, which ensures binding of PRO (in its derived position as a daughter of VP) by the matrix subject DP via Agree-based valuation of the index feature of PRO; see (46). This ensures that that PRO will now invariably be interpreted as bound by the matrix subject, even if the latter item leaves the representation again (as an instance of counter-bleeding).

\textsuperscript{23}I leave open the questions of how exactly this Agree operation is triggered, how minimality
(46) Agree (DP_{ext}, PRO_{[□]}):

\[
\begin{array}{c}
\text{vP} \\
\text{DP}_1 \\
\text{v'} \\
\text{VP} \\
\text{PRO}_1 \\
\text{v'} \\
\text{VP} \\
\text{DP} \\
\text{V} \\
\text{versucht wurde}
\end{array}
\]

der Traktor zu reparieren

The final operation triggered by the current root \( v \) is removal of \( \text{DP}_{ext} \), i.e., passivization in the matrix domain; see (47).

(47) Remove \( (v_{[\neg D_2]}), \text{DP}_{ext}):\n
\[
\begin{array}{c}
\text{vP} \\
\text{VP} \\
\text{PRO}_1 \\
\text{V'} \\
\text{VP} \\
\text{DP} \\
\text{V} \\
\text{versucht wurde}
\end{array}
\]

der Traktor zu reparieren

Next, \( T \) is merged with the vP in (47):

distinguishes between subject and object control, and how the additional role of the matrix predicate is implemented (e.g., with promise-type verbs that give rise to subject control in the presence of a matrix object); these considerations are orthogonal to present concerns.
At this point, \( T \) is in a position where it can assign nominative case to \( \text{DP}_{\text{int}} \) of the embedded clause (\textit{der Traktor}). This is necessary because \( \text{DP}_{\text{int}} \) has no other possible source of case; and it is possible because (i) \( \text{PRO} \) is not in need of regular (nominative or accusative) case assignment (but rather needs a c-commanding phase head), (ii) there is no \( \text{DP}_{\text{ext}} \) in the specifier of matrix \( v \) anymore that would be in need of case assignment, and (iii) case assignment via Agree is independently known to be able to circumvent the blocking effect of an intervening phase boundary (possibly because Agree can be cyclic, as proposed by Legate (2005), or because Agree is simply exempt from the PIC, as suggested by Bošković (2007)).

Finally, \( C \) is merged with the TP, and the derivation terminates (more precisely, the relevant part of the derivation does so – the CP thus derived is of course still an embedded one since it is headed by a complementizer signalling embedding).
According to the PRO Licensing condition in (38), PRO must be c-commanded by a PIC-accessible phase head in the output representation at the end of the derivation. There are two a priori candidates for the licensing of PRO in (49), viz., C and v. However, the status of v as a phase head in (49) is not fully uncontroversial. The functional head v qualifies as what has sometimes been viewed as a weak phase head: On the one hand, it does not take a DP\textsubscript{ext} as a specifier (at least not in the eventual output representation, under present assumptions) – i.e., it does not qualify as a complete v*, in Chomsky’s (2008) terms –, but on the other hand, there is evidence that it is active from a PIC perspective – i.e., it creates a locality domain for (some) syntactic operations (see Legate (2003) for empirical evidence for this). I would like to conclude from this that v in (49) does project a phase that can give rise to PIC effects, but that v itself is too weak to belong to the class of licensing heads for the purposes of (38) (which, then, rather includes C and v*). Consequently, it has to be matrix C that satisfies (38) by formally licensing PRO in (49). Such licensing is compatible with the PIC, given that (i) VP is actually a part of the edge domain of v (this is a straightforward assumption since VP is a daughter of vP as a consequence of DP\textsubscript{ext} removal), and (ii) edge domains are defined recursively (i.e., if α is in the edge domain of β, and β is in the edge domain of γ, then α is also in the edge domain of γ; see, e.g.,
Richards (2011)). In contrast, in standard restructuring environments where matrix v does not trigger passivization (as they were discussed in section 4 above), it is this latter functional head that licenses PRO. And finally, if no restructuring takes place, it is the embedded C head that fulfills this role, essentially as envisaged in Adger (2003).

A remarkable property of the present analysis is that the final output representation in (49) in extremely small; and as a matter of fact, if one considers only the projections in the clausal spine, representationally it is almost indistinguishable from Wurmbrand’s (2001) proposal according to which long-distance passives in German involve embedding of a bare VP by the matrix verb. I take this to be a welcome consequence (and I adduce some independent evidence for this in subsection 5.3 below). However, it is worth noting that (49), unlike Wurmbrand’s structure for long-distance passives, still manages to integrate the embedded DP$_{ext}$ PRO, with its interpretation fixed via control in the course of the derivation. It seems clear that alternative approaches envisaging a small output representation that do without structure removal in the course of the derivation cannot easily reconcile the presence of PRO with the little bits of structure that are present in (49) (there is no obvious way to externally merge PRO in the matrix VP domain from the start). In view of this, in approaches like the ones developed in Wurmbrand (2001, 2007, 2015a,b), it is postulated that control in long-distance passivization structures in German (and perhaps more generally in restructuring environments involving control verbs) is treated lexically, via (some operation like) functional composition that brings about the identification of an argument of the matrix predicate with the external argument of the embedded predicate (also see Müller, St. (2002), Stiebels (2007), and Haider (2010), among others). However, if control in non-restructuring contexts (where the final representation is still biclausal) is handled syntactically, this gives rise to a non-homogeneous theory where there are two independent means to identify argument positions in control. And ceteris paribus this means that the phenomenon of control shift (see Růžička (1983), Wurmbrand (2002), Stiebels (2007)) is wrongly predicted to be possible with restructuring and with long-distance passivization. However, control shift never takes place with restructuring in German (i.e., if a certain control relation is present in a given non-restructuring environment, it is invariably also present if restructuring takes place in the same environment), and it certainly also never takes place with long-distance passivization: Subject control with an active matrix verb always corresponds to
subject control in a long-distance passive construction (cf., e.g., the examples with *versuchen* (‘try’) in (33)), and object control with an active matrix verb always corresponds to object control in a long-distance passive construction (see the examples with *empfehlen* (‘recommend’) in (3)). The non-existence of control shift under long-distance passivization, while potentially problematic for mixed approaches to control, follows directly under present assumptions because Agree-based control of an empty category PRO is uniformly involved, even if the final output representation is very small.

5.2. Open Issues

In the present analysis, there is no intrinsic relation between passivization (i.e., removal of DP$_{ext}$) in the matrix domain and massive restructuring (i.e., removal of the CP, TP and vP layers) in the embedded domain. The question is how this consequence can be enforced in a null theory of long-distance passivization that just combines passivization and restructuring (given that vP is not removed in the latter process), and whether the approach may lead to overgeneration. Another question arising under the analysis developed in the previous section is whether more needs to be said about nominative case assignment to DP$_{int}$ than has been said so far. I address these two issues in turn.

5.2.1. Tackling Overgeneration

The two single most important operations in the analysis laid out in the previous subsection are (i) *Remove* (V$_{[\neg v0]}$, vP) (and the ensuing reassociation of PRO in the VP domain) in (43) and (ii) *Remove* (v$_{[\neg D2]}$, DP$_{ext}$) in (47). What if one applies without the other? Would this lead to ungrammatical sentences that the theory wrongly predicts to be well formed? Closer scrutiny shows that this is not the case.

Consider scenario (i) first. Here, in the restructuring infinitive, *Remove* (V$_{[\neg v0]}$, vP) applies but is not ultimately followed by *Remove* (v$_{[\neg D2]}$, DP$_{ext}$) in the matrix domain. This would give rise to something like (50).

(50) *dass der Karl der Traktor zu reparieren versucht hat \\
    that the Karl$_{nom}$ the tractor$_{nom}$ to repair tried has

As it turns out, (50) is excluded by the analysis without further ado: In this
case, the embedded DP\textsubscript{int} cannot get accusative case (because of the removal of the embedded vP layer), but it can also not get nominative case from matrix T because that is still needed for DP\textsubscript{ext} in the matrix vP. Given that there is no other source for case assignment to the embedded DP\textsubscript{int}, ungrammaticality must arise.\footnote{Also note that unlike Wurmbrand (2001), I do not assume that the control verb itself can ever assign case to the embedded DP\textsubscript{int}.}

Next consider scenario (ii). Here, Remove (\(v[-D_2-]\), DP\textsubscript{ext}) applies in the matrix domain, generating a passive structure, but Remove (\(V[-v_0-]\), vP) fails to apply in the embedded domain. If restructuring does not take place in the infinitive (i.e., if matrix V not only lacks a \([-v_0-]\) feature for structure removal, but also \([-C_0-]\) and \([-T_0-]\) features), a well-formed derivation can be generated, which corresponds to a sentence like (51) (= (1a-i)).

\begin{equation}
\text{(51) dass gestern [CP PRO den Traktor zu reparieren ] versucht that yesterday the tractor\textsubscript{acc} to repair tried wurde was}
\end{equation}

However, if restructuring (removal of the CP and TP layers via \([-C_0-]\) and \([-T_0-]\) features) takes place but the final step of removal of the vP projection does not occur, the result should be a sentence that behaves exactly like the long-distance passive, except for the case change on an embedded DP\textsubscript{int} (nominative instead of accusative). As shown in (52) (where \textit{den Traktor} (‘the tractor\textsubscript{acc}’) undergoes scrambling to the matrix domain, thus indicating restructuring, i.e., the presence of CP/TP removal), a sentence of this type is not grammatical.

\begin{equation}
\text{(52) *dass den Traktor gestern zu reparieren versucht wurde that the tractor\textsubscript{acc} yesterday to repair tried was}
\end{equation}

Similarly, there is a sharp contrast between verb cluster topicalization (as another indication of restructuring) in (53a) (where DP\textsubscript{int} bears nominative, as expected if the embedded vP layer has been removed) and (53b) (where DP\textsubscript{int} bears accusative, which implies the continued presence of the embedded vP layer despite CP/TP removal).

\[]\text{dass gestern [CP PRO den Traktor zu reparieren ] versucht that yesterday the tractor\textsubscript{acc} to repair tried wurde was}\[]\text{(51) dass den Traktor gestern zu reparieren versucht wurde that the tractor\textsubscript{acc} yesterday to repair tried was}\[]\text{Similarly, there is a sharp contrast between verb cluster topicalization (as another indication of restructuring) in (53a) (where DP\textsubscript{int} bears nominative, as expected if the embedded vP layer has been removed) and (53b) (where DP\textsubscript{int} bears accusative, which implies the continued presence of the embedded vP layer despite CP/TP removal).}\[]\text{Also note that unlike Wurmbrand (2001), I do not assume that the control verb itself can ever assign case to the embedded DP\textsubscript{int}.}\]
The key to a solution of this problem is the realization that it is not in fact case assignment to DP\textsubscript{int}, or, for that matter, any property of DP\textsubscript{int} that is responsible for the illformedness of (52) and (53b). Rather, it is a violation of the licensing condition for PRO in (38) that poses the problem here. Recall from the discussion below (49) that, with matrix v being too weak, it has to be matrix C that licenses the embedded PRO, and this is possible (given the PIC) only if PRO is displaced into the matrix domain – which it is automatically after removal of the embedded vP. If, however, the embedded vP projection remains present, PRO is too deeply embedded to be accessible by the only potential licensing phase head C; see the abstract representation in (54) (with phases underlined).

\begin{align*}
(54) & \quad [\text{CP} \ C \ [\text{TP} \ [vP \ [vP \ PRO \ [v' \ VP \ v \ ]] \ V \ ] \ v \ ] \ T \ ]
\end{align*}

Independently of what the exact structure of the matrix VP in (54) looks like, it seems clear that even with a recursive concept of edge of phase, matrix C cannot reach into the embedded vP phase across an intervening vP phase and license PRO in its in situ position. This reasoning guarantees that if passivization takes place in the matrix domain and restructuring takes place in the embedded domain, the latter procedure has to go all the way, including a removal of vP, so that PRO becomes accessible to matrix C and can satisfy condition (38). Thus, no specific assumptions for long-distance passivization are required to ensure that removal of vP must take place here, and the null hypothesis can be fully maintained.

5.2.2. Nominative Assignment

I have so far assumed that DP\textsubscript{int} of the embedded verb is assigned nominative case in long-distance passives in a structure like (49) by matrix T, via Agree. This is fully in line with the general observation that nominative can be assigned to a vP-internal position, or even to a VP-internal position (in the case of standard passive structures) in German; and an EPP feature on T only shows up optionally, not obligatorily (see den Besten (1981), Grewendorf
(1989), and Haider (2010), among many others). However, based on contrasts like the one in (55a) vs. (55bc) (also cf. Müller, St. (2002)), it has been argued by Wurmbrand (2001), Bobaljik & Wurmbrand (2005) and Bobaljik (2015) that \( \text{DP}_{int} \) needs to undergo case-driven movement to the matrix SpecT position in a representation like (49).

\[
\begin{align*}
(55) & \quad \text{a. } *[a \text{ Ein blauer Wagen zu reparieren }] \text{ wurde hier nicht vergessen} \\
& \quad \text{a blue car}_{\text{nom}} \text{ to repair was here not forgotten}
\end{align*}
\]

Wurmbrand and Bobaljik's conclusion is that whereas (55bc) show that topicalization of (the higher or lower) VP in a structure like (49) is completely unproblematic in long-distance passives, the fact that a \( \text{DP}_{int} \) bearing nominative cannot undergo topicalization together with the verb that it is the internal argument of can be taken to show that the only way \( \text{DP}_{int} \) could move together with the verb would be as a TP, and TPs are independently excluded from undergoing topicalization in German. This thus qualifies as an argument for obligatory movement of \( \text{DP}_{int} \) to SpecT. In principle, this approach could be incorporated into the present analysis without problems. It would essentially suffice to follow Bobaljik & Wurmbrand (2005) in assuming an extended, slightly more complex (relativized) concept of phase from which it may then follow that structures of the type in (49) do in fact not permit nominative assignment by T to \( \text{DP}_{int} \). However, there are a number of observations which shed doubt on the idea that the illformedness of (55a) can be traced back to obligatory case-driven movement of \( \text{DP}_{int} \) in the long-distance passive construction in German.

The first thing to note is that a sentence like the one in (56) should be excluded for exactly the same reasons as the one in (55a): If \( \text{ein blauer Wagen} \) (‘a blue car\(_{\text{acc}}\)’) is in SpecT, and TP cannot be fronted, there is no way how it could be topicalized together with the two verbs. An example of this type is indeed starred in Bobaljik & Wurmbrand (2005, 824), but such a data judgement does not seem to be widely shared by speakers: There does not seem to be anything wrong with (56).
Second, recall from the discussion of the examples in (27) in section 4 that fronted unstressed pronouns can optionally be preceded by a moved nominative subject, but not by non-pronominal objects or adverbials, and that this provides evidence not merely for movement to SpecT of nominative DPs that show up to the left of an unstressed pronoun, but also crucially for a lower position of nominative DPs that show up to the right of an unstressed pronoun. On this basis, an example like (57) (= (3a), also cf. (3b), (3c)) shows unambiguously that DP$_{int}$ does not have to undergo case-driven movement to SpecT in long-distance passives: DP$_{int}$ (der Artikel (‘the article$_{nom}$’)) follows the fronted unstressed pronoun ihm (‘him$_{dat}$’), and thus cannot be in SpecT but must have remained vP-internally.

(57) dass ihm der Artikel zu lesen empfohlen wurde
    that him$_{dat}$ the article$_{nom}$ to read recommended was

And third, it seems that examples of the type in (55a) can be somewhat improved if, e.g., a bare plural DP$_{int}$ replaces the singular indefinite DP$_{int}$, and the argument is resumed TP-internally via stranded pre-nominal material as part of an NP-split construction; see (58).

(58) ??[a Traktoren zu reparieren ] sind hier keine versucht worden
    tractors$_{nom}$ to repair are here none tried been

All in all, I would like to conclude that there is no obligatory movement of DP$_{int}$ to the matrix SpecT position in the long-distance passive construction. However, I will have to leave open what the illformedness of (55a) is due to.\textsuperscript{25} \textsuperscript{26}

\textsuperscript{25}Wurmbrand (2001) and Bobaljik & Wurmbrand (2005) also provide another argument for such an obligatory movement step, based on obligatory wide scope of nominative DP$_{int}$ in long-distance passives; but, as shown in Keine & Bhatt (2016) (also cf. Bobaljik (2015)), the same effect also shows up with adverbs, and is in both cases ultimately due to verb cluster formation.

\textsuperscript{26}Based on the melioration effect documented in (58), one might speculate that (55a) indicates a parsing problem due to a surprising morphological realization of DP$_{int}$ as non-accusative (the bare plural DP$_{int}$ in (58) is morphologically locally ambiguous between a nominative and accusative (but syntactically unambiguously a nominative).
5.3. Evidence for Selective Accessibility

If long-distance passives in German are derived by structure removal affecting (i) DP\textsubscript{ext} of the matrix domain (regular passivization), (ii) CP and TP layers of the embedded domain (regular restructuring), and (iii) the vP layer of the embedded domain (because displaced PRO cannot otherwise be licensed), the prediction is that temporary accessibility effects as they show up with passivization (see section 3) and restructuring (see section 4) should occur with all the items affected by structure removal in long-distance passives. As shown in the following three subsections, this prediction is borne out.

5.3.1. Selective Accessibility of DP\textsubscript{ext}

The tests documenting downward accessibility and upward inaccessibility of DP\textsubscript{ext} in the passive can be replicated with long-distance passivization. Consider downward accessibility first. As shown in (59), DP\textsubscript{ext} of the matrix domain can control a PRO in an adjunct clause ((59a) involves subject control, in (59b) there is object control):

(59)  
\begin{align*}
\text{a. } & \text{ dass DP}_{\text{ext}} \text{ PRO}_1 \text{ der Traktor zu reparieren versucht wurde } \quad \text{that } \text{the tractor}_{\text{nom}} \text{ to repair tried was } \\
& \quad [\text{CP um PRO}_1 \text{ die Arbeit fortsetzen zu können }] \quad \text{in order to the work continue to can} \\
\text{b. } & \text{ dass DP}_{\text{ext}} \text{ ihm}_2 \text{ PRO}_2 \text{ der Artikel zu lesen empfohlen } \quad \text{that } \text{him}_{\text{dat}} \text{ the article}_{\text{nom}} \text{ to read recommended } \\
& \quad [\text{CP um PRO}_1 \text{ seine}_2 \text{ Reaktion überprüfen zu was } \text{in order to his reaction check to können }] \quad \text{can}
\end{align*}

Matrix DP\textsubscript{ext} can also control into secondary predicates. For the subject control environment in (60a), this is shown by the fact that widerwillig (‘unwillingly’) can only be construed with the matrix verb versuchen (‘try’), not with the embedded verb reparieren (‘repair’). For the object control case in (60b), matrix DP\textsubscript{ext} is the only candidate for control of PRO in the secondary predicate to begin with if these two items share an index – i.e., under a reading where the recommendation (and not the reading) takes place under the influence of drugs.
Apart from these considerations, it is clear that under the present analysis, long-distance passive, by its very nature, inherently implies control by matrix DP\textsubscript{ext} into a complement clause, determining the index of an embedded PRO via Agree.\textsuperscript{27}

Similarly, matrix DP\textsubscript{ext} can locally bind, and thus license, a reciprocal or reflexive pronoun in the matrix domain, cf. (61).\textsuperscript{28}

On the other hand, matrix DP\textsubscript{ext} cannot be bound by a quantified DP in a matrix clause (see (62a)).\textsuperscript{29} It also cannot by itself provide an accessible subject argument for control from a higher clause either (see (62b)). Thus, as regards accessibility, matrix DP\textsubscript{ext} behaves exactly as predicted under a Remove-based approach.

\textsuperscript{27}Needless to say, the same goes for cases where matrix passivization of a control verb does not go hand in hand with restructuring (see, e.g, (15b) above).

\textsuperscript{28}(61b) is somewhat marked under a non-reciprocal interpretation of sich, presumably because a scenario where someone recommends something to oneself is unusual.

\textsuperscript{29}As one would expect, if DP\textsubscript{ext}, is resumed by a by-phrase von ihm (‘by him’) in (62a), binding becomes possible.
5.3.2. Selective Accessibility of CP and TP

The same picture arises with regard to the embedded CP and TP projections in long-distance passive constructions. This is obvious for upward inaccessibility – there are no CP and TP layers that might block monoclausal properties like scrambling or unstressed pronoun fronting from the embedded domain. This is illustrated in (63).30

(63) dass ihr1 \(\text{DP}^{ext2}\) gestern widerwillig \(\text{PRO}_2\) [ t1 das Buch \(\text{zu}\) schenken ] versucht wurde
give tried was

Similar considerations apply in the case of all the other tests probing the absence of CP and TP (wide scope of negation, compactness, etc.). However, in the same way, there is also evidence for an initial presence of these layers in long-distance passive constructions. Thus, recall from section 4 that the very existence of the third construction (i.e., a combination of extraposition and transparency of an infinitival complement) provides an argument for an initial CP status of infinitival complements that participate in restructuring.

As observed by Wöllstein-Leisten (2001) and Haider (2010), extraposition is also an option in long-distance passivization; see the lack of a clear contrast between (64a) and (64b). Given that only CPs can undergo extraposition in German (not TPs, vPs or VPs), this provides an argument for an initial CP status of the infinitive.

30 Also cf. already an example like (53a), where the two non-finite verb forms are topicalized as a single constituent, stranding nominative DP\(_{int}\), which thus must have undergone scrambling from the embedded domain.
(64)  
a. dass der Hund zu füttern vergessen wurde
   that the dog\textsubscript{nom} to feed forgotten was

   b. dass der Hund vergessen wurde zu füttern
   that the dog\textsubscript{nom} forgotten was to feed

Another empirical argument for an initial CP status of the embedded infinitive in long-distance passivization comes from the distribution of unstressed pronouns. Recall from the discussion of (29) that unstressed pronoun fronting to the left edge of vP can leave the embedded infinitive (thus providing an argument for a smaller structure), but can (as a somewhat more marked option) also find enough space in the infinitive itself; this presupposes an initial presence of C that licenses the local movement operation. Essentially, this option would also seem to exist with long-distance passives; cf. (65).

(65)  
\[\text{dass mir}_1 \text{ schon letzte Woche } [ t_1 \text{ es}_2 \ t_2 \text{ zu geben } ] \text{ versucht}
\]
   that me\textsubscript{dat} already last week it\textsubscript{acc} to give tried

\[\text{wurde}
\]
   was

5.3.3. Selective Accessibility of vP

Deriving the variable accessibility of the embedded vP layer by structure removal is the central feature characterizing the present approach to long-distance passives. On the one hand, as already noted by Wurmbrand (2001), an absence of the embedded vP projection is required to account for the change from accusative to nominative of the embedded DP\textsubscript{int}. On the other hand, the only way to let an embedded DP\textsubscript{ext} (i.e., PRO) be part of the sentence is via an embedded vP projection (more specifically, in the present proposal, via [\(\bullet D_2 \bullet\)] on embedded v). In turn, a syntactic presence of PRO in long-distance passives is required if a systematic explanation of the complete absence of control shift under long-distance passivization is sought. This PRO subject is then predicted to be active throughout the derivation without qual-
ification; among other things, it can be held responsible for providing a local binder for the reflexive and reciprocal pronouns in (66).\(^{31,32}\)

\[(66)\]

\[
\text{a. } \text{dass PRO}_1 \text{ sich}_1 \text{ der Plan}_2 \text{ zu erklären versucht wurde} \\
\text{that PRO} \text{ refl to explain tried was}
\]

\[
\text{b. } \text{PRO}_1 \text{ einander}_1 \text{ zu reparieren empfohlen würden den} \\
\text{RECIP to repair recommended were the}
\]

\[
\text{Leuten}_1 \text{ die Traktoren nicht} \\
\text{people dat the tractors not}
\]

\(^{31}\)It should be noted, though, that Wurmbrand (2001, 278) has divergent judgements for examples which are structurally similar, and she takes this to be an argument that PRO is not present in German long-distance passives. However, even if PRO did not exist in (66a) and (66b), one should still expect DP of matrix V to be able to locally bind the reflexive pronoun in a subject control sentence like (66a), on a par with sentences like (17) (den Leuten in the object control sentence (66b) cannot do so under Wurmbrand's assumption hat dative DPs cannot bind – but cf. Featherston & Sternefeld (2003) for evidence to the contrary). Incidentally, I would surmise that the examples in Wurmbrand (2001, 278) involve an embedded verbal domain that does not lend itself naturally to passivization in the first place for many speakers.

\(^{32}\)Here are two further arguments in support of postulating an initial presence of vP. (Thanks to an anonymous reviewer for pointing them out.) First, since structure removal is intrinsically highly local, the illformedness of sentences like those in (i) can be derived. Here, an auxiliary (see (i-a)), or an additional control verb that is compatible with long-distance passive as such (see (i-b)), intervene between the V head responsible for removal and the structure that needs to be removed.

\[(i)\]

\[
\text{a. } \text{*dass der Traktor } \text{ repariert zu haben } \text{ beabsichtigt wurde} \\
\text{that the tractor nom fix to have inteded was}
\]

\[
\text{b. } \text{*dass der Traktor } \text{ zu reparieren zu versuchen } \text{ beabsichtigt wurde} \\
\text{that the tractor nom to fix to try inteded was}
\]

A second potential argument for an initial vP involves semantic interpretation. If it is assumed that the functional head v is in general responsible for introducing an abstract predicate like cause with agentive verbs like reparieren (‘fix’), then the fact that this meaning component is present with the embedded verb in a typical long-distance passives like (2a) finds a natural explanation in the present approach, where there is a vP at an earlier stage of the derivation (given that its removal is recoverable), but cannot be accounted for straightforwardly without further assumptions in an approach like Wurmbrand (2001), where there is never an embedded vP in German long-distance passives.
References


Tough-Displacement without Movement

Marie-Luise Schwarzer*

Abstract
While tough-constructions in English are standardly analyzed as biclausal structures that involve an A’ dependency in the embedded clause, German tough-constructions are generally conceived of as a type of passive. I challenge this view by presenting evidence of bi-clausal behavior of German TCs. I argue that these structures show both, typical monoclausal, passive-like, and typical bi-clausal properties. I propose (following Müller 2017a o.) that this paradoxical state of affairs is an indication of the derivational history of these constructions: they start out with an embedded CP that is at a later point of derivation depleted to VP size by a syntactic structure removal operation. The CP shell is removed from the derivation, which allows a DP in Spec,CP to be transported into the matrix clause without incurring an Improper Movement violation.

1. Introduction

Constructions like (1a), in which the object of an embedded verb shows up as the subject of a matrix predicate, pose interesting questions for syntactic theories. These so called tough-constructions, named after their characteristic predicates like tough, hard, easy, fun, etc., have been the object of much debate in recent years (see e.g. Hicks 2003, Rezac 2006, Hartman 2011, 2012, Pesetsky 2013, Fleisher 2013, Keine & Poole 2015, Longenbaugh 2016, Gluckman 2017, Keine & Poole 2017 and many more). They alternate with a semantically equivalent expletive construction (1b).

(1)  a. John is easy [to please (John)].
    b. It is easy [to please John].

The discussion has focused primarily on the construction in English. German possesses structures like (2), which look superficially similar.

*For comments and feedback I would like to thank Gereon Müller, Susi Wurmbrand, Omer Preminger, Philipp Weisser, and Martin Salzmann, as well as the audience at ConSOLE XXVI at University College London. All errors are my own.

Structure Removal, 167–214
Andrew Murphy (ed.)
LINGUISTISCHE ARBEITSBERICHTE 94, Universität Leipzig 2019
(2) weil dieser leckere Käsekuchen [leicht zu backen] ist
  'because this tasty cheese cake is easy to bake’

There are two groups of analyses for English-type *tough*-constructions: base-generation accounts and long-movement accounts. Both have in common (i) the clausal embedding structure – the *tough*-predicate takes a CP complement, and (ii) the existence of an A′ dependency in that CP (see e.g. Chomsky 1982). They differ on their assumptions regarding the element involved in the A′ dependency. Base-generation approaches suggest that an empty operator is merged in the embedded object position which moves to the edge of the embedded CP. It is coreferent with the subject that is base-merged in the matrix clause. Long-movement analyses, on the other hand, assume that the matrix subject itself is base-generated as the embedded object and moves (improperly) through the embedded clause into the matrix clause.

In contrast, German *tough*-constructions are standardly analyzed as monoclausal, passive-like structures (e.g. Höhle 1978, Comrie 1997, Demske-Neumann 1994, Holl 2010). However, there are also analyses that equate it to its English counterpart (e.g. Brinker 1969, Breckenridge 1975, Rosengren 1992).

These two strands of analyses are already indicative of the main puzzle that is addressed in this paper: I will demonstrate that German *tough*-constructions offer evidence for both, a small, monoclausal, and a bigger, embedding underlying structure. I propose that these seemingly conflicting properties can be reconciled under an analysis in which syntactic structure is removed at a certain point in the derivation. Thus, these structures behave like smaller and like bigger structures, because at different stages of the derivation, they are bigger, and then smaller. The crucial idea of the analysis is that *tough*-predicates embed CP-complements and trigger the removal of the CP shell. This removal of the CP projection allows the object DP of the embedded clause to enter the matrix clause and become the matrix subject.

The subsequent parts are structured as follows: section 2 investigates the size of German *tough*-constructions. It sheds light on the paradoxical behavior of these structures: while they are similar to a passive in some respects, other tests suggest that they involve clausal embedding like English *tough*-constructions. Section 3 reviews and compares previous analyses. A new analysis is proposed in section 4. Section 5 explores the question of improper movement and shows
how the observed properties of tough-constructions can be accounted for. Section 6 concludes.

2. The size of German tough-constructions

This section illustrates some properties of German tough-constructions. As will become clear, German tough-constructions differ from English ones in interesting ways. They show paradoxical behavior: in some respects they behave as if the tough-predicate embeds a clausal structure, parallel to their counterpart in English. At the same time, other tests suggests that the construction is monoclausal and passive-like.

I will first summarize arguments from the literature on German tough-constructions that argue for a small, monoclausal structure. The second part of this section presents mainly new evidence for the underlying biclausal, CP-embedding structure of these sentences.

2.1. In favor of a monoclausal structure

2.1.1. Object promotion

It has been observed that the passive and tough-constructions behave alike in their ability to promote the theme argument to grammatical subject (e.g. Hawkins 1986). In German, the indirect object may not be promoted to subject in passivized structures and it is also unable to be the subject in a tough-construction, see (3).

(3)  a. *dass der Junge geholfen wurde

   that the.NOM boy.NOM helped become.PASS

   b. *dass der Junge einfach zu helfen ist

   that the.NOM boy.NOM easy to help is intended: ‘that the boy is easy to help’ (Hawkins 1986)

Dative case has to be retained in both constructions.

(4)  a. dass dem Jungen geholfen wurde

   that the.DAT boy.DAT helped become.PASS

   ‘that the boy was helped’
b. dass dem Jungen einfach zu helfen ist
that the.DAT boy.DAT easy to help is
‘that the boy is easy to help’

English shows the mirror image: indirect and oblique objects can be *tough*-moved, and they can also be promoted to subject in passive constructions, see (5).

(5) a. He is easy to help.
b. He was helped.

This observation illustrates the superficially common property of *tough*-constructions and passives: the promotion of the direct object and the impossi-
bility to target non-direct objects in German. The line of argumentation in Hawkins (1986) and others is, that since they behave the same with respect to object promotion, these structures can be equated with one another.

2.1.2. *Tough*-movement is bounded

*Tough*-constructions in German are typically short dependencies: *tough*-movement may not cross multiple clause boundaries. As (6) shows, *tough*-movement is impossible across an embedded infinitive in German, in contrast to English.

(6) a. *dass dieses Buch schwer [CP Hans zu überzeugen
that thisNOM book.NOM hard Hans.DAT to convince
[VP (dieses Buch) zu lesen]] ist
this book to read is
b. that this book was easy to convince John to read

(Wurmbrand 2001:29)

English *tough*-movement is allowed across multiple clauses, as long as there is no intervening element in a Spec,CP position (like why in (7b)), which suggests an A′-dependency in the embedded clause (Nanni 1980, Hicks 2003). Compare (7a) and (7b)\(^1\).

\(^1\) Note that Gluckman (2017) judges *tough*-constructions like (i) which involve a finite embedded clause as ungrammatical, in contrast to Hicks (2003:43). Gluckman concludes that *tough*-constructions show a ‘weak’ A-bar dependency, since they do pass most tests for A′-movement, but are not completely unbounded.
(7) a. A guy like John is hard [to imagine [any woman believing [she could ever resist falling in love with ___ ]]]
b. ?? A guy like John is hard [to imagine [any woman wondering [why she could never resist falling in love with ___ ]]]

(Hicks 2003:43)

If a predicate embeds a CP and there is $A'$-movement in that CP (regardless of the element that moves), one would expect that this $A'$-movement has the typical property of applying successive-cyclically. It is expected to cross any number of CPs. In German, that does not seem to be the case.

This led some scholars to believe that there is no clausal embedding in these constructions in German. Instead, they propose a monoclausal structure like a canonical passive.

2.1.3. No attributive use

Tough-constructions can be used attributively in prenominal position in English, as in (8).

(8) a. a difficult book to read
b. an easy thing to say

Since they pattern similarly to adjectives with respect to their distribution, it is plausible to assume that tough-constructions consist of an adjective that embeds a clause. Tough-constructions in German do not show the same distribution.

Tough-predicates in German do not behave like adjectives in the sense that they cannot be used attributively in prenominal position (e.g. Comrie 1997, Holl 2010) (9).

(9) a. ein schwer zu les-end-es Buch
   a hard to read-PTCL-NOM.SG.NEUT book.NOM
b. ?? ein schwer-es Buch zu lesen
   a hard-NOM.SG.NEUT book.NOM to read

In German, the tough-element is invariant and does not bear inflection – properties of an adverb rather than an adjective, as argued by Comrie (1997).

(i) *The book was difficult to say that John read ___
Instead, the inflection is hosted on the participle verb form, which marks it as
the head of the attribute. Consequently, Comrie argues that German tough-
constructions do not consist of an adjective that embeds a complement clause,
unlike their English counterparts. Rather, they are a passive (morphologically
realized as a zu ‘to’-infinitive instead of a participle) which is modified by
an adverb (i.e. the ‘tough-predicate’; see also 2.1.4, Demske-Neumann 1994,
Comrie 1997).

2.1.4. Tough-predicate is optional

German tough-constructions may lack a tough-predicate, as in (10). These
sentences are ambiguous between two possible interpretations: they can receive
a reading of necessity or possibility. Depending on the lexical verb, one of the
readings is more salient.

(10) dass die Kälte jetzt ∅ [VP zu spüren war]
that the.NOM cold.NOM now to feel was
‘that it was possible to feel the cold now’ or ‘that one had to feel the cold
now’

(Höhle 1978)

The possibility of tough-constructions without an overt tough-predicate has
been used as an argument for the underlying passive nature of these construc-
tions (Höhle 1978, Comrie 1997). In these analyses, tough-constructions are
monoclausal structures with an infinitival predicate. The tough-adjective is
merely an optional adverb that modifies it. This structure has been called a
‘modal passive’ because of the modality readings it can receive (possibility or
necessity).

Additional evidence for this view comes from the fact that elements that are
unambiguously adverbs (and not adjectives) can occur in the tough-predicate
position, as in (11).

(11) a. dass der Brief kaum zu lesen war
that the.NOM letter.NOM hardly to read was
‘that one could hardly read the letter’

b. *der kaum-e Brief
the.NOM barely-NOM.SG.M letter.NOM
2.2. In favor of a biclausal structure

2.2.1. Non-passive verbs and non-tough verbs

Some German verbs cannot be passivized, but can still be part of a tough-construction: bekommen ‘to get’, erhalten ‘to receive’, erfahren ‘to learn’, haben ‘to possess’ (Holl (2010), Rosengren (1992)). In (12a) and (13a), passivization of bekommen and haben is impossible, while the tough-constructions in (12b) and (13b) are licit.

(12) a. *dass die Bücher am Schalter im Lesesaal
   that the.NOM books.NOM at desk in reading.room
   bekommen werden
   get.PTCL become.PASS
   intended: ‘that one can get the books at the desk in the reading room’
   b. dass die Bücher am Schalter im Lesesaal
      that the.NOM books.NOM at desk in reading.room
      (einfach) zu bekommen sind
      easy to get are
      ‘that one can get the books at the desk in the reading room (easily)’
      (Holl 2010, modified)

(13) a. *dass es gehabt wird
   that it.NOM possessed become.PASS
   intended: ‘that it is possessed’
   b. dass es noch zu haben ist
      that it.NOM still to possess is
      ‘that it can still be obtained’
      (M. Salzmann, p.c.)

In addition to that, intransitive unergative verbs can be passivized, but cannot be part of a tough-construction. (14)\(^2\) shows the impersonal passive of some unergative verbs.

\(^2\)Example (14a) was obtained from https://www.wertingen.de/rathaus-und-verwaltung/aktuelles/612-allein-der-spass-zaehlt.html, 27 July 2018. Example (14b) was obtained from Müller & Rieland (2006:289).
(14)  
   a. dass bei jedem Wetter gelaufen wird
      that in every weather run become.pass
      ‘that one runs in every weather’
   b. wenn von der Vereinbarung innerhalb der
      if from the.dat agreement within the.gen
      Klagefrist §4 des KSchG
      period.of.appeal §4 the.gen employment.protection.law
      zurückgetreten wird
      withdraw become.pass
      ‘if one withdraws from the agreement within the period of appeal
       according to §4 of the employment protection law…’

These verbs are illicit in *tough*-constructions, like (15).³

(15)  
   a. ?*dass bei jedem Wetter gut zu laufen ist
      that in every weather easy to run is
      intended: ‘that running is easy in every weather’

³Some authors (e.g. Holl 2010) report that *tough*-constructions with intransitive verbs are possible. Holl (2010) presents examples like (i):

(i)  
   a. ??Nun ist aber wirklich einzuschlafen.
      now is mod.part really to.go.to.sleep
      ‘One really has to go to sleep now.’ (Holl 2010:17, grammaticality judgment from Holl 2010)
   b. Ab 22 Uhr ist zu schlafen.
      from 22 o’clock is to sleep
      ‘One has to be asleep as from 10pm.’ (Holl 2010:18)

Since these examples involve unaccusative verbs, which cannot be passivized, they pattern like *bekommen* and *haben* in (12) and (13).

Note that there is a discussion about the ability of unaccusative verbs in German to be passivized. Recently, some authors have argued that genuine passivization of these verbs is possible (Primus 2010b, 2011, 2010a, Kiparsky 2013). Such passive realizations are standardly analyzed as a reinterpretation of unaccusative as unergative verbs (Růžička 1989, Fanselow 1992, Müller 1999, 2002). I follow this line of thought in concluding that unaccusative verbs cannot be passivized directly (see also arguments in Müller 2018).
b. *weil von der Vereinbarung innerhalb der Klagefrist
because from the agreement within the period of appeal
schwer zurückzutreten ist
hard to withdraw is
intended: 'because it is hard to withdraw from the agreement
within the period of appeal'

Altogether, it seems that there is a double dissociation between the ability
to occur in a passive and in a tough-construction, suggesting that these two
structures are not related to one another. This does not mean that tough-
constructions have an underllygroundly bigger structure than passives. It only
suggests that these two constructions cannot be identical.

2.2.2. Licensing of parasitic gaps

A’-movement is known to be able to license parasitic gaps (e.g. Engdahl
1983). Parasitic gap licensing is taken as a diagnostic of A’ movement in
tough-constructions since Chomsky (1982), Montalbetti et al. (1982), as in (16).
There is no consensus regarding the element that undergoes this movement:
in base-generation theories it is an empty operator that moves to the edge of
the embedded clause and is coreferent with the subject that is first merged
in the matrix clause (e.g. Fiengo 1980, Chomsky 1982, Rezac 2006, Keine &
Poole 2015). In long-movement accounts it is the matrix subject itself that
moves from the embedded into the matrix clause (e.g. Brody 1993, Hicks 2003,

(16) CDs are easy [to copy ___ [without having to pay good money for
___pg]].

(Hicks 2003:43)

The existence of parasitic gaps in German has been debated in the literature
concludes that the contexts which allow parasitic gaps are much more restricted
in German than in English. While parasitic gaps can occur in English in tensed
and tenseless adjunct clauses (17a), (17b) relative clauses (17c), subjects and
complement clauses (17d) a.o., in Standard German4 they are only licit in

4I only consider parasitic gaps in Standard German. Parasitic gaps in Bavarian and other
non-tensed adjunct clauses. They are licensed by A′-movement, *wh*-movement in (18a), and scrambling\(^5\) in (18b).

(17) **Parasitic gaps in English**

a. Which article did Ted copy ___ [without reading ___\(pg\)]?
   (Postal 1994)

b. Which colleague did John slander ___ [because he despised ___\(pg\)]?
   (Engdahl 1983)

c. the woman [who your attack on ___\(pg\) enraged ___]?
   (Postal 1994)

d. Who did you tell ___ [that we are going to vote for ___\(pg\)]?
   (Engdahl 1983)

(18) **Parasitic gaps in German**

a. Wen hat er [ohne ___\(pg\) zu mögen] freundlich ___
   who.ACC has he without to like friendly
   gegrüßt?
   greeted
   'Who did he greet nicely without liking?'

b. dass dieses Buch alle [ohne ___\(pg\) zu lesen] ___ ins
   that this.ACC book everyone without to read into
   Regal gestellt haben
   shelf put have
   'that everyone put this book on the shelf without reading it'
   (Müller 1995:173)

---

southern German varieties show different behavior and are structurally closer to English, see e.g. Lutz (2004).

\(^5\)There is much discussion about the nature of scrambling in German. It has been argued to be either A-movement (Fanselow 1987, 1990, Frey 1989, Moltmann 1990), A′-movement (Stechow & Sternefeld 1988, Sternefeld 1990, 1991, Müller & Sternefeld 1993, 1994, Vikner 1994), or have mixed A- and A′-properties (Webelhuth 1989, 1992). As (i) shows, parasitic gaps cannot be licensed by A-movement (see also the discussion in Müller 1995). I will consider scrambling to be A′-movement.

(i) *dass dieses Buch [ohne ___\(pg\) zu lesen] dem Peter ___ gegeben
   that this.NOM/ACC book without to read the.DAT Peter given
   wurde
   become.PASS
   intended: ‘that this book was given to Peter without reading’
Both *tough*-constructions and parasitic gap constructions are marginal structures in German. Crucially however, sentences in which they are combined are not significantly worse than sentences with only parasitic gaps. In the context where parasitic gaps are licit in German, they seem to be possible in *tough*-constructions as well. The sentences in (19) were judged to be as grammatical as sentences like (18a) and (18b) by 35 participants in an informal grammaticality judgment survey.

(19) **Parasitic gaps in German *tough*-constructions**

a. dass der Text einfach [ohne ___pg gründlich, that the.NOM text easy without thoroughly
durchzulesen] ___ zu verstehen ist
to.read to understand is
‘that the text is easy to understand without reading thoroughly’

b. dass die Türen [ohne ___pg zu beschädigen] ___ zu
that the.NOM doors without to damage to
schließen sind
close are
‘that the doors should/ can be closed without damaging them’

c. ?dass diese Skulpturen leicht [ohne ___pg zu
that these.NOM sculptures easy without to
beschädigen] ___ zu transportieren sind
damage to transport are
‘that these sculptures are easy to transport without damaging’

The acceptability of parasitic gaps in *tough*-constructions represents strong evidence that there exists an A′ dependency in German *tough*-constructions, parallel to the English construction. This in turn suggests that *tough*-constructions are not underlying monoclausal passive structures, since no theory of the passive assumes an A′ dependency. Consider as well (20), which shows that parasitic gaps cannot be licensed in canonical passives.

(20) **No parasitic gaps in German passives**

a. *weil er [anstatt ___pg freundlich zu behandeln] ___
because he instead friendly to treat
glängert wurde
bothered was
intended: ‘He was bothered instead of treating him nicely.’

b. *dass das Buch [ohne ___ pg zu kaufen] ___ bekritzelt wurde
   that the book without to buy doodled.on was
   intended: ‘that the book was doodled on without buying’

c. ?*dass Maria [ohne ___ pg anzusehen] ___ geküsst wurde
   that Maria without to.look.at kissed was
   intended: ‘that Maria was kissed without looking at’

2.2.3. Topicalization

A second argument against an underlying passive-representation of tough-constructions comes from topicalization: (active and passive) VPs can be topicalized to Spec,CP in German, while APs generally cannot, see (21).

(21) a. [VP Langusten gegessen] wurden nicht.
   
   ‘Crawfish wasn’t eaten.’

b. *[AP Langusten lecker] sind nicht.

   ‘Crawfish are not tasty.’

If tough-constructions are passives, we would expect the infinitival VP to be able to topicalize with the tough-adverbial leicht, parallel to (21a). This is, however, not what we find⁶, see (22).

(22) *[VP Linguisten leicht zu überzeugen] sind nicht.

   ‘Linguists are not easy to convince.’

Even structures that do not contain an overt tough-adjective cannot be topicalized⁷, (23).

---

⁶Note that (22) cannot be ruled out on the grounds of the subject being too low, inside vP. Haider (1990) shows that (indefinite) subjects in German can be topicalized with the vP, as in (i).

(i) [VP Ein Außenseiter gewonnen] hat hier noch nie.
   
   ‘An underdog has never won here.’ (Haider 1990:94)

⁷Even though the sentence in (23) is not completely ungrammatical for all speakers and
(23) \[ \text{Tough-Displacement without Movement} \]

\[ \text{Tough-Displacement without Movement} \]

\[
\begin{align*}
\text{Tough-Displacement without Movement} \quad \text{without Movement} \quad \text{VP Briefbomben zuzustellen] sind nicht. mail.bombs.NOM to.deliver are not intended: 'Mail bombs should not be delivered.'}
\end{align*}
\]

\text{Tough-constructions seem to pattern with APs rather than VPs with respect to topicalization, suggesting that their underlying structure is not a VP, but an adjective embedding a bigger structure.}

\text{2.2.4. Long scrambling}

\text{Some accounts of tough-movement assume that a tough-construction and the corresponding expletive construction are derivationally connected (e.g. Rosenbaum 1967, Postal 1971).}

\text{Expletive constructions show evidence for the existence of a CP shell. Scrambling is generally clause-bound in German (e.g. Ross 1967, but see Grewendorf & Sabel 1994 for exceptions). Thus, (24), where an element has been scrambled out of an embedded clause into the matrix clause, is ungrammatical.}

\[ \text{(24) weil ich den Mann, glaube, dass sie i liebt] because I the.ACC man.ACC believe that she loves intended: 'because I believe that she loves the man'} \]

\text{Likewise, in an expletive construction, scrambling out of the embedded clause is ungrammatical, (25).}

\[ \text{(25) dass es den Kuchen, leicht zu backen] war that it the.ACC cake.ACC easy to bake was intended: 'that it was easy to bake the cake'} \]

\text{Scrambling behavior thus indicates that a clause boundary is present in the expletive sentences. If tough-constructions are derived from expletive con-}

\text{becomes more grammatical with more material in the middle field, as in (i), the contrast between (21) and (22)/(23) certainly still obtains.}

\[ \text{(i) [VP Briefbomben zuzustellen] sind nur montags bis freitags von 12 bis mail.bombs.NOM to.deliver are only on.mondays until on.fridays from 12 to 14 Uhr. 14 o'clock 'Mail bombs should only be delivered from Monday to Friday from 12–2pm.'} \]
structions, the tough-predicate must at some point in the derivation embed a CP.

Interestingly, scrambling out of the tough-predicate complement in tough-constructions is permitted, (26).

(26) weil meine Nachbarin [CP meinem neuen Freund]_i
because my.NOM neighbor my.DAT new.DAT boyfriend.DAT
leicht [VP ___i vorzustellen] ist
easy to.introduce is
‘because my neighbor is easy to introduce to my new boyfriend’

2.2.5. Scope of negation

The scope of embedded negation is in the embedded clause (Haider 2010), (27).

(27) Sie hat versucht, [ihn nicht zu beunruhigen].
she has tried him not to alarm
‘She has tried not to alarm him.’ (versuchen » NEG, Haider 2010:19)

In expletive and tough-constructions, the negation cannot scope out of the complement of the tough-predicate, (28).

(28) a. dass es schwer war [ihm das nicht zu versprechen]
that it hard was him that not to promise
‘that it was hard not to promise that to him’ (schwer » NEG)
b. ?dass ihm dieses Versprechen schwer [nicht zu geben]
that him.DAT this.NOM promise hard not to give
war
was
‘that this promise was hard not to give him’

This scope restriction is indicative of a clause boundary.

2.2.6. Unstressed pronoun fronting

Unstressed pronouns have to be fronted, i.e. they have to be moved to the left periphery of vP. They can be preceded only by the subject (Müller 2016a). This is illustrated in (29) with the verb lehren ‘to teach’ which takes two accusative objects, one of which is the unstressed pronoun ihn ‘him’.

...
Tough-Displacement without Movement

(29) a. dass *ihn* die Maria den korrekten Umgang mit that him ACC the NOM Maria the ACC proper handling with Schusswaffen gelehr* hat* firearms taught has

b. dass die Maria *ihn* den korrekten Umgang mit that the NOM Maria him ACC the ACC proper handling with Schusswaffen gelehr* hat* firearms taught has

c. *dass die Maria den korrekten Umgang mit that the NOM Maria the ACC proper handling with Schusswaffen *ihn* gelehr* hat* firearms him ACC taught has

d. dass die Maria den korrekten Umgang mit that the NOM Maria the ACC proper handling with Schusswaffen den Peter gelehr* hat* firearms the ACC Peter taught has

‘that Maria taught him/Peter the proper handling of firearms’

It seems to be a fact about the syntax of German that this fronting can only happen in the presence of a CP in the relevant domain (Müller 2016a). Compare the embedded infinitivals in (30). (30a) shows a raising construction. Independent tests show that raising verbs do not embed a CP, but a smaller complement. Thus, fronting in the complement of a raising verb results in ungrammaticality. On the other hand, (30b) shows a control infinitival in which fronting of *es* is possible.

(30) a. *dass sie mir schon letzte Woche [*es zu lesen]* schien that she NOM me DAT already last week it to read seemed ‘that she seemed to me already last week to be reading it’

b. dass sie *mir* schon letzte Woche [*t_i es zu geben]* that she NOM me DAT already last week it to give versucht hat tried has ‘that she already tried to give it to me last week’ (Müller 2016a)

When we turn to embedded tough-infinitives, we find that they pattern with control-infinitives rather than with the complements of raising verbs, suggesting
that they possess an equally big structure, see (31). (31) again uses the double-accusative verb *lehren*. One of the accusative objects can be *tough*-moved, while the other, the pronoun, can be fronted or left inside the complement of the *tough*-predicate *leicht*.

(31) dass (ihn) der korrekte Umgang mit Schusswaffen that him.acc the.nom proper handling with firearms (ihn) nicht so leicht [ (?ihn) umfassend (*ihn) zu him.acc not so easy him.acc thoroughly him.acc to lehren ] war teach was ‘that the proper way of handling firearms was not all that easy to teach him thoroughly’

The parallel behavior of *tough*-complements and control-complements suggests that they have in common what raising constructions lack: a CP.

2.2.7. Reconstruction

Approaches to *tough*-constructions that posit a long movement chain from the position in the embedded clause in which the pivot DP is first merged to its surface position as the matrix subject, can be distinguished from base-generation accounts, where an empty operator is $A'$-moved in the embedded clause and connected to an overt subject merged in the matrix clause, by the possibility of reconstruction of the pivot DP into a position in the matrix clause (Longenbaugh 2017).

Analyses in which an element starts out in a lower position and moves into a higher one, predict the possibility of reconstruction into the lower position. It has been shown by Pesetsky (2013), Fleisher (2013), and Longenbaugh (2017) (a.o.) that reconstruction of the *tough*-subject into the object position of the embedded clause is possible in English\(^8\). The prediction is borne out in the same way in German. (32a) shows reconstruction for anaphor binding: the anaphor *sich selbst* can reconstruct into a position where it is c-commanded by

---

\(^8\)Unfortunately, it is not completely clear to what extent reconstruction really is possible in English. While the authors cited in this subsection argue that reconstruction for binding and scope is possible, Poole et al. (2017) and Gluckman (2017) argue against it. I will not discuss intervention effects in reconstruction. For that, see Hartman (2011), Keine & Poole (2017) and references therein.
its antecedent. (32b) shows scope reconstruction which is entirely parallel to English.

(32)  a. dass Bilder von sich selbst für Max schwierig ___ zu verschenken sind 'that pictures of himself are hard for Max to give as a present'

b. dass fünf Leute schwierig gleichzeitig ___ zu zufriedenzustellen sind 'that five people are hard to please at the same time' (hard > five)

Since reconstruction behavior is so similar in German and English, it is plausible to assume that the structure underlying the constructions in the two languages is similar, too (see also Salzmann 2018). The consensus in the literature is that the complement of an English tough-predicate is a CP. The parallel reconstruction behavior suggests that the same is true for German.

2.3. Interim summary: paradoxical data

We have seen that tough-constructions have puzzling properties: in some ways they resemble a passive, suggesting that they consist of a small, monoclusal structure. On the other hand, there is evidence for standard long-movement analyses of tough-constructions, for the identity of the matrix subject and the embedded object, and for clausal embedding. This leaves us with a paradox. How can these conflicting properties be reconciled?

Before we try to solve this paradox, we will look at previous analyses that have been proposed for tough-constructions in the next section.

3. Previous accounts

There are three groups of analyses of tough-constructions that are relevant for our purposes: monoclusal accounts specifically for German (e.g. Höhle 1978, Demske-Neumann 1994, Comrie 1997, Holl 2010), and two strands of biclausal accounts that have mainly dealt with English. Of these, the long-movement
group proposes that the gap in the embedded clause is derived by movement out of that infinitival into the matrix clause (e.g. Postal & Ross 1971, Chomsky 1981, Hicks 2003, 2009, Hartman 2011, 2012, Longenbaugh 2016, 2017), while the base-generation group proposes that an (empty) operator A’ -moves in the embedded clause and enters a predication relation with the matrix clause subject (e.g. Chomsky 1977, Browning 1987, Jones 1991, Mulder & Den Dikken 1992, Rezac 2006, Salzmann 2017, Keine & Poole 2015, 2017). This section offers a brief overview over these accounts and evaluates their advantages and disadvantages.

3.1. Monoclausal accounts

German tough-constructions have standardly been analyzed as monoclausal modal passive/ A-movement constructions (see e.g. analyses in Höhle 1978, Hawkins 1986, Demske-Neumann 1994, Comrie 1997, Wurmbrand 2001, Holl 2010). The main idea of these analyses is that sentences like (33a) are a type of passive, with the verb surfacing with non-canonical morphology (zu-infinitive plus a form of sein ‘to be’, whereas normal passives consist of the past participle and a form of werden ‘to become’)9. The ‘tough-predicate’ is merely an optional adverb that modifies the verb.

(33) a. weil der Kuchen [VP (schwer) [VP zu backen ist]]
   because the.nom cake.nom hard to bake is
   ‘because the cake is hard to bake’
   b. weil der Kuchen [VP (gerade) [VP gebacken
   because the.nom cake.nom just.now baked wurde ]]
   become.pass
   ‘because the cake has just been baked’

As we have seen in 2.1, evidence for this view comes from the parallel object promotion behavior of tough-constructions and passives, the impossibility

---

9 Although note that there are types of passive in German that also do not show canonical passive morphology, like lassen-passive in (i).

(i) dass der König den Wein hereinbringen ließ
that the.nom king.nom the.acc wine.acc bring.in let
‘that the king had someone bring in the wine’
to use a *tough*-predicate attributively, the optionality of the *tough*-adjective, and most of all the boundedness and non-$A'$-character of *tough*-movement in German.

Demske-Neumann (1994) offers a historical examination of the *tough*-construction and the modal passive in German and English. She argues that due to the loss of adverbial morphology, adjectives like *leicht* 'easy', *schwer* 'hard' etc. could no longer be identified as being used adverbially or predicatively. This led to a reanalysis of genuine *tough*-constructions in German as passives. The infinitival verb is not understood as the complement of a *tough*-predicate anymore, but reanalyzed as a complex predicate that emerged from the abstract incorporation of the infinitival verb with *sein* ‘to be’ (Demske-Neumann 1994:184f). The internal argument of the infinitive is raised to subject position to receive case, since the incorporated complex predicate cannot assign accusative case anymore. The *tough*-predicate is reanalyzed as an adverbial that is adjoined to the VP.

(34) **Passive analysis of German tough-constructions**

a. des er billich zuo breisen waer

   REL.PRON he.NOM appropriate to praise be.COND

   ‘of which it would be appropriate to praise him’

   *(Early New High German; Demske-Neumann 1994:184)*

b. 

\[
\text{TP} \\
\text{NP} \quad \text{T'}
\]

\[
\text{VP} \quad \text{T}
\]

\[
\text{AP} \quad \text{VP}
\]

\[
\text{billich} \quad \text{VP} \quad \text{V}
\]

\[
\text{NP} \quad \text{V}_2 \quad \text{V}_2 \quad \text{V}_1
\]

\[
er \quad zuo \quad breisen \quad waer
\]

Based on the typically biclausal properties discussed in section 2.2, which are
unaccounted for in such a passive-like analysis, I will dismiss a straightforward passive analysis of *tough*-constructions. Something more needs to be said. I will turn to biclausal proposals next.

3.2. Biclausal accounts

Section 2.2 presented evidence that *tough*-constructions in English and in German involve clausal embedding. The subject of these constructions is related to a gap in an embedded infinitival CP. It is the thematic argument of the embedded verb, not of the *tough*-predicate, see (35) (Chomsky 1981, Browning 1987, Brody 1993, Hicks 2003 and many others). If (35b) is grammatical, Gluckman (2017) notes that it can only be interpreted with respect to an implicit event of chopping down, suggesting that the theta role the subject receives stems from the embedded verb.

(35)  
a. This tree was easy to chop down.  
b. ??This tree was easy.  

(Gluckman 2017:5)

While this relation follows straightforwardly in monoclausal analyses (since there is only one predicate, the lexical verb), biclausal proposals have to develop different means to account for it.

3.2.1. Base generation accounts

In base generation analyses, the relation between the matrix subject and the gap in the embedded infinitival is indirect, mediated by an operator in the embedded clause. This operator is merged as the object and A′-moves to the edge of the embedded clause. There, it enters into a relation with the subject that is base-merged in the matrix clause: via semantic predication in e.g. Ross (1967), Lasnik & Fiengo (1974), Chomsky (1977, 1981), Rezac (2006), Keine & Poole (2015, 2017) and via syntactic Agree in Rezac (2006), Fleisher (2013).

Base generation analyses face a number of problems: first, they cannot account for the possibility to reconstruct into a position in the embedded clause (Hicks 2003, Longenabaugh 2017). Reconstruction is possible for anaphor binding and for scope, as the German data in (32a) and (32b) have illustrated. German and English behave entirely parallel in this respect. Reconstruction into a lower position suggests that the reconstructed element inhabited that position at some point in the derivation. This is excluded in base generation
accounts. Second, Longenbaugh (2017) points out that *tough*-constructions show properties of A-movement, and the A′-movement they exhibit seems to be more restricted than A′-movement elsewhere (it is degraded for non-subjects and impossible for subjects out of finite CPs). An analysis that relies solely on A′-movement in the embedded clause cannot offer a straightforward account for that. Third, base generation accounts violate the Theta Criterion (Hicks 2003, 2009). *Tough*-predicates do not assign theta-roles. The *tough*-subject is interpreted as a theta-argument of the embedded verb. If that role is assigned to an operator in the embedded clause, the matrix subject will be without a theta-role. It is not clear how it can receive its interpretation.

I conclude that these observations have cast sufficient doubt on the validity of base generation approaches to *tough*-constructions. I turn to long movement analyses next.

3.2.2. Long movement accounts

In long movement accounts (e.g. Postal & Ross 1971, Chomsky 1981, Hicks 2003, 2009, Hartman 2011, 2012, Longenbaugh 2016, 2017), there is a direct relation between the matrix subject and the gap: they are connected by a movement path. A DP is merged as the object of the embedded predicate and moves into matrix subject position improperly, i.e. by following an A′ movement step with an A movement step.

Evidence for this view and against base generation approaches comes from reconstruction data, as discussed above. All other biclausal properties discussed in 2.2 are compatible with this analysis.

The major disadvantage of long movement accounts is that they violate the Improper Movement constraint. This constraint was formulated (Chomsky 1973, May 1979, Chomsky 1981) to rule out derivations in which the same XP undergoes A′-movement and subsequently, A-movement, accounting for the ungrammaticality of hypherraising as in (36).

(36) a. *Minnie seems [ ___ that ___ adores custard].
    b. *Minnie scheint [ ___ dass ___ Windbeutel liebt].

However, this is exactly the configuration that is assumed in long movement analyses of *tough*-constructions: the object DP A′-moves to the edge of the embedded clause and A-moves from there into subject position. Various ways
around this violation have been proposed: Longenbaugh (2017) proposes that the relevant movement step in English *tough*-constructions from Spec,CP to Spec,vP is an instance of *composite movement*, showing both A- and A′-properties. He follows van Urk (2015) in assuming that the A/A′ distinction is due to properties of movement-triggering probes. When a head contains both A (i.e. φ)- and A′-probes, they can trigger movement together and this movement would then show mixed A/A′ behavior. Longenbaugh (2017) argues that v in English is such a head. Thus, since v does not distinguish between A- and A′-movement, movement into its specifier does not constitute an Improper Movement violation. Composite movement cannot account for *tough*-constructions in German, though. Mixed A/A′-properties for German scrambling have been proposed by Webelhuth (1989, 1992). However, it has been shown by Mahajan (1994) that German offers no evidence for these kinds of mixed A/A′-positions.

Hicks (2003, 2009) adopts a smuggling account for *tough*-movement (based on Collins 2005a,b). The *tough*-movement pivot possesses the complex structure in (37).

\[(37) \quad [\text{DP D [NP Op [DP ...]]}]\]

Hicks’ account, the outer DP moves to the embedded clause edge and the inner DP is subextracted into the matrix clause. Smuggling derivations face the problem of violating the freezing principle (according to which movement out of a moved constituent is ruled out, see section 3.2.2). Even in Abels (2007), where it is argued that the freezing principle as such is too strong, the dependencies relevant to *tough*-constructions (wh-movement feeding A-movement) are still ruled out, see Abels (2007:76f).

It thus seems that neither of these proposals can offer a solution for the improper movement problem in German *tough*-constructions. In fact, certain German *tough*-movement structures even show evidence that no movement takes place from the embedded into the matrix clause.

First, evidence from scrambling in the German middle field challenges long-movement approaches to *tough*-constructions. It has been observed (e.g. Reuland 1988, Webelhuth 1989, Geilfuß 1991, Haider 1996) that certain elements cannot be scrambled, i.e. moved into a position in the middle field. Compare (38a) – (38c). In (38a) the indefinite pronoun *was* ‘something’ is located in its base position, following the subject *einer* ‘someone’. (38b), where the indefinite
has been scrambled across the subject, is ungrammatical. This movement is allowed when the scrambled element is a full DP like in (38c).

(38) Indefinite pronouns cannot scramble

I can’t imagine ...

a. dass hier einer was begreift
   that here someone something understand
   ‘that someone gets anything here’

b. *dass hier was_i einer t_i begreift
   that here something someone understand

c. dass hier das Problem_i einer t_i begreift
   that here the problem someone understand
   ‘that someone gets the problem’

(Haider 1996, modified)

Yet, while scrambling is impossible, indefinite pronouns can be part of a tough-construction, as in (39).

(39) dass was schwer zu verstehen ist
    that something hard to understand is
    ‘that something is hard to understand’

If both short scrambling and the final movement in tough-constructions are A-movement (for scrambling see a.o. Fanselow (1990), Saito (1992), Mahajan (1994)), their divergent behavior is unexpected.

Something similar is at play in the interaction between scrambling of objects and focus. It has been observed that there is an asymmetry in the ability to scramble over a focussed or non-focussed phrase (a.o. Lenerz 1977, Stechow & Sternefeld 1988). A direct object can scramble over a focussed indirect object, see (40b) (Stechow and Sternefeld 1988:452, the focussed argument is shown in small caps, (40a) shows the base-generated word order).

(40) [Context: Who did you give the book to?]

a. dass ich dem SCHÜLER das Buch gegeben habe
   that I the.DAT student.DAT the.ACC book.ACC given have
b. dass ich das Buch dem Schüler ⟨das Buch⟩
   that I the.ACC book.ACC the.DAT student.DAT the book
   gegeben habe
   given have
   ‘that I gave the book to the student’

A focussed direct object may generally not scramble over a non-focussed indirect object, see (41).

(41) [Context: What did you give to the student?]
   a. dass ich dem Schüler das Buch gegeben habe
      that I the.DAT student.DAT the.ACC book.ACC given have
   b. ?*dass ich das Buch dem Schüler ⟨das Buch⟩
      that I the.ACC book.ACC the.DAT student.DAT the book
      gegeben habe
      given have
      ‘I gave the book to the student.’

If tough-constructions are derived by moving a DP into the matrix clause, we should expect the restriction that rules out (41b) to hold here as well. This is not what we find. (42), where the DO ends up in a position higher than the non-focussed IO, is grammatical.

(42) dass das Buch_i schwer dem Schüler ⟨das Buch⟩
    that the.NOM book.NOM hard the.DAT student.DAT the book
    zuzustellen war
to.deliver was
    ‘that the book was hard to deliver to the student’

In both of these contexts, tough-‘movement’ was allowed, while scrambling was not. This contrast suggests that the same restrictions do not hold for tough-movement and scrambling. Given the general immobility (in the middle field) of indefinite pronouns like was, it is not implausible to assume that the tough-pivot is not transported from the embedded into the matrix clause by movement.

A second problem for standard long-movement approaches is the lack of freezing effects in complex tough-‘moved’ DPs. If the step from the embedded clause into the matrix clause is movement, the tough-moved phrase should be
opaque for further extraction, according to the freezing principle (Ross 1967, Wexler & Culicover 1980, Abels 2007).

Freezing describes certain movement configurations in which an element moves out of a moved constituent. Ungrammaticality ensues when a trace in a moved item is c-commanded by its antecedent outside of that moved item (43).

(43) Freezing (Müller 1998)

\*X [\_Y \_X] \_Y

One construction to test this hypothesis with is the was-für (‘what kind’) split construction. Was-für splits are complex discontinuous DPs, as in (44).

(44) Was haben dich denn für Leute besucht?
what have you..acc mod.part for people visited
‘What kind of people have visited you?’

Was-für splits are standardly analyzed as remnant movement constructions (Abels 2003, Leu 2003, see also Leu 2008 for a detailed analysis of their internal syntax), in which a subpart of the complex DP moves out and allows the bigger DP containing the trace to move. Simplifying somewhat, they are derived as in (45).

(45) a. [DP was für Leute]
b. [XP für Leute] \_DP was (XP)]
c. [DP was (XP)] \_XP für Leute] \_DP

Remnant movement is subject to the freezing principle. Compare the structure in (46), in which scrambling and subsequent wh-movement out of the DP leads to ungrammaticality.

(46) *[DP Was (XP)] (denkst du) hat [XP für Leute] keiner (DP)
what think you has for people nobody

seen
intended: “What (do you think, what) kind of people did nobody see?”

Was-für DPs in tough-constructions should be as ungrammatical as (46). But this is not the case. In tough-constructions, a DP can be split after arriving in
the matrix clause, see (47). This violation of the freezing principle suggests that the complex DP is not transported into the matrix clause by movement\textsuperscript{10}.

\begin{Verbatim}
(47) [DP Was ⟨XP⟩] sind denn [XP für Studenten] leicht [⟨DP⟩ zu
what are MOD.PART for students easy to beeindrucken]?
impress
‘What kind of students are easy to impress?’
\end{Verbatim}

A constraint on the surface order like (43), however, cannot take into account differences in the derivational development and is thus unable to differentiate between the grammatical (47) and the ungrammatical (46). There have been numerous attempts to reduce freezing effects to a violation of the improper movement constraint (e.g. Müller 2014, Abels 2007, Grewendorf 2003, 2004).

3.3. Interim summary: theoretical paradox

We have seen that monoclausal analyses cannot account for the complex behavior of German tough-constructions. Among the biclausal approaches, base generation accounts have been dismissed due to reconstruction data, bipartisan properties of tough-movement (A and A’ ) and the violation of the Theta Criterion. It seems that long movement accounts are largely on the right track, but there is evidence against them from scrambling and freezing effects in German.

In the next section, I will show how the paradoxical properties of German tough-constructions can be resolved under a removal analysis.

\textsuperscript{10} Note that there is an alternative analysis for sentences like (47), given that we have just learned that long scrambling is possible in tough-constructions. In this alternative analysis, was is the sole target of tough-movement and the rest of the DP scrambles up, apparently long-distance, at a later stage of the derivation. This analysis predicts the possibility of leaving the remnant in the embedded clause, since scrambling is an optional operation. Curiously, this is not what we find. Sentences like (i) are ungrammatical.

\begin{Verbatim}
(i) *Was sind denn einfach für Studenten zu beeindrucken?
what are MOD.PART easy for students to impress
\end{Verbatim}

Under such an analysis, scrambling would have to be obligatory.
4. Structure Removal analysis

The crucial idea of the analysis is the removal of syntactic structure at a certain point in the derivation. German *tough*-constructions behave like they consist of both a small structure and a larger structure, because they do. Early in the derivation, they have a biclausal structure and at this point the properties typical of such a large structure are established. At a later stage, parts of the embedded clause are removed, yielding a smaller structure, and here the operations whose results suggest the small structure apply.

Concretely, I propose that the *tough*-predicate triggers deletion of the CP projection of its complement clause. This removal of the CP shell brings about an environment in which a DP in former Spec,CP can be transported into an A-position in the matrix clause. This gives us an analysis of *tough*-constructions that belongs to the family of long movement accounts (in the sense that the element in object and subject position is identical), but does not face their problem of violating the Improper Movement constraint, since no syntactic movement occurs, and thus also has the advantage of accounting for the properties of German *tough*-constructions that indicate that no movement out of the embedded clause took place.

In the rest of this section I will first outline the concept of structure removal I adopt, then go through an example derivation, and finally review how the properties of German *tough*-constructions can be accounted for under this view.

4.1. Theoretical assumptions

Long movement analyses generally face the problem of violating the Improper Movement constraint whereby an XP may not move from an A’- into an A-position. This is the exact movement that occurs in *tough*-constructions, as proposed by long-movement accounts – the *tough*-subject DP moves from Spec,CP of the embedded clause into the matrix clause subject position.

4.1.1. Structure removal

The analysis is based on the operation Remove introduced in Müller (2016b, 2017a,b). Remove is proposed to offer a systematic account for cases in which empirical evidence leads to conflicting representations which cannot plausibly
be reconciled by movement Müller (2016a). This concept of syntactic removal has forebears in operations like tree pruning (Ross 1967) and S-bar deletion (Chomsky 1981; see also Exfoliation in Pesetsky 2016).

Remove deletes structure from the derivation that has previously been built by Merge. It is imagined to be a syntactic operation that is a mirror image to Merge. As such, it has similar properties as Merge: it is triggered by $[-F_{0/2}-]$-features that are ordered on lexical heads (see Stabler 2013, Georgi 2014 among others; $[-F_{0}-]$ removes a head, $[-F_{2}-]$ a phrase); it can be external or internal; it obeys the Strict Cyclicity condition (Safir 2010, 2015 building on Chomsky 1973 in (48).

(48) **Strict Cyclicity Condition (SCC, Müller 2016a)**
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 $\alpha$.

(49) **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 $X$.

Remove can apply to heads and phrases, but in the analysis of tough-constructions I propose, only heads are the target of Remove$^{11}$.

A lexical item $X$ carrying a $[-Y_{0}-]$ feature will trigger the removal of the head $Y_{0}$ of a projection in $X$'s minimal domain. Under the assumption that a projection cannot exist without the head that projects it, the whole projection of $YP$ (i.e. $Y_{0}$, $Y'$, $YP$) will be deleted, see (50).

(50) **Merge ($X_{[\bullet Y^{\bullet}]>[−Y_{0}−]}$, YP)**

```
X'
  /
X_{[−Y_{0}−]} YP
     /
    /
   /
  ZP Y'
    /
   /
  Y_{0} WP
```

$^{11}$For Remove applying to phrases see Müller (2016b) for passives, Müller (2017a) for applicative constructions and Murphy (2015), Murphy & Müller (2016) for ellipsis in VP and TP.
In contrast to phrasal removal, where YP and every element contained in YP are deleted, the complement and specifier of the removed head Y₀ in (50) survive the head removal. However, since the connecting tissue is gone, they are temporarily unassociated from the tree. Müller (2017b) argues that the elements in the former specifier and complement of the removed head have to be reassociated into the remaining structure in a way that preserves the hierarchical and linear relations between the elements pre-removal as much as possible. This reassociation¹² is motivated independently for restructuring constructions in German and Russian (Müller 2016a, Dschaak 2017) and for multiply-filled-prefield constructions in German (Müller 2017b). The reassociation is shown in (51).

(51)  
\[
\begin{array}{c}
\text{XP} \\
\text{ZP} \quad X' \\
\text{X} \quad \text{WP}
\end{array}
\]

ZP, formerly in the specifier position of Y, has been dislocated to Spec,XP after reassociation. It has reached a higher syntactic position without having moved there.

In the following section, it is shown how German tough-constructions are derived via structure removal.

4.2. Analysis

I propose that the basic structure of tough-constructions is one of copular clauses with a complex predicate, see (52).

¹² Müller (2017b) argues that this reassociation is not an instance of Merge: unlike Merge it only applies to phrases, not heads, it does not show an external/internal distinction and it is not triggered by features; instead it is a byproduct of Remove that reintegrates material whose compositional contribution cannot be recovered otherwise.

The subject XPref is the tough-movement pivot and the predicative XP is of course the adjectival phrase. I suggest that tough-adjectives take clausal complements in German. Tough-adjectives differ from non-tough-adjectives in their ability to (optionally) remove the head of their complement. In what follows I will illustrate the details of the derivations of tough-constructions. I will equate PredP in the matrix clause with vP.

First, the tough-movement pivot dieser Kuchen ‘this cake’ is merged as the object of the embedded infinitival verb. The embedded clause is constructed as in (53).

(53) Embedded clause

The object DP A’-moves into the left periphery of the embedded clause (54).
This intermediate A′ movement can license parasitic gaps in the embedded clause, as discussed in 2.2.2. At this point in the derivation, the Improper Movement constraint comes into effect: the object DP has moved up until Spec,CP. The final landing site of the DP is an A-position. A movement step into that position would be improper and lead to a crash of the derivation.

Removal of the CP shell solves that problem.

Suppose that in a next step, the tough-predicate einfach ‘easy’ is merged. Tough-predicates have the lexical property of (optionally) removing the head of their complement. Thus, they have a feature that removes C₀, [−C₀−], in addition to the structure building feature [●C₀●]. The features are ordered intrinsically in the following way, [●C₀●] > [−C₀−], triggering the merger with a CP and subsequently, the deletion of C₀. In (55), the adjective and CP are merged and [●C₀●] is satisfied and deleted. C₀ is about to be removed.
The unassociated elements have to be reintegrated into the structure in a way that preserves the original structure as much as possible. The Strict Cycle condition in (48) provides a restriction concerning the possibilities of reintegration. Thus, it rules out a structure in which the DP is reassociated as a specifier of vP, since in that case changes to the structure would exclusively target a non-root subset of the phrase marker. The only possible structure obeying the SCC is the one in (57).\footnote{Note that the discussion surrounding (56) just serves to illustrate the remove-and-reassociate mechanism and is strictly speaking not a ‘real’ derivational step. As noted in footnote 12, Reassociation is not internal Merge and as far as I can see, no operations may interact with Reassociation. Rather, it is a part of the removal operation. Thus, the change of structure from (55) to (57) proceeds in one derivational step.}

This is the key part of the derivation. Crucially, the DP is reassociated as the specifier of the tough-predicate. Thereby, it could reach the matrix clause without having moved there. This is of course a change in the c-command relations between the DP and the adjective compared to the structure pre-removal (55). This kind of change is allowed, however, by the SCC, since the
modification affected the root node. In its re-associated position the DP is accessible for v. No intermediate movement is necessary, the DPs can move directly into Spec, vP to check the [●D●] feature.

(58) **Final movement of DP**

In the remaining part of the derivation, the DP can now move on into the prefield.

(59)

Recall that tough-constructions contrast with the semantically identical expletive constructions, as in (60).
(60)  a. dass dieser Kuchen einfach \([v_P \langle \textit{dieser Kuchen} \rangle zu backen] \) ist
that this cake easy this cake to bake is
b. dass es einfach ist, \([CP \textit{diesen Kuchen zu backen}] \)
that it easy is this cake to bake

In an expletive structure, intermediate movement of the tough-movement pivot is triggered in the same way as in the derivation of a tough-construction above. However, the intermediate clause edge cannot be removed. The DP stays in Spec,CP and the \([\bullet D \bullet] \) feature of matrix v is satisfied by the expletive pronoun es, see the structure in (61)\(^4\).

(61)  \textit{Structure of expletive constructions}

5. Consequences of the analysis

5.1. Improper Movement

As has been noted in section 3.2.2, the violation of the Improper Movement constraint is a big conceptual problem for long movement analyses. I argue that a structure removal analysis has the advantage of not running into problems with improper movement.

Various scholars have proposed to reduce the improper movement constraint to the Williams Cycle (Williams 1974, 2003), e.g. Abels (2007), Neeleman &

\(^4\)The embedded CP can be extraposed, yielding the structure in (i).

(i) dass es einfach ist, \([CP \textit{diesen Kuchen zu backen}] \)
that it easy is this cake to bake
van de Koot (2010), Bader (2011), Müller (2014). The Williams Cycle captures the idea that syntactic operations that apply to an item in a specific domain on one level of embedding may only be followed by the same operations in the same or a higher domain in a higher level of embedding. Thus, movement of XP to Spec,TP in an embedded clause may only be followed by XP moving at least to matrix Spec,TP. The movement chain may not end in a matrix position below Spec,TP. Williams (2003) formulates it as (62).

(62)  **Generalized ban on improper movement** (Williams 2003)
Given a clausal structure $X_1 > ... > X_n$ (where $X_i$ takes $X_{i+1}P$ as its complement), a movement operation that spans a matrix and an embedded clause cannot move an element from $X_j$ in the embedded clause to $X_i$ in the matrix clause, where $i < j$.

A tough-moving DP violates the condition in (62). Every time an XP moves through Spec,CP and ends up in a position lower than Spec,CP of the next higher clause, (62) is disobeyed.

The idea of a structure removal analysis is that with the removal of CP, the XP ‘forgets’ it ever stopped over there. Thus, Removal of CP repairs a violation that the derivation would otherwise incur (see also repairs by ellipsis in Merchant 2001, 2009).

5.2. Accounting for monoclausal and biclausal properties

5.2.1. Monoclausal properties

**Object promotion**  As discussed in 2.1.1, tough- and passive structures share the commonality of promoting the direct object to subject.

Dative-marked arguments cannot be subjects neither in passives, nor in tough-constructions, (63). They cannot receive nominative case or control verb agreement.

(63)  a. *dass die Kinder geholfen wurden  
        that the.NOM children helped  become.PASS
   b. *dass die Kinder leicht zu helfen sind  
        that the.NOM children easy to help are

As has been observed many times, dative arguments are opaque for certain oper-
ations on independent grounds (e.g. Chomsky 2000, Holmberg & Hróarsdóttir 2004, Alexiadou et al. 2014).

Unboundedness  We have seen that tough-movement cannot cross multiple clause boundaries in German. Recall example (6a), repeated here as (64).

(64) *dass dieses Buch schwer [Hans zu überzeugen [___ zu lesen ]] ist
read is
‘that this book was easy to convince Hans to read’

(Wurmbrand 2001:29)

This ungrammaticality is due to a violation of the ban against Improper Movement. Recall that improper movement was evaluated on the basis of the Williams cycle and f-seq. The improperness of tough-movement was resolved by removing the CP shell from the derivation and thereby removing its ‘memory’ on the f-seq.

Structure removal is very local. It may only affect the specifier or complement of the licensing head. More deeply embedded CPs cannot be removed by the tough-predicate. Thus, they contribute to an improper f-seq and thereby to an improper movement chain. It follows that only one layer of embedding is predicted to be allowed for tough-movement to cross in a structure removal analysis. In English tough-constructions, improper movement is resolved in another way, as in Longenbaugh (2017) or Hicks (2009).

No attributive use  Fleisher (2008) notes that semantically, the prenominal adjective does not modify the noun, but the postnominal infinitival clause. He calls structures like (64) clausal attributive-with-infinitive constructions (clausal AICs) and argues that they are not derived from tough-constructions. He points out the following differences between clausal AICs and tough-constructions: (i) clausal AICs allow adjectives like ‘odd’ and ‘nice’ that are not tough-adjectives (65) and (ii) the infinitival clause in AICs allows an expletive subject, which is not allowed in tough-constructions (66).

(65)   a. Bob is an odd person (for me) to see in Berkeley.
   b. *Bob is odd (for me) to see in Berkeley.  (Fleisher 2008:163)
Tough-Displacement without Movement

(66)  
\begin{align*}
  &a. \text{July is an unusual month for it to snow (in).} \\
  &b. \text{*July is unusual for it to snow (in).} \quad \text{(Fleisher 2008:163)}
\end{align*}

It seems that at least in English, tough-constructions and the attributive structures are not identical. Properties of the attributive structures do not necessarily tell us anything about tough-constructions. It is not clear how much can be inferred about tough-constructions in German. I will return to this question in section 6.

Optional tough-predicate  
In case of tough-constructions without an overt tough-predicate, as in (10), repeated as (67), I assume that a phonologically null tough-adjective with the meaning ‘necessary’ or ‘possible’ is present.

\begin{align*}
  &\text{(67) dass die Kälte jetzt } \emptyset \text{ zu spüren war} \\
  &\quad \text{that the.NOM cold now to feel was} \\
  &\quad \text{‘that it was possible to feel the cold now’ or ‘that one had to feel the cold now’}
\end{align*}

\text{(Höhle 1978)}

Since this null adjective may not occur in expletive constructions as in (68), I assume that it carries the CP-removal feature obligatorily.

\begin{align*}
  &\text{(68) *dass es die Aufgaben zu lösen ist} \\
  &\quad \text{that it the.ACC problems to solve is} \\
  &\quad \text{intended: ‘that it is possible/necessary to solve the problems’}
\end{align*}

5.2.2. Biclausal properties

Some characteristics of a biclausal structure fall out quite naturally from the proposed analysis: parasitic gaps in tough-constructions are licensed by A’-movement of the direct object in the embedded clause. The topicalization behavior discussed in 2.2.3 can be accounted for under the view that tough-constructions involve an AP rather than a VP. In that case, the structure of (22) should be more accurately portrayed as (69). Since the tough-subject is still in Spec,AP, the matrix verb is subjectless, which may be a reason for the ungrammaticality. But even if the matrix subject position is filled by an expletive, topicalization of the AP is ungrammatical (regardless of the position of the thematic tough-subject inside the AP).
The possibility of reconstruction as discussed in 2.27 follows trivially as well: since the tough-movement pivot is merged as the embedded object, it is predicted to be able to reconstruct into that position.

The other properties are discussed below.

**Double dissociation**  Section 2.2.1 illustrated that verbs do not behave uniformly with respect to passivization and tough-movement, which indicates that these two constructions are not identical. Generally speaking, the ability to passivize is a property of the verb, while tough-movement is regulated entirely by the tough-predicate and matrix v in the present analysis. The argument structure of the embedded verb is only relevant insofar as it should supply a direct object DP. Thus, all verbs that take a DP complement are predicted to occur in a tough-construction. This seems to be borne out. The verbs that cannot be part of a tough-constructions are unergative ones. Verbs like bekommen are unable to passivize for independent reasons, but can occur in a tough-construction.

**Long scrambling**  Recall that long distance scrambling is allowed in tough-constructions but not in expletive structures ( (25) vs. (26)). Under the present analysis, scrambling out of a tough-complement is only apparently long distance. If we assume that scrambling can happen only after the CP layer has been removed in the complement clause, we can maintain the generalization that scrambling in German is clause-bound. Removal of CP feeds the possibility to scramble in this case.

**Scope of negation and unstressed pronoun fronting**  In contrast to scrambling, scope relations and pronoun fronting in the embedded clause have to be established before the CP shell is removed.

Recall that negation in the embedded clause can only take scope there and fronted unstressed pronouns are grammatical in the complement clause of control-predicates and tough-predicates, but not in raising constructions. Descriptively, fronting of the pronoun (to Spec,vP) happens in the context
of a higher CP projection. Fronting happens inside the embedded CP phase. Once the tough-predicate triggers removal of CP, the phase has already been completed and the pronoun has fronted. Likewise, once the scope relations are established in the embedded clause, they are set. Removal of CP comes too late to have any effect on these phenomena.

5.3. Case-assignment on the tough-moving DP

An issue that remained unaddressed so far concerns the case of the tough-pivot. If it remains in situ in the embedded clause, it receives accusative case. If it tough-moves up, however, it is marked with nominative. These facts can be accounted for in a Dependent Case framework (Marantz 1991, McFadden 2004, Baker 2015 a.o.), in which case is assigned configurationally, i.e. it is not licensed by a functional head, but rather by the presence of another c-commanding DP.

In expletive constructions the embedded DP is in a c-command relation with the pro subject of the embedded clause and is therefore assigned accusative (e.g. by the addition of a [+inferior] feature as argued by McFadden 2004, for different implementations see Richards 2010, Preminger 2014, Levin 2015) (70).

I argue that German tough-constructions are parallel to raising-to-object constructions in Sakha (Baker & Vinokurova 2010, Levin & Preminger 2015). Under the assumption that the relevant case domain is C₀’s phasal domain, raising can feed case assignment in these constructions: a DP moves to the edge of the embedded clause, out of C₀’s phase domain, and thereby escapes case assignment. In Sakha, the subject DP of the embedded clause is now in the case domain of the matrix clause and receives accusative case under the presence of a matrix subject. A tough-moved DP also joins the matrix case domain. Since no higher DP is merged in the matrix clause, it receives nominative case (71).
6. Conclusion and outlook

This paper proposed a novel analysis of *tough*-constructions that is similar to other long-movement analyses but does not face their problem of violating the improper movement constraint. I showed that German show evidence that *tough*-constructions consist of both, a smaller and a larger structure. I argued that this contradicting evidence can be accounted for if considered to not exist simultaneously, but in succession. This can be implemented by way of removing parts of the larger structure at a certain point in the derivation, yielding the smaller structure. In the analysis I propose, *tough*-predicates can optionally possess a feature \([-C_0-]\) that removes the head of their complement. In the presence of that feature, *tough*-constructions are derived, while its absence leads to the semantically equivalent expletive structures. Reassociation after removal transports the *tough*-movement pivot DP from an A‘ position in the embedded clause into an A position in the matrix clause. Since reassociation is not (internal) Merge, it is not subject to the same constraints, like the one on improper movement and therefore cannot violate them. Technically this is implemented in changes in the buffer of the moving DP. *Tough*-constructions are complicated structures and not all puzzles could be solved. The exact relation between them and the pre-nominal clausal attributive-with-infinitive constructions in German (see (9)) remains a topic for future research.

This view on *tough*-constructions and other phenomena that exhibit evidence for conflicting representations can give insight into the range of operations that are available in UG.
References


Glückman, John (2017). The syntax and semantics of the tough-construction. Ms., UCLA.

Tough-Displacement without Movement


Leu, Thomas (2003). What moves where and what for. An account of Swiss German *was für*. Ms., New York University.
Longenbaugh, Nicholas (2017). Composite A/A′ movement: Evidence from English tough-movement. Ms. MIT.
Moltmann, Friederike (1990). Scrambling in German and the specificity effect. Ms., MIT.


Salzmann, Martin (2018). Apparent intervention in German tough-constructions and its implications. *Talk given at the 33rd Comparative Germanic Syntax Workshop, Göttingen*


Neutralization in Balinese Subjects by Structure Removal

Imke Driemel & Sören E. Tebay*

Abstract

Levin (2015) provides an account of restrictions on Subjects in Balinese object voice constructions, based on postsyntactic case licensing, local dislocation and KP-shell removal. We argue that a more formalized version of Structure Removal based on a D₀ removal feature (Müller 2017) can derive the data without further assumptions. The ban on definite subjects is explained as a neutralization effect on LF, whereas adjacency requirements follow from c-command requirements on the reintegration of syntactic structure. The NP-status of these subjects at a later point of the derivation accounts for their inability to undergo any category-specific DP-movement, thus deriving the independently observed symmetry property of the Balinese voice system (Wechsler & Arka 1998).

1. Introduction

The DP/NP distinction constitutes a frequently used explanandum in accounts of Differential Object Marking and (Pseudo-)noun Incorporation (Massam 2001, Dayal 2011, López 2012). Less prominently discussed are subjects/agents which can show similar restrictions in certain scenarios, one of which is the Balinese object voice construction (Levin 2015). Definite subjects are banned from subject in-situ positions in object voice, whereas pronouns, proper names and indefinite noun phrases are licensed, although they cannot extract from such positions. Additionally, weak quantifiers and adverbials trigger adjacency effects, as they are not allowed to appear between in-situ subjects and verbs.

We argue against the post-syntactic approach, proposed in Levin (2015), that derives the cluster of properties by local dislocation (Embick & Noyer 2001). Besides post-syntactic case licensing and Distinctness (Richards 2010), Levin employs structure removal of a KP-shell as a last resort operation. In contrast

*For comments and discussion, we would like to thank Gereon Müller. The research was supported by the DFG-funded graduate program Interaction of Grammatical Building Blocks (IGRA) and the Maria Weber Grant of the Hans Böckler Stiftung.

Structure Removal, 215–246
Andrew Murphy (ed.)
LINGUISTISCHE ARBEITSBERICHTE 94, Universität Leipzig 2019
to Levin, our proposal will make structure removal single-handedly derive almost all the Balinese in-situ properties.

We assume a structure removal feature targeting the D heads of in-situ subjects, which enters the derivation by object voice heads. These subjects can be selected as DPs in the vP cycle, but any later operation, e.g. raising to pivot, will treat them as NPs. This also leads to a neutralization of definite and indefinite DPs towards indefinite interpretations. The semantics of proper nouns and pronouns are left intact, as their D heads are not contentful.

The rest of this paper will be structured as follows. Section 2 will introduce data on Balinese noun phrases and object voice. It will also state the crucial empirical generalizations that need to be derived. Section 3 will present the approach in Levin (2015) and point out several problems. These will be solved in our approach, presented in section 4, divided into a syntactic and a semantic part, following the general introduction of the structure removal operation. We will conclude in section 5.

2. Empirical Generalizations

Like many Western Austronesian languages, Balinese exhibits a voice system in which one argument, called the pivot, is privileged in some way. The Balinese voice marking system is symmetrical. Both object voice and subject voice require two arguments, where the order subject-object is reversed in object voice. Subject voice is marked with a prefix ng- whereas object voice is morphologically unmarked.

(1) Arka (2003:106,45)

a. Nyoman ejuk polisi
   Nyoman ov.arrest police
   ‘A policeman arrested Nyoman.’

b. Polisi ng-ejuk Nyoman
   police av.arrest name
   ‘A policeman arrested Nyoman.’
2.1. Definiteness in Balinese Noun Phrases

Definiteness in Balinese is marked by the suffix \(-ne\) which can additionally occur with the postnominal determiners \(ene\) and \(ento\) (cf. 36a). Both can co-occur with possessors, shown in (2b).

(2) Arka (2003:105)

a. Celeng-e ene
   pig-DEF this
   ‘this pig’

b. Celeng tiang-e ene
   pig I-DEF this
   ‘(Lit. this my pig)/this pig of mine’

Modifiers like PPs and adjectives usually follow the noun they modify, cf. (3a) and (3b). Not much is known about the distribution of quantifiers in Balinese. Arka (2003) describes three universal quantifiers \(onya\), makejang, and \(sami\) (= all), uniformly occurring to the right of a noun, as in the examples given in (3a) and (3b) for \(onya\).

(3) Arka (2003:106)

a. Dagang celeng uli Badung ento onya
   trader pig from Badung that all
   ‘All of the pig traders from Badung’

b. ?*Dagang celeng uli Badung onya ento
   trader pig from Badung all that
   intended: ‘All of the pig traders from Badung’

They, furthermore, select for a definite noun phrase (Arka 2003:45,48), cf. (4a) vs. (4b), while also being able to float (5). These quantifiers are, thus, much in line with what has been reported about quantifier \(all\) cross-linguistically, see Giusti (1990) for German, Sportiche (1988) for French, McCloskey (2000) for Irish, or Shlonsky (1991) for Hebrew, among many others. Arka (2003:107) concludes from the distributional properties that \(onya\), makejang, and \(sami\) are merged in a rightward specifier of DP.
(4) Arka (2003:45)
   a. ?* [Jaja makejang] jemak=a
cake all ov.take=3
?‘(S)he took all cake.’
   b. Jaja-ne makejang jemak=a
cake-DEF all ov.take=3
‘(S)he took all the cake.’

(5) Arka (2003:46)
   Cerik-cerik-e meli jaja-ne ibi onya.
child-child-DEF AV.buy cake-DEF yesterday all
↝ The children bought all of the cake.
↝ All the children bought cake.

Levin (2015) provides data with respect to weak (adjectival) quantifiers. He discusses liu (=many) which, according to his fieldwork, can occur to the right or to the left of the head nouns. Whereas all examples in Arka (2003:16-17,183-184) show liu following the noun, Levin (2015) provides examples where liu precedes the noun (6).

(6) Levin (2015:76)
   a. (Liu) cicing (lui) ngugut Nyoman.
many dog many sv.bite Nyoman
‘Many dogs bit Nyoman.’
   b. Cicing-e ngugut (liu) anak cerik (liu)
dog-DEF sv.bite many person small many
‘The dog bit many children.’

Unfortunately, the data on these weak quantifiers are very limited, we do not know e.g. about the definiteness values that are possible with prenominal weak quantifiers. We also have no data on the possibility of quantifier float with weak quantifiers.
2.2. Object voice in-situ subjects

In Balinese object voice constructions, non-pivot definite DP subjects are banned, cf. the example in (8a). This contrasts with indefinite noun phrases, pronouns and proper names, as shown in (8b) and (8c).

(8) Levin (2015:77)

a. *I Wayan gugut cicing-e (ento).
   ART Wayan ov.bite dog-def that
   ‘That dog bit Wayan.’

b. Be-e daar ida/Nyoman.
   fish-def ov.eat 3/Nyoman
   ‘(S)he/Nyoman ate the fish.’

c. Nyoman sing gugut cicing.
   Nyoman neg ov.bite dog
   ‘A dog didn’t bite Nyoman.’

Additionally, two further restrictions apply to these subjects. According to Levin (2015), weakly quantifying determiners as well as adverbs cannot intervene between the in-situ subjects and the verb, cf. (9a) and (9b).

(9) Levin (2015:76)

a. Nyoman gugut (*liu) cicing (liu).
   Nyoman ov.bite many dog many
   ‘Many dogs bit Nyoman.’

---

1This set of noun types is completely unexpected from a functional perspective (Bossong 1991, Aissen 2003). An account which is e.g. based on definiteness scales will be difficult to argue for, as indefinites, pronouns, and proper names do not form a natural class to the exclusion of definite arguments.

(7) Definiteness scale according to Aissen (2003:437)

pro > pn > def > indef spec > indef non-spec
b. (Sanget-sanget) gamelan-e (sanget-sanget) depak
   very-very gamelan-DEF very-very ov.hit
   (*sanget-sanget) ia (sanget-sanget).
   very-very 3 very-ver
   ‘(S)he was hitting the gamelan really hard.’

Moreover, they cannot be dislocated from the post-verbal position: (12a) shows that non-pivot subjects cannot extrapose, (12b) shows that they cannot undergo topicalization either.²

(12) Levin (2015:71-73)
   a. Siap-e uber (cicing) ke jalan-e (*cicing)
      chicken-DEF ov.chase dog into street-DEF dog
      ‘A dog chased a chicken into the street.’ extraposition
   b. *Cicingi ia uber t_i.
      dog 3 ov.chase
      ‘A dog, it chased him/her.’ topicalization

²Levin also provides an example for illicit wh-movement, given in (10). It should be noted that this is a complex, ex-situ wh-phrase in object voice. The grammatical wh-in-situ counterpart is missing from his descriptions. A simplex wh-in-situ object voice sentence is given instead, cf. (11).

(10) Levin (2015:75)
   * [Anak cerik cen]_i be-e daar t_i? wh-movement
      person small which fish-DEF ov.eat
      ‘Which boy ate the fish?’

(11) Levin (2015:73)
   ?Montor anyar beli nyen?
      car new OV.buy who
      ‘Who bought a new car?’

There seems to be a general dispreference for wh-phrases in object voice sentences in-situ as well as ex-situ (Mike Berger, p.c.). As will become clear in the following sections, our account is able to derive the general ungrammaticality of wh-phrases in object voice sentences. Assuming that the wh-feature is situated on the D head (i.a. Ouhalla 1996), removal of D will lead either to neutralization with non-wh pronouns or to a crash at the interfaces. We will, however, not consider wh-phrases for our analysis any further.
3. Against local dislocation of Balinese in-situ subjects

Recent theories of Differential Object Marking (DOM) and Pseudo-noun Incorporation (PNI) operate under the assumption that certain types of arguments have to be exceptionally case-licensed by establishing a syntactic (Kalin 2014, 2018, Tyler 2018, Levin 2019) or post-syntactic (Levin 2015, van Urk 2019) dependency with another licenser in the clause. Levin (2015) provides a uniform analysis for object PNI in Sakha, Tamil, and Niuean as well as for in-situ subjects in Balinese and Malagasy. By situating the case licenser on the argument itself in form of a K head and allowing arguments to enter the derivation either as KPs or structurally reduced DPs/NPs, Levin creates scenarios in which non-KP arguments need to be licensed by local dislocation with V, a PF-operation that requires linear adjacency with V. This section will discuss his case-licensing approach and its application to Balinese in-situ subjects.

3.1. The post-syntactic account

The main empirical observation Levin wants to capture is that the highest nominal projection within an in-situ subject in object voice seems to require surface adjacency with the lexical verb. While Balinese generally displays free word order, in situ subjects in object voice must be postverbal and linearly adjacent to the verb, recall the data set in section 2.2. Levin redefines the case filter as a restriction on size rather than a restriction on feature valuation. All categories must be part of a complete extended projection which in case of noun phrases is a KP.

(13) Levin’s case filter (Levin 2015:46)
Noun phrases must be KPs.

(14) Structure of the noun phrase:

```
            KP
             /
            /   
           K     DP
           /
          /     
         D      NP
```
If noun phrases are merged as anything less than a KP, say a DP or an NP, then the head of the highest nominal projection must get licensed by forming a complex head with the lexical verb via adjunction as part of post-syntactic local dislocation. This step obviates the case filter since the nominal becomes part of the verbal projection. Non-KPs have to be linearly adjacent to V, as this is the only configuration where local dislocation is permitted.

While the reduction of objects to NP-size in PNI languages is the result of idiosyncratic c-selectional properties of PNI-verbs in Sakha and Tamil, reduction to DP-size for in-situ subjects in Malagasy and Balinese is triggered by a Distinctness violation of the form <KP,KP> within the vP phase. Following Richards (2010), this requirement on syntactic structures prohibits the occurrence of identical categories too close to each other.

(15) **Distinctness** (Richards 2010:5)

If a linearization statement <α, α> is generated, the derivation crashes.

Although Levin does not formalize it, he proposes that a distinctness violation is remedied by removing the subject’s KP-layer (Levin 2015:132), compare the boxed nodes in (16). The DP status requires subjects to stay merged in their in-situ positions so that they can get case-licensed by local dislocation on PF. This PF-operation is a type of adjunction which is only licit between two elements if they are linearly adjacent (Embick & Noyer 2001), see (17) where • encodes immediate precedence. Further operations to consider in (16) are the obligatory promotion of the object in object voice and the roll-up head movement (whose derivational timing is not made explicit).
(16) **Removal of KP-layer for Balinese subjects** (Levin 2015:132)

\[
\begin{array}{c}
\text{TP} \\
\text{KP} \\
\text{T} \quad \langle \text{KP} \rangle \\
\text{v}_P \\
\text{v}^\prime \\
\text{v}^\prime \\
\text{v} \\
\text{VP} \\
\end{array} \quad \Rightarrow \quad
\begin{array}{c}
\text{TP} \\
\text{KP} \\
\text{T} \\
\langle \text{KP} \rangle \\
\text{v} \\
\text{v}^\prime \\
\text{v}^\prime \\
\text{v} \\
\text{DP} \\
\text{VP} \\
\end{array}
\]

(17) **Local dislocation of the highest nominal head in Balinese** (Levin 2015:104)

\[
[T^o \ T+..+V] \bullet [DP \ D^o (\bullet NP)] \Rightarrow [T^o \ T+..+V+D]
\]

The ban on displacement, shown in (12), and intervention of adverbs (9b) follow directly from the linear adjacency requirement. In order to derive the class of nominals which are licensed as in-situ subjects, Levin has to introduce an additional assumption.

(18) **Distributional Constraint on non-KP nominals** (Levin 2015:47)

The highest overt head in the extended nominal projection, whatever it is, must be linearly adjacent to the verb.

Levin treats weak quantifiers like adjectives, presumably because they can occur pre- or post-nominally. Thus, a prenominal weak quantifier is blocked in (9a) since N as the highest overt head is not linearly adjacent to V. Definite subjects in (8a) are not licensed since the highest nominal head D is not adjacent to V due to NPs interference. Pronouns and proper names in (8b) are analyzed as (monovalent) D heads, thereby obeying linear adjacency. Finally, indefinites are licensed because, by assumption, empty D heads are invisible to the application of local dislocation.
3.2. The problems

In this section, we will point out some conceptual problems with the post-syntactic removal of a KP-layer. Contra Richards (2010), Levin has to assume that distinctness takes unpronounced lower copies into account – a move that is highly counter-intuitive for a constraint that was originally motivated to avoid contradictory linearization statements. This step is unavoidable for Levin, however, since the pivot needs to move to pivot position which is necessarily outside of the local domain where it was first merged as an argument. Moreover, lower copies cannot be visible at the point when local dislocation applies, as the intermediate copy would count as an intervener between the in-situ subject and V in (16). Consequently, Levin must assume that lower copies are visible for some PF operations but not for others.

The second issue concerns the argument choice. It seems coincidental that it has to be the subject whose KP-shell gets removed. Levin (2015:140) addresses this issue and suggests there might be an advantage of removing an unvalued case feature over a valued one, suggesting that the object in object voice is assigned case by a head within in vP. This is not in line with how he derives the pivot properties in Balinese object voice in Levin (2014b:298) where he assumes that the object’s case feature remains unvalued until T enters the derivation. He uses this assumption as a trigger for movement into pivot position, i.e. movement to Spec,TP. Since both external and internal argument are unvalued for case, there is no reason to pick one KP over another to circumvent Distinctness.

According to Levin (2014b:300), the subject voice head does assign case to the internal argument, leaving the external argument without case. The T head, therefore, attracts the subject to pivot position, as it is the only argument without case at this point in the derivation. For the sake of completeness, we show the underlying structure for subject voice clauses in (19). Note that we extrapolate here from Levin’s assumptions in Levin (2014a) and Levin (2015) since he does not discuss subject voice with respect to KPs in Levin (2015). Again, it is not obvious to us how the distinctness filter is not also violated in subject voice clauses.
For the rest of this paper, we will develop an idea that makes Levin’s trigger for surface adjacency, i.e. removal of a nominal shell, the key component of the analysis. We dispense with Distinctness and encode structure removal via a feature on the object voice head. Thus, no effects are expected in subject voice clauses. In-situ properties of the subject in object voice fall out from removal of a DP-layer, hence no reference to post-syntactic case licensing under linear adjacency with V is needed. The choice of targeting the subject over the object for structure removal follows from locality considerations.

4. Structure Removal Analysis

Structure removal, as introduced in Müller (2017), is an operation that removes structure in a syntactic derivation. It is triggered by designated $[-F^-]$ features on syntactic heads. These features come in two flavours, phrasal removal features $[-F_2^-]$ and head removal features $[-F_0^-]$. We will only be concerned with the latter, more specifically in our case the little $v$ head in Balinese object voice constructions will bear a $[-D_0^-]$ feature, i.e. a structure removal feature for a D head.

We will follow Müller (2009, 2010, 2011) in assuming that features on heads are ordered. This crucially means that a head can merge with a specifier and
later remove part of this specifier. An example of such a structure removal derivation is schematically shown in (20). A head \( Z \) is merged with a complex specifier consisting of an XP dominating a YP. Merge is triggered by \( [\bullet X \bullet] \) on \( Z \) and subsequently gets discharged. The next feature in the stack is \( [-X_0] \), which can only be discharged by removal of \( X \). Here structure removal takes place and removes \( X \), thereby taking the XP-shell with it. This step affects YP in so far as it is now the highest node contained within the specifier position of \( ZP \).

\[ \text{(20) Structure removal} \]

\[
\begin{array}{c}
\text{ZP} \\
\text{XP} \\
\text{Z} \\
\text{X \ YP}
\end{array} \quad \Rightarrow \quad \\
\begin{array}{c}
\text{ZP} \\
\text{YP} \\
\text{Z} \\
\text{[\bullet X \bullet < -X_0]} \\
\end{array}
\]

Structure removal has been applied to analyses of German complex prefields (Murphy & Müller 2016) and passives (Müller this volume); sluicing in English, German, and Serbo-Croatian (Murphy & Müller 2016), tough-movement in English and German (Schwarzer this volume), DP-NP reanalysis in Serbo-Croatian (Puškar 2015), restructuring in Russian (Dschaak 2017), clausal determiners in Gā (Korsah & Murphy 2017), applicative constructions in German (Müller 2017) and Restructuring in German (Müller 2016a). Similar approaches have been pursued under the name of Exfoliation (Pesetsky 2016, Stojković this volume) and Tree Pruning (Ross 1967, Embick 2010).

### 4.1. The syntactic side of structure removal

We claim that the Balinese data can be derived straightforwardly, under the assumption that object voice \( \nu \) heads bear a \( [-D_0] \) feature which removes the DP shell of their first merged specifiers. This naturally leads to a neutralization of definite and indefinite subjects in object voice since the distinction between them is overtly encoded by the D head. Thus, the ban on definite subjects in object voice clauses, as shown in (8a), is only apparent since they are permitted in principle but exhibit a short life cycle, i.e. they are only accessible to other operations within a narrow time window of the derivation.
We illustrate our proposal in (22). For the derivation of an object voice sentence, a vP is built by first merging an object voice v head with VP that includes an object DP and then merging a DP in the specifier of the vP. The next feature on v’s feature stack is the removal feature [−D0−]. Removing the DP shell of a head’s specifier is a strictly local application. No other D head can be found in a sufficiently local domain. This operation only applies to object voice constructions since only the object v head bears a removal feature. Other v heads do not bear such a feature. We make this assumption explicit in (21).

(21) Feature stacks for voice heads

a. vobject voice [•V≺•D≺−D0−]

b. velsewhere [•V≺•D]

(22) Syntactic tree structures before and after D₀-removal

Crucially, the remove feature can only be discharged by removing the D head of the *external* argument. This is ensured by the Strict Cycle Condition (SCC), given in (23). Removing the D head of the internal argument would violate the SCC since syntactic operations cannot affect a proper subdomain of the vP created by merging a specifier. In other words, the D head to be removed in that operation is inside the VP which is itself inside the maximal projection at that point, i.e. the vP. Removing the D head of the external argument, on the other hand, is not a problem since the specifier is not included inside any other XP but the maximal vP, which is the highest phrase at that point of the derivation.

(23) Strict Cycle Condition (Müller 2017)
Within the current XP α, a syntactic operation may not exclusively target some item δ in the domain of another XP β if β is in the domain of α.
Interestingly, this approach captures the intuitive similarity between a regular passive voice and Balinese object voice. Müller (2016b) and Müller (2017) derive the German passive by a [–D₂] on the little \( v \) head. The difference can thus be reduced to the feature on the little \( v \) head. Regular passive is triggered by structure removal of a whole phrase, i.e. a DP, whereas Balinese object voice properties result from the removal of only a D head. In both cases the external argument is demoted and thus less prominent syntactically. In Balinese this is manifested as a restriction on possible argument types and a ban on movement. For regular passive, the external argument is completely banned from its base position. It should be noted, however, that Balinese still has a regular passive voice. This is expected in our approach, since some other voice heads could still bear a [–D₂] feature, thereby exactly mirroring Müller’s approach.

As an interesting side effect, our proposal also simplifies the movement to pivot position within the Balinese voice system. Pivot properties in Austronesian voice systems have been proposed to follow either by base generating the pivot in pivot position (Spec,TP or Spec,CP depending on the account) with a co-indexed empty operator in the argument position (Pearson 2005) or by moving the pivot form argument position to the phase edge of \( vP \) over potential non-pivots, thereby making it accessible for further movement into pivot position (Rackowski 2002, Aldridge 2004, Rackowski & Richards 2005, van Urk 2015).

We derive the pivot vs. non-pivot asymmetry via the categorical distinction DP/NP. DP-movement is often argued to be derived by a categorical feature [●D●], be it for scrambling generally or for EPP-movement and object shift specifically (Chomsky 1995, Kitahara 1997, Epstein et al. 1998, Alexiadou & Anagnostopoulou 1998, Kishimoto 2000, Bailyn 2003, Müller 2010). In our account, the subject has lost its DP shell, therefore only one argument DP is left. Hence, the object DP can easily be targeted for movement to Spec,TP, bypassing the subject NP. Movement into pivot position is schematically shown in (24), constituting the underlying structure for (1a). We assume with Levin (2015:104) that head movement of V via \( v \) to T ensures that the verb precedes the in-situ subject. DP-movement requires a DP, but at the point where the movement-inducing feature enters the derivation, the highest accessible argument, i.e. the in-situ subject, is not a DP anymore. There are two possibilities to explain the dislocated structures in (12). Either topicalization/extraposition is triggered in the same fashion as movement to pivot position or the operations are triggered...
by a category neutral feature which nevertheless targets the closest argument, which again is the pivot in (24).

(24)  *Movement to pivot position in object voice clause*

Moving on to adverb placement, recall that Levin (2014a, 2015) uses adverbial distribution, shown in (9b), as one piece of evidence in favour of the surface adjacency requirement between the in-situ subject and the verb. Furthermore, he argues that this intervention is related to the voice head as there is no such ban observable for subject voice, see (25).


(Sanget-sanget) ia (sanget-sanget) nepak (sanget-sanget)
very-very 3 very-very sv.hit very-very
gamelan-e (sanget-sanget).
gamelan-DEF very-very
‘(S)he was hitting the gamelan really hard.’

Since we dispensed with a requirement for surface adjacency, our theory overgenerates with respect to adverb placement at this point. Hence, we account for the contrast between (25) and (9b) by an assumption about the adjunction site of (low manner) adverbs. If we prohibit low manner adverbs from adjoining
to \( vP \), there is no position within an object voice clause such as the one in (24) to which adverbs can adjoin and occur between the verb and the in-situ subject.\(^3\) In subject voice clauses, however, the voice head does not come with a removal feature, leaving the subject intact, i.e. a full DP. Since \([\bullet D \bullet] on T\) attracts the locally closest DP, it triggers internal merge of the subject in \( \text{spec,TP} \). This derivation leaves enough space for the adverb to occur between the verb and the subject as an adjunct to \( VP \), see (28).

(28) Movement to pivot position in subject voice clause (\& adverb adjunction)

\[ TP \]
\[ \uparrow \]
\[ DP \]
\[ \uparrow \]
\[ T'[\bullet D \bullet] \]
\[ T \]
\[ v \]
\[ \langle DP \rangle \]
\[ v' \]
\[ \langle v_{sv} \rangle \]
\[ \langle VP \rangle \]
\[ VP \]
\[ \langle V \rangle \]
\[ DP \]

\(^3\)This might be an idiosyncratic property of some Austronesian languages. Whereas adverbs act as interveners between in-situ subjects and verbs in object voice in Balinese and Malagasy, they do not do so in Squiliq Atayal and Tagalog, see (26) and (27) for the latter. We thank Mike Berger for bringing these examples to our attention.

(26) Levin (2015:85)  
Squiliq Atayal

Wal=saku\( \text{kt-an hera ne Tali.} \)
\( \text{AUX.pst=1sg.nom see-pv yesterday gen Tali} \)
'Tali saw me yesterday.'

(27) Kroeger (1991:111)  
Tagalog

Binisita\( sa=palasyo ni=Juan ang=hari. \)
\( \text{visit-perf.oV dat=palace gen=Juan nom=king} \)
'Juan visited the king in the palace.'
With respect to the class of nominals licensed as in-situ subjects in object voice, we claim that the presence of \([-D_0-]\) on the object voice head leads to neutralization of definite and indefinite subjects. Under the assumptions that definiteness is encoded as the morphosyntactic feature \([\pm \text{def}]\) on D, this distinguishing feature gets deleted within subjects of object voice heads. As shown in (29), deletion of the D head with a \([\pm \text{def}]\) feature neutralizes the definiteness distinction such that the structure and featural content of a definite and an indefinite DP are identical. We assume that this neutralization always leads to an indefinite interpretation, see section 4.2 for more details. Removal of D and subsequent neutralization explains why (i) definite noun phrases can never occur as non-pivots in object voice and (ii) indefinite nominals are licensed, recall (8a) and (8c).

(29)  **Neutralization of definite and indefinite subjects in object voice**

a. Definite noun phrases

\[
\begin{array}{c}
\text{DP} \\
\text{NP} \quad \text{D} \quad [+\text{def}] \\
\text{vP} \quad \text{v'} \\
\end{array} \\
\begin{array}{c}
\text{vP} \\
\text{NP} \quad \text{v'} \\
\text{v} \quad \text{VP} \\
\end{array} \\
\begin{array}{c}
\text{\textcolor{red}{-[D_0-]}} \\
\end{array}
\]

b. Indefinite noun phrases

\[
\begin{array}{c}
\text{DP} \\
\text{NP} \quad \text{D} \quad [-\text{def}] \\
\text{vP} \quad \text{v'} \\
\end{array} \\
\begin{array}{c}
\text{vP} \\
\text{NP} \quad \text{v'} \\
\text{v} \quad \text{VP} \\
\end{array} \\
\begin{array}{c}
\text{\textcolor{red}{-[D_0-]}} \\
\end{array}
\]

In order to extend the analysis to proper names and pronouns, we will follow Abney (1987), Szabolcsi (1987), Longobardi (1994), and many others in assuming that all arguments constitute DPs. The DP status of proper names is supported by
the fact that many languages show either optional or obligatory co-occurrence of definite determiners with proper names.

(30) Longobardi (1994:651)  \hspace{1cm} \textit{Optional D in Italian}
    (I) Gianni mi ha telefonato.
        the Gianni me has called
        ‘Gianni called me up.’

(31) Longobardi (1994:653)  \hspace{1cm} \textit{Optional D in German}
    (Der) Hans ist angekommen.
        the Hans is arrived
        ‘Hans has arrived.’

(32) Marinis (2003:71)  \hspace{1cm} \textit{Obligatory D in Greek}
    *(O) Nikos chtizi spiti sti Myknono.
        the Nikos builds house in.the Mykonos
        ‘Nikos is building a house in Mykonos.’

In line with Longobardi (1994:650), we treat the determiner co-occurring with proper names as an “expletive” article, suggesting that it contributes no semantic content. This D head is different from the D heads shown in (29), in that it does not change the type of the embedded NP. The same D head will also introduce pronouns. Déchaine & Wiltschko (2002) propose that there is no uniform syntactic category for pronouns across languages. They can be of different sizes, either DP, $\phi P$, or NP where the projections encode semantic entailment relations.

(33) \textit{Pronoun types} (Déchaine & Wiltschko 2002:410)  
\begin{center}
\begin{tikzpicture}
  \node (D) {D} ;
  \node (Phi) [below of=D] {$\Phi$};
  \node (P) [right of=Phi] {NP} ;
  \node (NP) [below of=P] {N} ;
  \node (N) [below of=NP] {N} ;

  \draw [-] (D) -- (Phi) ;
  \draw [-] (Phi) -- (P) ;
  \draw [-] (P) -- (NP) ;
  \draw [-] (NP) -- (N) ;

  \draw [-] (D) -- (Phi) ;

  \draw [-] (Phi) -- (P) ;

  \draw [-] (P) -- (N) ;

  \draw [-] (N) -- (N) ;

\end{tikzpicture}
\end{center}
Neutralization in Balinese Subjects by Structure Removal

While we do not follow Déchaine & Wiltschko (2002) in allowing for different types of pronouns in Balinese, we nevertheless adopt the elaborate nominal structure. Balinese pronouns constitute full DPs, see (34). Importantly, phi-features are encoded below the DP by $\phi_P$. Removal of the D head, thus, keeps the $\phi$-feature structure intact.⁴

(34)  Balinese pronouns:

```
DP
   \Phi P  D
   \NP  \Phi
```

Just like proper names pronouns are equipped with expletive D. Hence, they start out with a D head, but removal of this head does not actually alter the semantic interpretation. For the sake of completion, we show the structure removal derivation in (35).

(35)  Expletive D removal in object voice constructions

a. Proper names

```
vP
   DP  \nu'
   \NP  \nu
   \D  \V P
   \[-D_0-] \triangle
```

```
vP
   \NP  \nu'
   \V P
   \{D_0\} \triangle
```

⁴According to Arka (2003:166) $\phi$-features on Balinese pronouns include person and status, but not number. Gender can only be encoded on second person pronouns.
b. Pronouns

We provide the vocabulary items for the D heads in (36). While D_{[+def]} is realized overtly, D_{[-def]} and D are spelled out by an elsewhere zero exponent. D markers compete for realization via the Subset Principle (Kiparsky 1973), thus (36a) wins over (36b) in [+def] contexts, i.e. for definite noun phrases like in (8a).

(36) **Vocabulary entries**

a. /ento/ ↔ [D, +def]

b. /∅/ ↔ [D]

Finally, let us address the ban on prenominal weak quantifiers in object voice structures, recall (9a). Levin assumes weak quantifiers to be adjectives which lets them adjoin to the left or to the right of the NP in (6), the former is blocked in object voice since it prevents the NP from undergoing local dislocation with V. We also analyze weak quantifiers as adjectives, although we take adjectives to constitute heads participating in the functional projection of the nominal spine. Following the observations in Arka (2003), where modifiers as well as the overt D head ento consistently occur to the right of the head noun, we take the nominal domain to be head-final. Prenominal modifiers, then, result from DP-internal movement of NP to Spec,DP. In (37) we show the two possible underlying structures, deriving the different surface positions of liu in (6).

---

5From the descriptions in Arka (2003) it seems that both the overt determiners and/or the suffix -e can mark definiteness. Since neither Arka (2003) nor Levin (2015) are explicit about the syntactic conditions and consequences, we will assume that the overt determiners instantiate the D heads and the suffix -e is a syntactic reflex of definiteness.

6We leave the question about the trigger of this movement open for now.
Neutralization in Balinese Subjects by Structure Removal

Postnominal vs. prenominal weak quantifiers in subject voice constructions

The DP-internal movement step leads to reassociation of more than one DP-internal constituent, after structure removal has applied. In such cases, constituents need to (i) reassociate within the projection of the head that caused the removal and (ii) respect the pre-Remove c-command relations while doing so (Müller 2017:6-7). The relevant derivational steps are shown in (38).

No prenominal weak quantifiers in object voice constructions

a. Postnominal weak quantifiers
b. Prenominal weak quantifiers

Since the NP moves out of AdjP and thereby asymmetrically c-commands AdjP, it gets reassOCIATED as an outer specifier of vP while the remnant AdjP is reassociated as an inner specifier. DP-internal movement can, in this sense, be “undone” via structure removal. This is a natural consequence for languages which allow for rightward specifiers inside the nominal domain and the removal of nominal heads.

4.2. The semantic side of structure removal

As already hinted at in the previous section, the absence of the expletive D head does not affect the semantic interpretation for pronouns and proper names. Neither does its presence. We model expletive D as an identity function from entities to entities, given in (39). The removal operation in this case targets nodes which are semantically recoverable.

(39)  *Determiner for proper names* (following Longobardi 1994:650)  

\[ [D] = \lambda x.e[x] \]

Pronouns denote indices and proper names individuals. Hence, they both provide the right input for D. We sketch the semantic composition for the relevant removal contexts in (40). In line with Kratzer (1996, 2000), we assume that a verb denotes a relation between an event and their internal argument while the external argument, introduced by the functional head v, is related to the verb indirectly by its \( \theta \)-role. Verbal projections denote predicates of
events until the event variable \( \langle v \rangle \) is existentially closed off. Sentences are thus propositions based on existential claims about events.

(40) **Expletive D removal does not affect semantic composition**

a. **Proper names**

\[
\begin{array}{c}
\text{vP} \langle v, t \rangle \\
\text{DP} \langle e \rangle \\
\text{NP} \langle e \rangle \\
\text{D} \langle e, e \rangle \\
\end{array}
\quad \Rightarrow 
\begin{array}{c}
\text{vP} \langle v, t \rangle \\
\text{NP} \langle e \rangle \\
\text{VP} \\
\end{array}
\]

b. **Pronouns**

\[
\begin{array}{c}
\text{vP} \langle v, t \rangle \\
\text{DP} \langle e \rangle \\
\text{ΦP} \langle e \rangle \\
\text{D} \langle e, e \rangle \\
\end{array}
\quad \Rightarrow 
\begin{array}{c}
\text{vP} \langle v, t \rangle \\
\text{ΦP} \langle e \rangle \\
\text{VP} \\
\end{array}
\]

As is apparent from the derivations in (40), structure removal can apply without any semantic consequences if in situ subjects are pronouns or proper names as in (8b). We will now turn to the more interesting definite/indefinite cases.

\( D_{[+\text{def}]} \) and \( D_{[-\text{def}]} \), in contrast to expletive D, are crucial in ensuring argumenthood. While the former constitutes a function from properties to individuals and is only defined for singleton properties, the latter takes a property as an argument and returns an existential quantifier. The denotations are given in (41).

(41) **Definite and indefinite determiner**

a. \[ [D_{[+\text{def}]}} = \lambda P_{(e,t)} : \exists ! x[ P(x) ] . \lambda x[ P(x) ] \]

b. \[ [D_{[-\text{def}]}] = \lambda Q_{(e,t)} \lambda P_{(e,t)} : \exists x[ Q(x) \land P(x) ] \]
In contrast to (41b), the definite determiner in (41a) outputs a semantic object of type \(\langle e \rangle\), making the subject directly composable with \(v'\). Following a proposal from Landman (2000), we assume that quantifiers, which are derived with (41b), move out of the event domain in order to get interpreted. This is schematically shown in (42).

(42) Quantifiers QR out of the event domain (Landman 2000:52–53)

Now let us consider the cases where structure removal takes place. Both \(D_{[+\text{def}]}\) and \(D_{[-\text{def}]}\) take NPs of type \(\langle e, t \rangle\) as arguments. If these heads are removed, we end up with a semantic incompatibility between the NP subject and \(v'\). This type clash cannot be circumvented by the movement in (42). We therefore propose that the type clash can be avoided by a type shifting operation of the A-kind (Partee 1986a,b), see (43). Together with \(\text{THE}\) and \(\text{BE}\), (43) is argued to be a natural type shifting operation, often expected to be lexicalised across languages.

(43) A-type shift (Partee 1986a:358)

\[Q_{\langle e, t \rangle} \Rightarrow \lambda P_{\langle e, t \rangle} \cdot \exists x [Q(x) \land P(x)]\]

Compare (43) to (41b): The result of NP undergoing A-type shift is equivalent with the result of NP taken as an argument of \(D_{[-\text{def}]}\). Thus, structure removal of \(D_{[+\text{def}]}\) as well as \(D_{[-\text{def}]}\) results in an indefinite interpretation. Proper names and pronouns do not have to type shift, as they are already of the right type
Neutralization in Balinese Subjects by Structure Removal

to serve as an argument. The structure removal denotations for DP\(_{[+def]}\) and DP\(_{-def}\) subjects are given in (44) and (45), respectively.

(44) Neutralization of definite subjects in object voice constructions

a. Merge of DP\(_{[+def]}\)

\[
\begin{aligned}
 vP \langle v, t \rangle \\
 DP \langle e \rangle \\
 NP \langle e, t \rangle & \quad D \langle \langle e, t \rangle, e \rangle \\
 v' \langle e, \langle v, t \rangle \rangle \\
 VP \langle e \rangle \\
\end{aligned}
\]

b. Structure removal and A-type shift

\[
\begin{aligned}
vP \langle v, t \rangle \\
NP \langle e, \langle v, t \rangle \rangle \\
\downarrow v \\
\langle \langle e, t \rangle, t \rangle \\
\end{aligned}
\]

c. QR out of event domain

\[
\begin{aligned}
vP \langle t \rangle \\
NP \langle e, t \rangle \\
\downarrow 1 \\
\langle \langle e, t \rangle, t \rangle \\
\exists vP \langle t \rangle \\
\end{aligned}
\]

\[
\begin{aligned}
 vP \langle v, t \rangle \\
 t_1 \langle e \rangle \\
\end{aligned}
\]
(45) **Neutralization of indefinite subjects in object voice constructions**

a. Merge of DP\([-\text{def}]\]

\[
\begin{array}{c}
vP \langle v, t \rangle \\
\text{DP} \langle \langle e, t \rangle, t \rangle & vP' \langle e, \langle v, t \rangle \rangle \\
\text{NP} & D & v & \text{VP} \\
\langle e, t \rangle & \langle \langle e, t \rangle, \langle e, t \rangle, t \rangle & [-D_0-] & \triangle \\
\end{array}
\]

b. Structure removal and A-type shift

\[
\begin{array}{c}
vP \langle v, t \rangle \\
\text{NP} & vP' \langle e, \langle v, t \rangle \rangle \\
\langle e, t \rangle & \downarrow & v & \text{VP} \\
\langle \langle e, t \rangle, t \rangle & [-D_0-] & \triangle \\
\end{array}
\]

c. QR out of event domain

\[
\begin{array}{c}
vP \langle t \rangle \\
\text{NP} & vP \langle e, t \rangle \\
\langle e, t \rangle & \downarrow & \exists & vP \langle v, t \rangle \\
\langle \langle e, t \rangle, t \rangle & t_1 \langle e \rangle & vP' \langle e, \langle v, t \rangle \rangle & \triangle \\
\end{array}
\]

Since semantic objects of type \(\langle \langle e, t \rangle, t \rangle\) have to undergo QR in order to get interpreted, we expect in-situ subjects to take flexible scope with respect to other operators. This prediction is borne out. Levin (2015) reports for (8c),
repeated here in (46), that the indefinite can take scope above or below negation. Under the assumption that negation applies not lower than existential closure (von Stechow 1993, Zeijlstra 2004, Chung & Ladusaw 2004, Zimmermann 2007, Penka 2010, Swart 2016), we can derive both readings by QR targeting a position either above or below negation.

(46) Levin (2015:77)

Nyoman sing gugut cicing.
Nyoman neg ov: bite dog
‘A dog didn’t bite Nyoman.’

One remaining question concerns the availability of type shifting. We argue that structure removal can trigger $A$-type shift to ensure successful semantic composition. We can ask ourselves now why it has to be $A$ and not e.g. $THE$ which creates an argument. The latter would result in definite interpretations of in-situ subjects, contrary to fact. We argue that $THE$ (or iota for that matter) is not permitted due to Chierchia’s Blocking Principle (1998).

(47) Blocking Principle (Chierchia 1998:360)

For any type shifting operation $\tau$ and any $X$:

$*\tau(X)$

if there is a determiner $D$ such that for any set $X$ in its domain,

$D(X) = \tau(X)$

Chierchia proposes (47) in order to account for the fact that English bare arguments receive a kind interpretation not a definite or indefinite one, whereas Russian bare arguments allow for all three interpretations. The reason is that English exhibits overt lexical entries in the form of the and a/an, in contrast to Russian. Overt lexicalised determiners win over covert type shift operations. We can paraphrase the Blocking Principle along the lines of Don’t do covertly what you can do overtly.

Now if we turn to the determiner system of Balinese, we immediately see why in-situ subjects can never be definite (unless they are pronouns or proper names). Any covert type shift that could create definite readings is blocked due the presence of the overt $D_{[+\text{def}]}$ head, given in (36a). Overt definiteness, on the other hand, is blocked due to the structure removal feature on the object voice head.7

7The approach presented here is similar in its syntactic assumptions to the structure removal
5. Conclusion

We have presented a structure removal approach for in-situ subjects in Balinese object voice constructions. Removal of the DP shell through a $[-D_0-]$ feature on object voice heads neutralizes the distinction between definites and indefinites, but leaves pronouns and proper names intact. Expletive D heads on pronouns and proper names are semantically vacuous. Their removal thus does not influence the semantic derivation. For nominals, we argued that noun phrases, whose DP shell has been removed, type-shift to an indefinite meaning at LF. Type-shifting to a definite meaning is blocked, since Balinese lexicalises on an overt definiteness marker.

The core of our criticism of Levin’s (2015)’s approach is the number of assumptions needed: if you assume an explicit formulation of structure removal as a regular syntactic operation, the other facts follow. In contrast, Levin’s approach with KP removal as a last-resort operation is based on further assumptions, e.g. the visibility of unpronounced copies to some post-syntactic processes but not to other. Similarly, the choice of the argument that undergoes KP removal is largely stipulated. In our approach, these facts follow from the strict cycle condition.

account of pseudo-noun incorporation, put forward in Müller (2018). He posits a structure removal feature for D heads $[-D_0-]$ on V, in order to derive constructions where internal arguments show mixed properties of NPs and DPs. The main difference to our analysis is the position of the structure removal feature and the argument it affects. In our approach the removal feature is located on $v$ and affects the external argument, whereas in Müller (2018) it is located on V and affects the internal argument. It is worthwhile pointing out that the semantic repair operation argued for in this section does not extend to PNI languages in general. Since PNI is mostly characterized by a correlation between scope inertness and a morphosyntactic defect such as case loss or immobility, the data in (46) strongly point against diagnosing in-situ subjects as pseudo-incorporated. Another reason to doubt the PNI status comes from the fact that in-situ subjects nevertheless are able to act as binders for Principle A (Wechsler & Arka 1998:406), in contrast to what has been claimed for PNI-ed objects in Spanish (Leonetti 2004, López 2012) as well as PNI-ed subjects and objects in Turkish (Öztürk 2005, 2009), see also Driemel (subm.) for a cross-linguistic overview of binding and scope capabilities within PNI languages. Hence, we conclude that while the DP/NP distinction triggered by structure removal is a resourceful tool to derive many of the properties we see with in-situ subjects in Balinese object voice, it does not extend to scope inertness and lack of binding, as it is commonly observed with PNI languages.
References


Müller, Gereon (2010). On deriving CED effects from the PIC. *Linguistic Inquiry* 41. 35–82.


Müller, Gereon (2016b). The short life cycle of external arguments in German passive derivations. *Ms., Universität Leipzig*


Removing Clausal Determiners in Kwa

Sampson Korsah & Andrew Murphy*

Abstract
This paper argues that the distribution of the clausal determiner in two Kwa languages, Akan and Gã, is best captured by appealing to Structure Removal. The clausal determiner is found in various, heterogeneous environments that seem difficult to unify as a natural class. However, we argue that they instantiate the contexts in which an underlying DP shell on a clause fails to be removed in the derivation. This approach captures the distribution of the clausal determiner in both languages and is compatible with additional facts about coordination. Some cross-linguistic implications of this analysis will also be discussed.

1. Introduction

The category of clausal constituents has been a topic of interest since the onset of generative grammar (e.g. Lees 1960, Rosenbaum 1967, Ross 1967, Kiparsky & Kiparsky 1970, Perlmutter & Soames 1979). While the existence of clausal constituents have motivated category-specific labels such as S and C(P), it has been noticed that clauses show a number of nominal properties, leading some to suggest that they are of the category D.

We can identify two stands of research in the literature that exemplify this. First, many have argued that sentential subjects are nominal in nature (e.g. Lees 1960, Rosenbaum 1967, Ross 1967, Delahunty 1983, Davies & Dubinsky 1998, 2009, Han 2005, Hartman 2012, Halpert & Schueler 2013, Lohndal 2014, Knyazev 2016). In practice, this means that CPs are contained in an outer DP shell that allows them to appear in argument positions (1).

(1) \[ [\text{DP } [\text{CP That John will win }]] \] seems likely.

*For comments at various stages of this work, we would like to thank audiences at SelectionFest at ZAS Berlin, GLOW 41 in Budapest, as well as audiences at Leipzig University and Queen Mary University London. Particular thanks go to Klaus Abels, David Adger, Hagit Borer, Benjamin Bruening, Emily Hanink, Itamar Kastner, Idan Landau, Fabian Heck, Gereon Müller, Coppe van Urk, among others, for discussion.

Structure Removal, 247–294
Andrew Murphy (ed.)
Linguistische Arbeiten Berichte 94, Universität Leipzig 2019
Beyond this, empirical arguments have been put forward in support of this view of sentential subjects as DPs. For example, McCloskey (1991) points out that clauses seem to be specified for (singular) $\phi$-features. This can be seen by the fact that conjoined clauses trigger plural agreement on the verb, as well as licensing averbs such as equally (2).

(2) **Sentential subjects have $\phi$-features** (McCloskey 1991:564):

\[
\text{DP} \begin{array}{l}
\text{That the president will be reelected} \\
\text{and that he will be impeached}
\end{array}
\]

are equally likely at this point.

Furthermore, clausal arguments seem to need Case licensing (Vergnaud 1977/2008). As (3a) shows, sentential subjects cannot be merged in non-Case positions such as the subject position of a raising infinitive, just like ordinary argument DPs (also see Müller & Sternefeld 1995:48 for data from German). In ECM-infinitives (3b), on other hand, sentential subjects are much more acceptable, due to this being a Case position.

(3) **Sentential subjects need Case** (cf. Bošković 1995:33):

a. *It is likely \( [\text{TP} \begin{array}{l} \text{that John loves Mary} \end{array} \text{to be surprising} ] \)

b. ?I find \( [\text{TP} \begin{array}{l} \text{DP} \begin{array}{l} \text{that John loves Mary} \end{array} \end{array} \text{to be surprising} ] \)

The fact that clausal arguments are sensitive to the conditions on nominal licensing further supports their status as DPs.

Another argument comes from the licensing of ‘emphatic’ reflexives. Subject DPs such as the professor in (4a) can associate with an emphatic reflexive herself that matches in $\phi$-features. Davies & Dubinsky (2001) illustrate that the same is also true for sentential subjects (4b). This parallel further supports the assumption of a DP-layer for such clauses.

(4) **Licensing of emphatic reflexives** (Davies & Dubinsky 2001:249f.):

a. \( [\text{DP} \text{The professor} \begin{array}{l} \text{herself} \end{array} \text{offered the student sage advice.} ] \)

b. \( [\text{DP} \begin{array}{l} \text{CP} \begin{array}{l} \text{That there were 25 miles to go} \end{array} \end{array} \text{was itself enough to discourage Edwin.} ] \)

A final suggestive parallel comes from the status of sentential subjects as islands for extraction (Ross 1967). While extraction from (some) complement clauses is possible (5a), sentential subjects do not allow this (5b). It is interesting to note that the DP shell analysis in (5b) gives rise to the same structural configuration
as Complex NP islands (5c). If it is this DP+CP configuration that renders movement impossible, as in the traditional Subjacency account of Chomsky (1973, 1977), then this supports the assumption of a DP shell in (5b).

(5) **Sentential subjects as Subjacency islands** (Ross 1967:241):

a. The teacher who, it was expected [CP that the principle would fire \_

b. *The teacher who, [DP [CP that the principle would fire \_\_\_] was expected

c. *The teacher who, I heard [DP a rumour [CP that the principle would fire \_\_\_]]

A second strand of research has argued that certain types of sentential complement can be nominal (e.g. Kiparsky & Kiparsky 1970, Perlmutter & Soames 1979, Müller 1995, Müller & Sternefeld 1995, Adger & Quer 2001, Potts 2002, Hankamer & Mikkelsen 2012, Takahashi 2010, Hartman 2012, Kastner 2015, Pietraszko 2019). In particular, one factor that has been deemed relevant for the nominal status of a clausal object is whether the predicate that selects it is factive or not. On this view, the clausal complement to a factive verb such as **regret** in (6) is nominal, i.e. contains DP-shell structure.

(6) We all regret [DP [CP that John won the race ]] (...#but he didn’t)

Kiparsky & Kiparsky (1970) in particular argued that factivity was central to the nominal status of a given sentential complement. One piece of evidence they point to in support of this is that there is no raising-to-object (R-to-O) from the complement of factive verbs such as **regret** (7a), in contrast to non-factive predicates such as **believe** (7b). Kiparsky & Kiparsky (1970) attribute this to the status of factive complements as Complex NP Islands, as described for (6).

(7) **No R-to-O from factive complements** (Kiparsky & Kiparsky 1970:16of.):

a. *He regrets [VP Bacon, [DP [CP \_\_\_\_\_\_\_\_, to be the real author ]]]

b. He believes [VP Bacon, [CP \_\_\_\_\_\_\_, to be the real author ]]

A further argument they present comes from the possible pro-forms that can correspond to clausal arguments. Non-factive complements can take either the DP pro-form *it, or the CP pro-form *so (8a). Complements to factive verbs such as **regret**, on the other hand, can only take the DP pro-form *it, suggesting that they are strictly nominal in nature.
Factive complements take DP pro-forms (Kiparsky & Kiparsky 1970:166):

a. John supposed [CP that Bill had done it ] and Mary supposed [DP it ] / [CP so ] too
b. John regretted [DP that Bill had done it ], and Mary regretted [DP it ] / *[CP so ] too

A similar diagnostic is provided by Kastner (2015), who pointed out the contrast in (9). In (9a), coordination of a nominal and clausal argument of the non-factive predicate claim fails, presumably due to a categorial mismatch (however, see Sag et al. 1985 and section 4.3 for discussion of some exceptions). With complements to factive verbs like deny (9b), however, Kastner (2015) notes that this mismatch does not seem to arise. Again, this follows if the sentential complement in (9) is actually a DP.

Factive complements can be coordinated with DPs (Kastner 2015:173):

a. *John claimed [DP responsibility ] and [CP that the building collapsed ]
b. ?John denied [DP the allegations ] and [DP that the building collapsed ]

A slightly different argument for the status of some object CPs comes from the observation by Postal (1994) that the trace of moved complement clauses have the distribution of nominals. First, notice that a CP cannot typically occur as the complement of a preposition (10a). If the CP is extracted, however, the structure is well-formed (10b). This seems to suggest that, at least when they are to be moved, clausal complements can be nominal (see Hartman 2012).

---

1Ormazabal (2005:104) mentions a potential challenge for the view that (factive) CPs are nominal, namely that that clausal complements cannot combine with the expletive correlate there (ia). As shown by Postal & Pullum (1988:643), they must instead take it (ib).

(i)  a. *They never mentioned there to the candidate [CP that the job was poorly paid ]
    b. They never mentioned it to the candidate [CP that the job was poorly paid ]

If the there is an expletive associated with DPs, then this could be viewed as an argument against the status of these clauses as DPs. However, much depends on the correct analysis of such cases, which might involve extraposition from an underlying single constituent (see e.g. Pullum 1979:47f.).

2It is important to mention, however, that an alternative account of (9) has been proposed, whereby the fronted CP is a ‘satellite’ associated to a nominal null operator that moves from the complement position of the preposition (e.g. Emonds 1972, Koster 1978, Stowell 1981, Alrenga 2005, Moulton 2009, 2013). For discussion of an alternative, see section 4.3.
(10) *CP gaps have nominal distribution* (Postal 1994:70):

a. *I couldn’t convince Frank [PP of [CP that Sonia was very competent ]] ]

b. [DP [CP That Sonia was very competent ]] I couldn’t convince Frank [PP of ___DP ]

Thus, it seems that there a large body of research that shows that clausal arguments at the very least *can* be DPs. This has repercussions for the theory of selection, however. For example, now we must envisage (perhaps optional) c-selectional features for C and D respectively, depending on the type of clause that is selected (11).

(11) *Optional subcategorization for clause-embedding:*

\[
\begin{array}{c}
\text{VP} \\
\text{V} \\
\text{[\text{C}]} \\
\text{C} \\
\text{TP} \\
\end{array} \\
\begin{array}{c}
\text{VP} \\
\text{V} \\
\text{[\text{D}]} \\
\text{D} \\
\text{CP} \\
\text{TP} \\
\end{array}
\]

While this is one way to capture the seemingly necessarily optionality of clausal selection, another is to adopt what we might call the *Radical DP Shell Hypothesis* in (12).


All finite clauses are DPs ([DP D [CP ... ]]).

On this view, there is only c-selection by verbs for the category D. This means that all clausal arguments are born as DPs. Any genuine CP complements must then be derived in the syntax. Indeed, this is what Kiparsky & Kiparsky (1970) argued for non-factive complements, namely that the nominal shell was removed by means of a ‘Fact-deletion’ operation. However, in general, one might wonder whether it is possible to distinguish this ‘derivational’ view of the DP/CP distinction from the ‘inherent’ view in (11).

In this paper, we will present an argument in favour of the former view that
all clausal arguments originate with DP a shell and may lose this shell in the
course of the derivation. The evidence for this comes from the distribution
of the so-called clausal determiner in two Kwa languages spoken in Ghana,
Gã and Akan. As the name suggests, clausal determiners are determiner-like
elements that are found on clausal arguments. Thus, they seem to provide
overt evidence for the kind of DP-shell on clauses that has otherwise been
motivated based on more indirect evidence for languages like English, as we
saw above. Examples of clausal determiners in the respective languages are
given in (13). In Gã, the determiner le, which is also found with nouns, is
obligatory in the complement clause in (13a). Similarly, the determiner nó in
Akan can optionally be realized on clauses, as in (13b).

(13) a. Mënì ni Kwei yése [CP akë Yëmo he ___1 *(le) ]?
    what foc Kwei realise that Yëmo buy CD
    ‘What did Kwei realise that Yemo bought?’ (Gã)

b. Déënì na Kofi káé [CP se Ám’má kítá ___1 (nó) ]?
    what foc Kofi remember that Ama hold CD
    ‘What does Kofi remember that Ama is holding?’ (Akan)

As we will see, the clausal determiners (CD) in Akan and Gã show a complex
distribution that is difficult to capture neatly under the inherent view that
clauses are sometimes DPs and sometimes CPs. Instead, we will show that a
syntactic process of Structure Removal that eliminates the DP projection in a
[DP [CP ... ]] configuration will provide a general and explanatory account of
the contexts in which the CD can occur and those in which it cannot.

In section 2, we elaborate on the nature and distribution of the CD in the
two languages, arguing that it is a genuine D head selecting a CP. We show that
the various contexts in which it arises cannot easily be captured as a natural
class. Section 3 argues that a unified account can be provided by Structure
Removal, that is, the contexts in which the CD can or must surface are those
contexts in which Structure Removal fails to apply. Finally, section 4 extends
the Structure Removal analysis to related puzzles about complementation in
other languages, namely complementizer-drop in English, extraction from DPs
in German and nominal properties of moved clauses.
2. Clausal determiners in Kwa

2.1. The CD in Kwa is not related to factivity

As mentioned earlier, a common idea is that the presence or absence of a DP-shell on clauses is related to factivity. If clausal determiners are the realization of such a nominal layer, we might expect them to also be sensitive to this property. Indeed, this is precisely what Kastner (2015) claims for Hebrew. A complement clause without a clausal determiner is not subject to a factivity restriction (14a), as shown by the bracket continuation. If this clause contains the clausal determiner ze, however, the CP must be interpreted as factive (14b).

(14) **Clausal determiner tracks factivity in Hebrew** (Kastner 2015:160):

a. hu hisbir [\[CP še-ha-binyan karas\] (aval hu lo he explained COMP-the-building collapsed but he NEG be’emet karas)
really collapsed 'He explained that the building collapsed (but it didn’t).

b. hu hisbir et [DP ze [\[CP še-ha-binyan karas\]]
he explained ACC this COMP-the-building collapsed (#aval hu lo be’emet karas)
but he NEG really collapsed 'He explained that the fact that building collapsed (#but it didn’t).

However, it is not universally the case that clausal determiners mirror factivity. In Greek, clausal arguments take a clausal determiner to (15).

(15) **Clausal determiner in Greek** (Roussou 1991:78):

\[DP to [CP oti ehis filus ]] simeni pola
the.NOM that have.2sg friends means much
‘That you have friends means a lot.’

However, Moulton (2017) points out that the presence of a CD in Greek does not entail the truth of that clause, as demonstrated by the example in (16).

(16) **Greek sentential subjects do not have to be factive** (Moulton 2017:296):

\[DP to [CP oti ine plusios ]] ine psema
the.NOM that is.3sg rich is lie
‘That he is rich is a lie.’
The situation is similar in Akan and Gã. The Gã examples in (17) show that there is a distinction between factive and non-factive clause-embedding predicates, i.e. *jwɛŋ* ‘think’ (17a) vs. *yɔse* ‘realize’ (17b). However, when the clause is interpreted as factive (17b), the CD still cannot surface.

(17) **Factivity does not correlative with clausal determiner in Gã:**

a. Kwei *jwɛŋ* [ _CP akɛ Yɛmo he ɣɛlɛ (*lɛ) _ ] ... Shi asomoa
   Kwei think that Yemo buy yam CD but apparently
   e-je-ee anɔakwale
   3SG-be-NEG true
   ‘Kwei thought that Yemo bought yam … but apparently it is not true.’

b. Kwei *yɔse* [ _CP akɛ Yɛmo he ɣɛlɛ (*lɛ) _ ] ... #Shi asomoa
   Kwei realize that Yemo buy yam CD but apparently
   e-je-ee anɔakwale
   3SG-be-NEG true
   ‘Kwei realized that Yemo bought yam … #but apparently it is not true.’

Furthermore, in cases where the CD is obligatory, such as the sentential subject in (18), there is no factivity restriction, analogous to the Greek example in (16).

(18) **Sentential subjects in Gã do not have to be factive:**

[ _CP Akɛ Kwei na Yɛmo *(lɛ) _ ] e-je-ee anɔakwale
   that Kwei see Yemo CD 3SG-be-NEG true
   ‘That Kwei saw Yemo is not true.’

From this, we conclude that the distribution of the CD in Akan and Gã cannot be reduced to factivity, unlike what has been claimed for some other languages.

2.2. The distribution of the clausal determiner

In this section, we turn to the distribution of the clausal determiner in Akan/Gã. As we will see, it has a rather complex distribution that involves a rather varied set of contexts.

2.2.1. Extraction from complement clauses

The first context for the clausal determiner in Akan/Gã is on complement clauses from which extraction has taken place. Starting with Gã, the embedded clause
without extraction in (19a) cannot have the CD le. As soon as we extraction from this clause though, the CD becomes obligatory (19b).

(19) *Extraction from verb complement clauses in Gã:*

a. Kwei yose 
   [CP akɛ Yɛmo he yele (*le)]
   Kwei realise that Yɛmo buy yam CD
   ‘Kwei realised that Yemo bought yam.’

b. Mɛni, ni Kwei yose 
   [CP akɛ Yɛmo he ___i (*le)]?
   what foc Kwei realise that Yɛmo buy CD
   ‘What did Kwei realise that Yemo bought?’

The same can be seen in Akan. Embedded clauses without extraction cannot take the CD nó (20a), whereas those containing an extraction site must (20b).

(20) *Extraction verb complement clauses in Akan:*

a. Kofi kaé
   [CP se Ám’má kita bayɛɛ (*nó)]
   Kofi remember that Ama hold yam CD
   ‘Kofi remembers that Ama is holding a yam.’

b. Déɛn₁, na Kofi ké
   [CP se Ám’má kitá ___₁ (nó)]?
   what foc Kofi remember that Ama hold CD
   ‘What does Kofi remember that Ama is holding?’

This pattern also holds for clausal complements to nouns. In the Gã example in (21a), the CP is the complement of the noun hhuuuhu and cannot have the clausal determiner. If we extract from this complement clause, however, then the CD becomes obligatory (21b).

(21) *Extraction from noun complement clauses in Gã:*

a. Wɔ-nu [DP hhuuuhu ko [CP akɛ gbɔ-i le e-ba
   3PL-hear rumour INDEF that visitor-PL DEF 3SG-come
   (*le) ]]
   CD
   ‘We heard a rumor that the visitors have come.’

b. [DP Gbɔ-i le₁ [CP ní wɔ-nu [DP hhuuuhu ko [CP akɛ
   visitor-PL DEF REL 3PL-hear rumour INDEF that
   amɛ₁-ba * (le) ]]]] nɛ
   3PL-come CD this
   ‘These are the visitors that we heard the rumor that they have arrived.’
This same pattern is shown for the analogous Akan example in (22).

(22) Extraction from noun complement clauses in Akan (Salzmann 2017:194):
   a. Me-te-e [DP atetêsêm bí [CP sê Kofi fe-e  n’ 1sg-hear-pst rumour indef that Kofi kiss-pst 3sg.poss anó (*nó)][mouth cd]
      'I heard a rumour that Kofi kissed her (mouth).'
   b. [DP ɔbaá nó_1 [CP áa me-tê-e [DP atetêsêm bí [CP sê woman the rel 1sg-hear-pst rumour indef that Kofi fé-e  n’_1 anó ]] *(nó)]]
      Kofi kiss-pst 3sg.poss mouth cd
      'The woman that I heard a rumour that Kofi kissed her.'

In sum, we see that complement clauses to both noun and verbs can only surface with the clausal determiner if they have been extracted from.3

2.2.2. Sentential subjects

Sentential subjects must obligatorily occur with the clausal determiner in both Gã (23a) and Akan (23b).

(23) a. [CP Åkɛ amlalo le tse too nɔ *(lɛ) ] fee maŋ-bii
   comp government def tear tax top cd do country-people
   le naakpɛɛ
   def wonder
   'That the government reduced taxes surprised the people.' (Gã)
   b. [CP Sɛ Kofi á-si dán *(nó) ] má-a Á’má ání gyé-eé
   that Kofi perf-build house cd give-past Ama eye collect
   'That Kofi has built a house made Ama happy.'
   (Akan; cf. Hein 2017:10)

---

3The CD is also found with monoclausal extraction. We put this aside for now, and return to it in section 3.3.1.
2.2.3. Relative clauses

The same is true for relative clauses modifying nouns. The Akan relative clause in (24a) obligatorily occurs with the clausal determiner nó. This can also been in Gã (24b).

\[ DP Õbáá₁ \ (nó) \ [CP áa \ ɾ₁-wářé-e \ Kofi \ *(nó) ] \ fi \]
woman \ definit \ relative \ 3SG marry-pst \ Kofi \ clausal \ determiner \ nó \ fi \]

Aburi.

Aburi ‘The woman who married Kofi is from Aburi.’

(24) a. \[ DP Õbáá₁ \ (nó) \ [CP áa \ ɾ₁-wářé-e \ Kofi \ *(nó) ] \ fi \]
woman \ definit \ relative \ 3SG marry-pst \ Kofi \ clausal \ determiner \ nó \ fi \]

(24) b. \[ DP Aṭaade₁ \ lɛ \ [CP ní \ ọ-he \ ____ ¹ \ *(lɛ) ] \ yɛ \ bie \]
dress \ definit \ relative \ 3SG buy \ clausal \ determiner \ lɛ \ bie \]

‘The dress that you bought is here.’

2.2.4. Preposed conditional clauses (in Gã)

Finally, Gã shows a positional asymmetry with regard to the clausal determiner on conditional clauses. If the conditional appears clause-finally, then the clausal determiner lɛ is not possible (25a). If the conditional clause precedes the consequent, however, then the clausal determiner is obligatory (25b).

(25) Conditional clauses in Gã (Kropp Dakubu 1992:9):

a. \[ CP Kɛ(jí) \ o-ba \ (*lɛ) ] \ m-á-yá \]
1SG FUT go \ conditional \ 2SG come \ clausal \ determiner \ lɛ \ m-á-yá \]

‘If you come, I will go.’

b. \[ CP Kɛ(jí) \ o-ba \ (*lɛ) ] \ m-á-yá \]
conditional \ 2SG come \ clausal \ determiner \ lɛ \ m-á-yá \]

‘If you come, I will go.’

2.2.5. Interim summary

We can therefore summarize the contexts for the clausal determiner as follows:
As is clear from the preceding discussion, the CD occurs in a rather heterogeneous set of environments. Thus, the relevant question at this point is what is the feature that unifies these contexts as a natural class?

2.3. What is the clausal determiner?

Before we turn to the unifying property behind the distribution of the CD in Akan and Gã, let us first clarify the syntactic status of the clausal determiner. In the previous literature on the CD in Kwa, there have been two claims that we will dispute in what follows: (i) the clausal determiner is a functional category other than D (Lefebvre 1992a,b, Aboh 2004), (ii) the clausal determiner is a ‘low’ (νP) adjunct (Renans 2016, Grubic & Renans 2017).

Instead, we will show that the clausal determiner in Akan and Gã is best analyzed as a D⁰ head that selects a CP directly, as shown in (27) (also see Pietraszko 2019 on Ndbele).

\[
(27) \quad [\text{DP} \ [\text{CP} \ldots ] \ [D^0 \ \text{nó/lé} ] ]
\]

2.3.1. The clausal determiner is a D⁰ head

For some languages, it has been claimed that clausal determiner encodes ‘familiarity’, for example in Haitian Creole and Fon(gbe) by Lefebvre (1992a) and also in Gungbe relative clauses (28) (Aboh 2005).

\[\text{(28) Clausal determiner encodes familiarity in Gungbe (Aboh 2005:270f.):}\]

a. Kofi xɔ \quad [D_P \ \text{agásá} \ [\text{CP} \ dẽ \ mĩ \ lɛ] ]

Kofi buy crab \text{ rel} we catch
‘Kofi bought the crab that we caught.’
b. Kofi xɔ [DP agásá [CP dē mí wlé ] lɔ ]
   Kofi buy crab REL we catch DET
   'Kofi bought the aforementioned crab that we caught.'

For this reason, the CD has been analyzed as a discourse/agreement-related functional head in some higher periphery of the clause, rather than an actual determiner (Agr⁰ by Lefebvre 1992a and Top⁰ by Aboh 2005). We argue that this characterization of the CD is not correct for Akan and Gã, however, and that the CD is a genuine determiner, i.e. D⁰. This can be demonstrated using the haplology effect identified by Saah (1994). Saah proposes the rule in (29), which deletes a sequence of adjacent, homophonous D⁰ elements.

(29) **Pleonastic Determiner Deletion Rule** (Saah 1994:155):
   \[ [D \text{ no }] [D \text{ no }] \Rightarrow \emptyset [D \text{ no }] \]

In Akan, a sequence of the resumptive pronoun nó and CD nó, which have identical forms, is tolerated (30a). Since the resumptive pronoun is not a D⁰ element, the rule in (29) does not apply. However, if the CD is preceded by a genuine definite determiner belonging to an object DP, for example, then the context for the rule in (29) is met and one of them must be deleted (30b).

(30) **Haplology effect in Akan** (Saah 1994:151f.):
   a. [DP Abrofrá [CP áa Kofi hú-u nó₁ *(nó)] ] á-ba
      child REL Kofi saw-PAST 3SG CD PFV-come
      'The child that Kofi saw has come.'
   b. [DP Onípá [CP áa ɔ₁-tó-o ] [DP ndwóm nó [*(nó)]]]
      person REL 3SG-throw-PAST SONG DEF CD
      ye-ɛ adé
do-PST something
      'The person who sang the song did well.'

Korsah (2017) shows that the same is true for Gã, where object resumptives do not lead to haplological dissimilation (31a), but object determiners do (31b).

(31) **Haplology effect in Gã** (Korsah 2017:157):
   a. Námɔ₁ ni Osa le [CP ákɛ Taki tsɛ lɛ₁ *(lɛ) ] ?
      who FOC Osa know that Taki call 3SG CD
'Who does Osa know that Taki called?'
b. Taki ni Osa le [CP áké e₁-hé [DP tsòné lé ] (*lé ) ] 
   Taki foc Osa know that 3SG-buy vehicle DEF CD
   'It is Taki that Osa knows bought the vehicle.'

Since the anti-haplology rule in (29) cares about the syntactic, as well as phonological properties of the adjacent items, it provides a useful diagnostic for the clausal determiner as a genuine determiner. We therefore conclude from this that the CD in both languages is the realization of the head of a DP shell dominating the clause.

However, this may not be true for all CDs cross-linguistically. Somewhat revealingly, the CD in Fon, which is argued by Lefebvre (1992a) not be a determiner, does not give rise to a haplological dissimilation:

(32) No haplology with clausal determiners in Fon (Lefebvre 1992a:138):
   Súnù ñ gbá [DP móto ñ ] ñ 
   man DET destroy car DET DET
   'The man destroyed the car, as we expected/knew.'

This is compatible with the claim that the final ñ in (32) is indeed not a D⁰ element, but perhaps a functional head of some other category. The consequence is that ‘clausal determiners’ may not form a homogeneous class cross-linguistically, however we will not explore this issue further.

2.3.2. The clausal determiner is high

Now we have determined what kind of element the CD is in the relevant languages, the question remains where exactly it is located. Since it typically occurs clause-finally, it is difficult to diagnose precisely how high the CD attaches to the clause in question. There have been two major proposal in the literature to so far. Saah (1994) argued that the CD occupies a relatively high position, attaching to the clause directly (33). Others have argued, mainly on
Removing Clausal Determiners in Kwa

(33) *High analysis of the clausal determiner* (Saah 1994:162):

```
   DP
  /\   /
CP   D
  /\   /
C    TP
  /\   /
T    vP
     /\...
```

(34) *Low analysis of the clausal determiner* (Renans 2016:175):

```
   TP
  /\   /
vP   D
  /\   /
T    vP
     /\...
```

We will adopt the analysis in (33), assuming that D selects the CP.\(^5\) One reason for rejecting the ‘low’ analysis in (34) is that the clausal determiner does not

---

\(^5\)It is interesting to note that this structure superficially seems to violate the *Final-over-Final Condition* (i), which prohibits a head-final phrase embedding a head-initial phrases within the same extended projection.

(i) *Final-over-Final Condition* (Biberauer et al. 2014:171; Sheehan et al. 2017):

A head-final phrase \(\alpha P\) cannot dominate a head-initial phrase \(\beta P\), where \(\alpha\) and \(\beta\) are heads in the same extended projection.

\*[\(\beta P [\alpha P \alpha P ] \beta\)]

This suggests that the clausal determiner does not belong to the extended projection of the verb (cf. Pietraszko 2019). Event nominalization in Akan appears to behave differently, when a head-final nominalizer attaches to a head-initial VP, the order of verb and object must be reversed to avoid a FOFC violating sequence ([\(nP [VP V O \alpha] \Rightarrow [nP [VP O V \beta] \Rightarrow\)])}, see Hein & Murphy (2018) for discussion.
behave like a low, vP-level adjunct with regard to pro-drop. First, consider the fact that inanimate pronouns are typically dropped in Gã (35b) (Korsah 2017).

(35) No inanimate object pronouns in Gã:
   a. Momo shá [DP mfonírí lé ]
      Momo snap photo DEF
      ‘Momo took the photo.’
   b. Momo shá (*lé)
      Momo snap 3SG
      ‘Momo took it.’

However, this effect is obviated in the presence of a clause-final adverb (36).

(36) Clause-final VP-adverbs trigger overt inanimate pronouns:
   a. Momo shá *(lé) [Adv oyá ]
      Momo snap 3SG quickly
      ‘Momo took it quickly.’
   b. Momo shá *(lé) [PP yɛ La]
      Momo snap 3SG at La
      ‘Momo took it in La.’

It is important to establish that this is not just an effect of clause-finality, since the sentence-final imperative marker mmó in Gã (cf. Kropp Dakubu 2008:96) does not block inanimate pro-drop (37b).

(37) a. Momo, shá [DP mfonírí lé ] mmó!
     Momo snap picture DEF IMP
     ‘Momo, take the photo!’
   b. Momo, shá (*lé) mmó!
     Momo snap 3SG IMP
     ‘Momo, take it!’

(ii) a. Kofi [VP á-si dán ]
     Kofi prf-build house
     ‘Kofi has built a house.’
   b. [VP dán sí ]-é na Kofi á-yó
     house build -NMLZ FOC Kofi prf-do
     ‘Kofi has BUILT A HOUSE. (not e.g. bought a boat)’
The imperative marker *mmó* in (37b) does not block pro-drop since it presumably occupies a higher left-peripheral position such as the head of ForceP. Thus, it is only a certain kind of clause-final element that blocks the regular process of inanimate pro-drop, namely low, *vP*-level adverbs. This gives rise to the descriptive generalization in (38), which we will utilize as a diagnostic for attachment of rightward elements (see Korsah 2017 for a suggestion on how it can be derived).

(38) pro-drop generalization (Saah 1994, Osam 1996, Korsah 2017):
Low, *vP*-level adverbs block deletion of inanimate pronouns.

Importantly, the clausal determiner behaves like the imperative marker in not blocking object drop, as shown by (39).

(39) Tsóñé ́lé₁ ni Osa le [CP áké Taki hé (.stopPropagation) *lé₁] vehicle DEF FOC Osakinó that Taki buym₃SG CD
‘It is the car that Osa knows that Taki bought.’

This follows if the CD occupies a high position, just like imperative *mmó*:

(40) The assumption that *mmó* is indeed higher than *vP*-adverbs is confirmed by the fact that it must follow adverbs (ia).

(i) a. Momo, shá ́lé₁ oyá *mmó!*
Momo snap 3SG quickly IMP
‘Momo, take it quickly!’

b. *Momo, shá ́lé₁ mmó oyá!*
Momo snap 3SG IMP quickly
‘Momo, take it quickly!’
An analogous argument can be made for Akan. As (41b) shows, inanimate pronouns are also ordinarily dropped in object position.

(41) a. Kofí di [DP aduane nó ]
    Kofi eat  food  DEF
    'Kofi eats the food.'

    b. Kofí di (*nó)
    Kofi eat  3SG
    'Kofi eats it.'

As in Gã, this process is blocked by clause-final adverbs in Akan (42a). Importantly, high clause-final functional heads such as the question particle anaa do not obviate inanimate pro-drop (42b).

(42) a. Kofí di *(nó) anɔpá
    Kofi eat  3SG morning
    'Kofi eats it in the morning.'

    b. Kofí di (*nó) anaa ?
    Kofi eat  3SG Q
    'Does Kofi eat it?'

Recall from example (13b), repeated below as (43), that the clausal determiner in Akan patterns with the question particle rather than low, vP adverbs in that it does not block pro-drop of the resumptive pronoun in (43) (note that resumptive pronouns and anaphoric pronouns have the same distributal properties; see Korsah & Murphy to appear).

(43) Déen₁ na Kofí káé  [CP sɛ Ām'má kítá (*nó) (nó ) ] ?
    what  FOC Kofi remember  that Ama hold  3SG CD
    'What does Kofi remember that Ama is holding?'  (Akan)

From this, we conclude that the CD in both Akan and Gã is a D⁰ that attaches at the CP-level.

3. Analysis

Let us now see how we can unify the contexts in which we find the clausal determiner in Akan and Gã. We will show that the distribution of the CD in Kwa can be captured by assuming that the DP shell is present underlyingly, but ordinarily deleted in the complement position of verbs (and nouns). This is shown schematically in (44).
On this view, the contexts in which the CD surfaces are the contexts in which Removal of DP fails, for principled reasons. Thus, we have to appeal to what might be the opposite of our first intuition. Rather than asking when a DP shell can be added to a CP, we have to ask when it cannot be removed. To this end, we will define the process of Structure Removal in more detail.

3.1. Structure Removal (Müller 2017, et seq)

In order to implement the Remove analysis, we will adopt the feature-driven approach to syntactic deletion has been recently proposed by Müller (2017, 2018, to appear, this volumea, this volumeb) (also see Murphy 2016, Zyman 2018). As summarized in (45), this approach expands the ontology of formal features to include features of the type $[-X-]$, which trigger removal of projections of category $X$.

(45) Structure Removal (Müller 2017, 2018):

As well as structure-building features $[\bullet X \bullet]$ that trigger Merge of a category $X$, there are structure-destroying features $[-X-]$ that trigger Removal of an existing projection $X$ in the structure.

To take an abstract example, if a head $Y$ selecting a phrase $XP$ bears a Remove feature $[-X-]$, then this feature is checked by deleting the $XP$ projection (46). Consequently, $ZP$ becomes the new complement of $Y$.

---

7 Müller (2017) suggests further diacritics on this feature to distinguish whether heads or entire phrases are removed, but for present purposes we will only deal with one type of Structure Removal, namely where a head and its associated projection are removed.
(46) Feature-driven Structure Removal:

\[
\begin{align*}
\text{YP} & \quad \rightarrow \\
\text{Y} & \quad \text{XP} \\
\text{[−X−]} & \quad \text{X} \\
\text{Z} & \quad \text{WP} \\
\end{align*}
\]

For present purposes, Y corresponds to a clause-embedding predicate and XP to a DP shell on the embedded clause.

Müller (to appear) also employs this idea in his analysis of restructuring. He argues that restructuring in German shows conflicting evidence for both bi-clausality (CP boundary) and monoclausality (lack of CP boundary). This can be reconciled by having restructuring predicates merge with a finite CP (47a), and then later removing the CP projection (47b) (also see Pesetsky 2019 on Exfoliation).

(47) Restructuring by Remove:

a. \([\text{VP} \quad [\text{−C−}, \text{−T−} \quad [\text{CP} \quad \text{TP} \quad [\text{VP} \quad \text{v} \quad \text{[VP V DP]]}]])\quad (\text{bi-clausality})

b. \([\text{VP} \quad [\text{[HG]}_\mathbb{H}, \text{−T−} \quad [\text{TP} \quad [\text{VP} \quad \text{v} \quad \text{[VP V DP]]}]])\)

c. \([\text{VP} \quad [\text{[HG]}_\mathbb{H}, \text{[HT]}_\mathbb{H} \quad [\text{VP} \quad \text{v} \quad \text{[VP V DP]]}]\quad (\text{mono-clausality})\)

In (47c), the TP projection of the embedded clause has also been removed. This means that Remove can apply cyclically. However, in order to respect the Strict Cycle Condition, higher projections must be removed before lower ones (see section 4.1 for further discussion).

3.2. A restriction on Remove

The restrictiveness of the following analysis will rest, in part, on the assumption that all CPs are actually born as DPs and that this DP shell can be removed by a head bearing an [−D−] feature. In addition, we make the following important two-part assumption about the Removal of DP shells in Kwa
(48) Condition on Removal of DP:
Removal of a DP shell can …
   a. only take place in complement (but not specifier) position, and
   b. may not remove material other than D(P).

The first assumption restricts the application of Removal of DP shells in the
languages in question to complements and not specifiers. While others have
assumed Remove is possible in specifier position (e.g. Müller 2018, Schwarzer
this volume), this could be a point of variation among languages. In the
particular case of clausal determiners, this assumption receives some cross-
linguistic support from the observation that, if a language has a subject/object
asymmetry with regard to when a clause shows the CD, the table in (49) shows
that it is always obligatory in subject position, and either optional or absent in
object position.

(49) Clausal determiners cross-linguistically:

<table>
<thead>
<tr>
<th>Clausal Determiner?</th>
<th>Subject position</th>
<th>Object position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persian</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Greek</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Russian</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Polish</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Uyghur</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>Hebrew</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>Spanish</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>Akan/Gā</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>Ndbele</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

This makes sense in light of (48a), since subject CPs do not occupy the comple-
ment position of a head that could remove their DP shell.\(^8\)

The second part of (48) asserts that Remove is blocked in contexts where
more than just the DP projection would be removed. In other words, if there

\(^8\)The opposite is proposed by Hartman (2012) who suggests that DP shells can be Late Merged
with clauses in order for them to be able to move to subject position (see Pietraszko 2019 for
discussion).
is an element in Spec-DP, Remove will not apply. This will be important in explaining why the CD surfaces in extraction contexts.

With this in mind, we can move onto the various contexts for the clausal determiner. Recall the various contexts for the CD that we saw in (26), repeated below.

(50) **Contexts for clausal determiner in Kwa:**

<table>
<thead>
<tr>
<th>Context</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Complement CP</td>
<td>✗</td>
</tr>
<tr>
<td>Complement CP with extraction</td>
<td>✓</td>
</tr>
<tr>
<td>Sentential subject</td>
<td>✓</td>
</tr>
<tr>
<td>Relative clause</td>
<td>✓</td>
</tr>
<tr>
<td>Initial conditional CP (Gã)</td>
<td>✓</td>
</tr>
<tr>
<td>Final conditional CP (Gã)</td>
<td>✗</td>
</tr>
</tbody>
</table>

The contexts in which the CD surfaces are the *exceptions* rather than the rule, i.e. the contexts in which Remove cannot apply.

3.3. Derivations

3.3.1. *Extraction from complement CPs*

The first context to consider is extraction from complement CPs, as in the Gã example in (51).

(51) a. Kwei yɔṣe [CP əkɛ Yɛmo he ɣeɛ (*le) ]

   Kwei realise that Yɛmo buy yam CD

   'Kwei realised that Yemo bought yam.'

b. Mɛni, ni Kwei yɔṣe [CP əkɛ Yɛmo he ___₁ (*le) ]?

   what foc Kwei realise that Yɛmo buy CD

   'What did Kwei realise that Yemo bought?' (Gã)

---

9 Note that this differs crucially from what is assumed in other Remove analyses (Müller 2018, Schwarzer this volume). Here, the projection in question can still be removed and the stranded former specifiers reattach as specifiers of the next higher head. It may well be that there is cross-linguistic variation with regard to how languages deal with specifiers of removed phrases. A language like German may have access to this kind of ‘reattachment’ repair, whereas other languages, such as Akan and Gã, simply delete the feature without checking it.
In clauses without extraction, the DP shell will be removed by the embedding verb (52), leading to the impossibility of the clausal determiner in (51a).

(52) **Removal of DP shell:**

\[
\begin{array}{c}
\text{VP} \\
V \\
\text{[-D-]} \\
\text{DP} \\
\text{CP} \\
\text{D} \\
(l'ε/nó) \\
\text{C} \\
\text{TP} \\
akε/sε \\
\triangle \\
\end{array}
\Rightarrow
\begin{array}{c}
\text{VP} \\
V \\
\text{CP} \\
\text{D} \\
\text{C} \\
\text{TP} \\
akε/sε \\
\triangle \\
\end{array}
\]

However, if the DP shell hosts material moving successive-cyclically to the edge of DP phase, given the condition in (48b), Remove cannot apply (53).\(^\text{10}\)

(53) **Removal of DP shell blocked:**

\[
\begin{array}{c}
\text{VP} \\
V \\
\text{[-D-]} \\
\text{DP} \\
\text{CP} \\
\text{D'} \\
wh \\
\text{C} \\
\text{TP} \\
akε/sε \\
\triangle \\
t_{wh} \\
\end{array}
\n≠
\begin{array}{c}
\text{VP} \\
V \\
\text{CP} \\
\text{D} \\
\text{C} \\
\text{TP} \\
akε/sε \\
\triangle \\
t_{wh} \\
\end{array}
\]

It is important to mention here that, although we have extraction from a Complex NP Island configuration, this need not be viewed as problematic, since these languages are not island-sensitive (see e.g. Saah 1994, Saah & Goodluck

\(^\text{10}\)Note that would also be an intermediate landing site in Spec-CP if C is a phase. For expository purposes, we will not include this landing site in the trees.
1995 for Akan). Korsah & Murphy (to appear) argue that this follows from the syntax of resumption.

A further important point is that the clausal determiner is also found with monoclausal extraction.\(^{11}\)

(54) a. Taki\(_{i}\) ni  [\(\text{TP} t_i\) tse Momo ] (le) ?
   \(\text{Taki} \quad \text{FOC} \quad \text{call} \quad \text{Momo} \quad \text{CD} \)
   'It is Taki who called Momo.'

b. Hwáñ\(_i\) na Kofi tán nó\(_i\) (nó) ?
   \(\underline{\text{who} \quad \text{FOC} \quad \text{Kofi} \quad \text{hate} \quad \text{3SG.OBJ} \quad \text{CD}} \)
   'Who does Kofi hate?'

This can be accounted for in the same matrix clauses are also DPs that are complements to some left-peripheral head, i.e. Fin (Rizzi 1997). In ordinary declarative clauses, Fin would bear a remove feature \([-D-]\) that removes the DP shell on the matrix CP. Assuming that focused phrases land in a FocP projection above Fin, then successive-cyclic movement would have to stop at the edge of the DP on its way to Spec-FocP. As with embedded clauses, this would block removal of the DP shell (55).

\(^{11}\) Renans (2016) and Grubic & Renans (2017) assume that the clausal determiner in monoclausal contexts is a D head that adjoins to the vP. In section 2.3.2, we saw some reasons to doubt this. Furthermore, their account does not attempt to explain why the CD can only adjoin to vPs from which extraction has taken place.
3.3.2. Sentential subjects

Recall that sentential subjects must always have a clausal determiner:

\[(CP \text{ Aké amlalö le } \text{ tse too no } \text{ *(le) } \text{ ] fee maŋ-bii le} \text{ that government } \text{ def tear tax top } \text{ cd do country-people def} \text{ naakpéé wonder}

‘That the government reduced taxes surprised the people.’ (Gâ)

\[(CP \text{ Së Kofi á-si dán } \text{ *(nó) ] má-a } \text{ Ámá ání gyé-é} \text{ that Kofi perf-build house cd give-past Ama eye collect}

‘That Kofi has built a house made Ama happy.’

This follows from the assumption in (48a) that Remove is restricted to complement position. Thus, even if a remove feature were present on V or v (58), it could not be checked by removing the DP shell of its specifier.

\[(58) \text{ No Removal of D in specifier position:}

\[\ast \]
\[
\begin{array}{c}
\text{DP} \\
\text{CP} \\
\text{le/nó}
\end{array}
\begin{array}{c}
\text{D} \\
\text{v}
\end{array}
\begin{array}{c}
\text{v'} \\
\text{vP} \\
\text{VP} \\
\text{V} \\
\text{XP}
\end{array}
\]

Consequently, derivations containing configurations such as (58) will crash. A convergent derivation with a sentential subject cannot contain a Remove feature on the head that introduces it.

3.3.3. Relative clauses

The third context for the clausal determiner involves relative clauses, where clausal determiners are obligatory on relative clauses in both Akan (59) and Gâ.
To account for this, we can adopt the traditional view that relative clauses are adjuncts to NP (e.g. Partee 1975). On this analysis, shown in (60), the relative is not in the complement position of the noun, for example, and therefore cannot be affected by remove feature.

(60)  

3.3.4. Conditional clauses

The final context we saw was conditional clauses in Gâ. Final conditionals do not show the CD (61a), whereas conditional clauses in initial position must occur with the CD (61b).

(61) Conditional clauses in Gâ (Kropp Dakubu 1992:9):
  a. M-á-yá [CP kë(ji) o-ba (*lɛ) ]
     1SG-FUT-go COND 2SG-COME CD
     ‘I will go if you come.’
  b. [CP Kë(ji) o-ba *(lɛ) ] m-á-yá
     COND 2SG-COME CD 1SG-FUT-GO
     ‘If you come, I will go.’

As with the preceding cases, we argue that this also conforms to a complement/non-complement asymmetry. That is, the CP in (61b) is
immune from Remove of the D shell since it is not in complement position. In particular, we propose that conditionals have a similar syntax to coordination, intuitively supported by their close affinity (62) (Culicover & Jackendoff 1997, Weisser 2015a,b).

(62) You drink another can of beer and I’m leaving (= If you drink another can of beer, I’m leaving).

(Culicover & Jackendoff 1997:197)

Thus, we propose that conditionals are hosted in a functional projection analogous to the &P that is frequently assumed for coordination (e.g. Zoerner 1995). We call this projection Cond(itional)P (63).

(63) Functional projection for conditional clauses:

\[
\text{CondP} \rightarrow \text{Cond} \rightarrow \text{TP} \rightarrow \text{CP}
\]

In this structure, a Remove feature \([-D-]\) can be placed on the Cond head. Consequently, Cond can remove the DP shell on a conditional CP in its complement, i.e. utterance-finally (64).

(64) Removal of D in complement position:
Given this structure, initial conditional clauses are not in complement position (cf. Adger & Quer 2001:125). Given the restriction in (48a), even if Cond bears the relevant Remove feature, it cannot remove the DP shell on the conditional clause.

(65) *No Removal of D in specifier position:*

This accounts for the positional asymmetry with regard to clausal determiners on conditional clauses. This unifies them with both sentential subjects and potentially relative clauses as non-complement configurations.

### 3.3.5. *Interim summary*

In this section, we have seen that the seemingly heterogeneous contexts in which the CD occurs in Akan and Gă can be unified as the contexts in which Structure Removal of a DP shell on a CP fails. This is summarized in (66).

(66)

<table>
<thead>
<tr>
<th>Context</th>
<th>Remove blocked?</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complement CP</td>
<td>✗</td>
<td></td>
</tr>
<tr>
<td>Complement CP with extraction</td>
<td>✓</td>
<td>Moving item in Spec-DP</td>
</tr>
<tr>
<td>Sentential subject</td>
<td>✓</td>
<td>Specifier position</td>
</tr>
<tr>
<td>Relative clause</td>
<td>✓</td>
<td>Adjoined position / Head-raising</td>
</tr>
<tr>
<td>Initial conditional CP (Gă)</td>
<td>✓</td>
<td>Specifier position</td>
</tr>
<tr>
<td>Final conditional CP (Gă)</td>
<td>✗</td>
<td></td>
</tr>
</tbody>
</table>

Other approaches that try to identify a natural class of contexts in which the clausal determiner can be added to a clause inevitably struggle to capture the distribution of the CD, since there is no clear semantic or morphosyntactic
Removing Clausal Determiners in Kwa

property that unifies. The Structure Removal approach advocated here instead characterizes the contexts for the CD as those in which Removal of the DP shell fails.

3.4. Evidence from coordination

The analysis outlined in the preceding section assumed that all clauses are born with a DP shell that may or may not be removed in the course of the derivation. This means that even object CPs that cannot normally show the clausal determiner overtly should were DPs at some point in the derivation.

Some striking evidence for this comes from coordination. In both Akan and Gâ, the form of the coordinator in conjunctions is sensitive to the category of the conjuncts. In Akan, for example, nominals can be coordinated either with ne or na (67a) (also see Larson 2013). Coordination of main clauses, however, cannot have the ne coordinator (67b) (see Renans 2016:98; Korsah 2017:14 for a similar pattern in Gâ).

(67) **Category-sensitive coordination in Akan** (Kobele & Torrence 2004):

a. Me-hu-u [DP Kofi] na/ne [DP Ama]
   1sg-see-pst Kofi and Ama
   ‘I saw Kofi and Ama.’

b. [TP Me-hu-u Kofi] na/ne [TP wo-bɔɔɔ Yaa]
   1sg-see-pst Kofi and 2sg-hit-pst Yaa
   ‘I saw Kofi and you hit Yaa.’

Importantly, when embedded TPs are coordinated, the ne coordinator found with nominals is not possible (68a). We can be sure that TPs are coordinated here since there is only a single complementizer. If the complementizer is repeated in the second conjunct, signaling that larger constituents are coordinated, then the nominal coordinator ne becomes possible (68b) (see Pietraszko 2019 for the similar data from Ndbele). We also find a similar pattern in Gâ.

(68) **Clausal coordination in Akan allows nominal coordinator ne:**

a. Kofi nim [CP se [TP Ama kita bayerɛ] na/ne [TP Yaw
   Kofi know that Ama hold yam and Yaw
   re-noa ɛmo]]
   PROG-cook rice
   ‘Kofi knows that Ama is holding yam and that Yaw is cooking rice.’
This is striking since the form of the coordinator seems to indicate that coordinated embedded clauses have nominal properties, even though the clausal determiner cannot occur in an example like (68b).

We can analyze this as follows. Let us assume that there are two types of coordinators in Akan. The first is the general coordinator in (69a) that selects phrases of any category (●X●) and is realized as na. In addition to this coordinator, there is a more specific variant of & that only selects phrases of category D. Thus, in cases of coordination of object DPs (67a), either of the variants in (69) will be possible, whereas coordination of non-nominals will only be compatible with the coordinator in (69a).

(69) Two lexical entries for & in Akan:

a. \&_1 = PHON: /na/, SEL: [●X●]

b. \&_2 = PHON: /ne/, SEL: [●D●]

Thus, it is prima facie a puzzle why the coordinated clauses in (68b) can occur with the nominal coordinator even though the CD cannot surface overtly. However, on the present view that they are born with a DP shell that is later removed, this is in fact what we expect. At the point at which the clauses are conjoined, they are still DPs. Consequently, they can check the c-selectional requirements of the coordinator in (69b) that will be realized as ne (70).
(70) Step ①: Merge clausal DPs with CD

When the clause-embedding predicate is merged, its remove feature [−D−] will remove the shell of its complement. Note that the &P must be invisible for the purposes of selection and we assume that the same holds for checking Remove features. In (71), Remove applies to the DP shell in each conjunct (i.e. ATB-Removal). The absence of the DP shell on either clause accounts for the why the CD cannot occur in these contexts.

(71) Step ②: Remove DP shells

At this point, we have an opaque output. In other words, it looks like the coordinator ne, which is inherently specified to select only conjuncts of category D, has selected CPs. However, at an earlier point in the derivation, the
conjuncts were DPs. We can therefore view this as a kind of derivational opacity (counterbleeding). Selection applies early, while the clauses are still DPs, and only later are they stripped of this DP shell via Structure Removal. Had the operations been able to apply in the reverse order (Removal ≻ Selection), then the resulting derivation would inevitably crash due to the inability to check the c-selectional features on &. To put it another way, Remove counterbleeds Selection in this analysis (see Benz this volume on another opaque interaction with Remove).

3.5. An allomorphy alternative?

Given that the previous section has provided evidence from coordination for the fact that DP shells are present (at least initially) even when they cannot ultimately be realized, let us briefly consider an alternative approach. Instead of assuming that something like syntactic Structure Removal is responsible for the (non-)realization of a clausal determiner, we could treat this as an instance of movement-conditioned allomorphy of a DP projection that is always syntactically-present.

It is a well-established cross-linguistic fact that many languages show dedicated morphemes to signal the presence of a movement dependency in a given domain (see e.g. McCloskey 2002, Georgi 2014, van Urk 2015). Some of the contexts for the CD could potentially be accommodated under this view. In particular, those involving movement such as extraction from CPs and also relative clauses. Since the moved operator passes through Spec-DP, it would be conceivable that the CD is a form of the universally-present DP shell that is realized under a kind of Spec-Head agreement with a trace or copy (cf. Georgi 2014).

However, some other contexts would remain accounted for under this view. For example, in sentential subjects, it is not clear what allomorphy would be tracking. While some analyses of sentential subjects may posit null operator movement within the CP, there is clear evidence that is not the correct analysis for Akan. For example, Korsah & Murphy (to appear) show that Â-dependencies in the language are signaled by a reflex of high-tone overwriting on the verb. This also holds for some null operator dependencies inside adverbial clauses. Importantly, sentential subjects do not show this reflex, indicating the absence

---

12 This approach has been suggested to us by, among others, David Adger (p.c.).
of a null operator dependency (72a). When an element is extracted from a sentential subject, the high tone reflex on the verb surfaces (72b).

(72) a. \[CP \text{ Sé } \text{ Kofi } d° \text{ Ámá } nó \text{ ] } yε \text{ asɛm } pά \]
    that Kofi love Ama cd be matter good
    ‘That Kofi loves Ama is good news.’

b. \[CP \text{ sé } \text{ Kofi } dό \text{ nó } \text{ nó } \text{ ] } ε-ŋε \text{ asɛm } pά? \]
    who foc that Kofi love 3SG.OBJ cd 3SG-be matter good
    ‘Who is that Ama loves her good news?’

Furthermore, it is unclear how the external distribution of conditional clauses in Gã could be related to the presence of a null operator dependency inside the conditional. For this reason, it seems that this alternative approach cannot adequately capture the full distribution of the CD in the languages in question, unlike the Structure Removal account advocated here.

4. Cross-linguistic perspective

4.1. Restrictions on complement clauses in English

The previous account based on the idea that clausal arguments are born as DPs that may or may not undergo Removal of this DP shell can also be extended to contexts beyond the clausal determiner in Kwa. In particular, it can account for two puzzling, and on the surface seemingly unrelated, puzzles regarding English complement clauses. First, as we saw in section 1, there is evidence that factive complement clauses contain nominal structure (e.g. Rosenbaum 1967, Ross 1967, Kiparsky & Kiparsky 1970, Müller 1995, Kastner 2015). One prominent argument for this is that the nominal structure makes them islands for extraction, i.e. complex NP islands (73).

(73) Factive clauses do not allow extraction (Adams 1985):
    a. *Who, does John hate [DP \[CP \text{ that Mary loves } \_\_\_\_ \]]?
    b. *What, does society care [DP \[CP \text{ that global warming is } \_\_\_\_ \]]?
    c. *How, do you regret [DP \[CP \text{ that Arsenal lost the cup final } \_\_\_\_ \]]?

Furthermore, it has been noted that factive clauses differ from non-factives

(74) No complementizer drop with factives (cf. Bošković & Lasnik 2003):
   a. Flat earthers hate *(that) the world is round (#… but it isn’t round)
   b. Society cares *(that) global warming is a problem (#… but it isn’t a problem)
   c. The players regret *(that) they lost the cup final badly (#… but they didn’t lose)

A similar set of restrictions hold for non-bridge verbs, e.g. manner-of-speaking verbs such as *quip, simper, lisp (Ross 1967, Erteschik-Shir 1973). Complements to manner-of-speaking verbs are similar to factive complements in generally disallowing extraction (75).

(75) Non-bridge verbs do not permit extraction (Ross 1986:154):
   a. *Which hat did Mike quip [CP that she never wore ___]?
   b. *What did she simper [CP that home economics was ___]?
   c. *Who did John lisp [CP that he hated ___]?

Furthermore, there is an observation going back to Ross (1967) that these non-bridge verbs also disallow complementizer drop (76), just like the factive verbs in (74).

(76) No complementizer drop with non-bridge verbs (Ross 1967:252):
   a. Mike quipped [CP that she never wore this hat]
   b. *Mike quipped [CP Ø she never wore this hat]
   c. She simpered [CP that home economics was a bore]
   d. *She simpered [CP Ø home economics was a bore]

These two types of verbs therefore show the same pattern: inability to extract from a clausal complement correlates with inability to drop the complementizer of that clause. It is particularly puzzling why the non-overt realization of a C head as a complementizer, as is often assumed for complementizer drop, should be related to the availability of extraction from that complement.\footnote{13 Of course, that-trace effects may come to mind here, but factive and manner-of-speaking complements also generally disallow extraction from object position (as shown above). Thus, they cannot simply be unified under some explanation of complementizer-trace effects.}
These two seemingly unrelated properties can be given a unified analysis under a Structure Removal approach. As mentioned earlier, we can assume that, in English, extraction is blocked from clausal complements with a DP layer:

(77) *What did she simper [DP [CP that home economics was ___ ]]?

Furthermore, bridge predicates bear the feature [-D-], which removes the DP shell and permits extraction (Kiparsky & Kiparsky 1970). Factive and manner-of-speaking verbs are non-bridge predicates, meaning that they cannot, as a lexical property, bear a [-D-] feature. Since the DP layer cannot be removed, extraction is impossible (77).

Furthermore, let us assume that complementizer drop is derived by Removal of the CP layer via an optional [-C-] on the embedding verb. In the approach to restructuring proposed by Müller (to appear), Removal of the outer layer of the complement of a head bearing a Remove feature can take place recursively. In order to derive complementizer drop, we require two Remove features on the embedding verb. As illustrated in (78) the first removes the DP shell on the CP. Subsequently, the CP layer can be removed, resulting in complementizer drop.

(78)

An important assumption here is that Remove obeys the Strict Cycle Condition, as defined in (79) (Müller 2017, to appear).

(79) **Strict Cycle Condition** (Chomsky 1973, McCawley 1984):

No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also cyclic node.
The result of this is that, assuming each XP is a cyclic node, a CP cannot be removed across an intervening DP shell due to the Strict Cycle (80).

Recall above that we assumed that what factive and manner-of-speaking predicates have in common is that they cannot bear an $[-D-]$ feature. Thus, even if they happen to have the optional $[-C-]$ feature responsible for complementizer drop, this feature cannot be checked in the presence of a DP shell (80).

4.2. Lexical restrictions on sub-extraction from nominals in German

Given what we have seen so far, we might also expect to find $[-D-]$ on non-clausal embedding predicates, removing DP shells on objects. We argue that this could be what we find in the German examples in (81). Sub-extraction of a PP from a DP object is possible if it is the complement of *schreiben* (‘write’) (81a), but not as the complement of a different predicate such as *klauen* (‘steal’) (81b) (similar facts obtain in English; see Bach & Horn 1976).

(81) *Predicate determines subextraction* (Müller & Sternefeld 1995:41):

a. *[PP Über wen ] hat der Fritz [DP ein Buch ___PP ] geschrieben?*
   about whom has the Fritz a book written
   ‘Who did Fritz write a book about?’

   about whom has the Fritz a book stolen
   ‘Who did Fritz steal a book about?’

Previous accounts of this variability involve positing a reanalysis operation based on subcategorization frames (Bach & Horn 1976, Chomsky 1977) or
assuming abstract incorporation that removes the ‘barrierhood’ of the DP (Müller & Sternefeld 1995, Müller 1995). It can also be accounted for under the Remove analysis outlined above.

First, let us adopt the theory of Bošković (2005) where inability to extract from a noun phrase is due to the presence of a DP phase. In particular, Bošković (2005) argues that the impossibility of PP extraction results from a combination of the phasehood of D and an Anti-Locality constraint stating that movement must cross a maximal projection other than the one in which it is immediately contained (Abels 2003, Grohmann 2003).

Given the Phase Impenetrability Condition (Chomsky 2000, 2001), a PP cannot be directly extracted from its base-position inside the DP (82a). Instead, it must move to the edge of the DP first, however this movement violates Anti-Locality since it only crosses the maximal projection which dominated it (NP) (82b).

\begin{align*}
\text{(82) a. } & \quad [ \text{PP} \ldots [\text{DP} [D' D [\text{NP} \text{ N tPP } ]]]] \quad \text{(violates PIC)} \\
\text{b. } & \quad [ \ldots [\text{DP} \text{ PP} [D' D [\text{NP} \text{ N tPP } ]]]] \quad \text{(violates Anti-Locality)}
\end{align*}

It is the combination of these constraints that renders extraction from DP impossible. Equally, the presence of a DP layer in examples such as (81b) would also result in the same movement problem, with direct extraction of PP violating the PIC (83).
Where a predicate such as *schreiben*, which does allow for sub-extraction from its complement, differs from those not allowing extraction is that it can optionally bear a Remove feature $[-D-]$. This feature will lead to the object lacking a DP-layer.

**Removal of DP shell:**

After Removal of DP, the conflict in (82) no longer exists (since there is no longer a phase) and direct extraction of the PP becomes possible (85), as in Slavic languages without a DP (Bošković 2005).
This approach suggests that $[-D-]$ features are by no means restricted to clause-embedding predicates.

4.3. Nominal properties of moved clauses

Recall the puzzling fact from section 1 that the ‘traces’ of moved CPs seem to have the same distribution as nominals. In (86a), we see that CPs cannot ordinarily be the complement of a preposition. However, if this CP is moved, then the sentence becomes possible (86b). In other words, CP gaps seem to behave more like DP gaps.

(86) **CP gaps have nominal distribution** (Postal 1994:70):

a. *I couldn’t convince Frank [PP of [CP that Sonia was very competent ]]  

b. [CP That Sonia was very competent ] I couldn’t convince Frank [PP of ____DP ]

This has lead to an analysis of moved CPs as ‘satellites’ (e.g. Emonds 1972, Koster 1978, Stowell 1981, Alrenga 2005, Moulton 2009, 2013). In other words, the CP itself does not move, but instead a null operator DP with which it is coindexed.

The Remove approach here offers an alternative approach to this puzzle. Let us assume that prepositions bear a Remove feature that can delete the DP shell on a clause in complement position only, in with the assumption in (48a).
Furthermore, we will assume that the structure derived by Removal in (87) leads to a violation of the following PF filter:

(88) *P CP filter:
A CP may not be the complement of a preposition to which it is PF-adjacent.

Evidence for this filter comes from what appear to category mismatches in coordination (Sag et al. 1985:165). In (89a), we have coordinated what appears to be a DP and CP, in violation of the ‘Law of Coordination of Likes’ (Neijt 1979, Sag et al. 1985). As we have seen, a preposition such as on cannot normally take a CP as its complement (89b). Furthermore, Bruening & Al Khalaf (to appear) show that reversing the order of the conjuncts in these examples results in ungrammaticality (89c).

(89) a. You can depend on [DP my assistant] and [CP that he will be on time]
   b. *You can depend on [CP that he will be on time]
   c. *You can depend on [CP that my assistant will be on time] and [DP his intelligence]

This means that, while there can be at least a surface mismatch in category under coordination, this is not possible if the CP is adjacent to the preposition. If we assume, as we did in section 3.4, that clauses are generated with a DP shell and this can be removed in an ATB-fashion in conjunctions, then the structure in (89c) will violate the filter in (88).

Furthermore, we see that the *P CP filter can be circumvented by movement of the clause. Since we assume that it holds at PF, it is possible that moving the clause will result in the filter in (88) not being violated. This is what we see
when clauses that were originally complements to a preposition surface under ellipsis, i.e. in fragment answers (90b) and split questions (90c).

(90) \textit{P+CP possible as fragments} (Merchant 2004:690; Arregi 2010:577):

a. *I am ashamed [\text{PP} of [\text{CP} that I ignored you ]] 

b. A: What are you ashamed of?
   B: [\text{CP} That I ignored you ] I am ashamed [\text{PP} of t_{\text{CP}} ] 

c. What are you ashamed of, [\text{CP} that you ignored me ] you are ashamed [\text{PP} of t_{\text{CP}} ]?

Concretely, we assume that movement of the clause with its DP shell to the edge of the phase can be extrinsically-ordered so as to occur before Removal of DP. Since we have assumed that Removal of the DP shell is limited to complement position, successive-cyclic movement bleeds subsequent Structure Removal of D.

We can see this by returning to the original puzzle of why moved CP complements to prepositions are possible. We assume that PPs are universally phases and that P-stranding languages allow movement to the edge of PP phase (Abels 2012). For examples like (86), if a CP is fronted, it must first move to the phase edge (Spec-PP). Since Remove is feature driven, it can be extrinsically ordered after the edge feature ([\text{●X●}]) triggering intermediate movement to Spec-PP, as shown in (91).

(91) \begin{figure}
\centering
\begin{tikzpicture}[level distance=1.5cm,
                     level 1/.style={sibling distance=3.5cm},
                     level 2/.style={sibling distance=1.5cm},
                     level 3/.style={sibling distance=1.5cm},
                     edge from parent path={(	ikzparentnode.east) -- (0,0) -- (	ikzchildnode.west)}]

  \node {PP}
    child {node {DP}
        child {node {D}}
        child {node {CP} edge from parent node[below] {\textbf{that}}}
    }
    child {node {P'}
        child {node {P} edge from parent node[below] {\textbf{t_{\text{DP}}}}}
        child {node {TP} edge from parent node[below] {\textbf{[\text{●X●}]}}
            child {node {C}}}
    }
\end{tikzpicture}
\end{figure}

Now that the clause is in a specifier position, the DP shell can no longer be removed to check the [\text{−D−}] feature. Consequently, the trace/copy of the clause in its base-position is a DP, since Remove failed to apply. Thus, its nominal distribution can be accounted for.
As a point of cross-linguistic variation, it might be expected that other languages do not have this particular extrinsic ordering of Remove. In this case, Remove of the DP shell could apply before movement and we would expect that the copy/CP-trace would violate (90), if such a filter were also active in the language.

5. Conclusion

This paper has discussed clausal determiners in two Kwa languages: Akan and Gâ. We have seen that the CD across these two languages has a rather complex syntactic distribution, surfacing in extraction contexts, with sentential subjects and certain types of conditional clauses. It was shown that this seemingly heterogeneous class of contexts is difficult to capture as the natural class of contexts in which the a DP can be added to a clause. Instead, we have argued that the contexts in which the CD can be realized constitute cases in which the DP shell inherent to clauses in the languages fails to undergo a process of syntactic Structure Removal. This alternative view of the CD’s distribution unifies these contexts as those in which the clause occupies a non-complement position. We have suggest that, in Akan and Gâ, feature-driven Structure Removal is only able to apply to constituents in complement position.

This analysis necessarily assumes that all clauses are born with a DP shell, what we might call the Universal DP Shell Hypothesis. Further evidence for this was provided from coordination. We saw that the coordinator in Akan can have a different morphological form (ne) when it coordinates nominals (i.e. DPs). Interestingly, this form is possible when full embedded clauses (including a complementizer) are conjoined, but not with simple TP coordination (where the coordinator is shared). This was argued to show that, at the point at which selection takes places, the clauses are actually DPs. They may lose this status by later Removal of the DP shells, however this comes too late to bleed satisfaction of the subcategorization requirements of ne (a case of counterbleeding opacity).

Finally, it was shown that the Structure Removal approach can be extended to some well-known puzzles from other languages. For example, assuming that Remove is subject to the Strict Cycle Condition, it can explain the otherwise puzzling connection between the islandhood and resistance to complementizer-drop found with clausal complements of both factive and manner-of-speaking verbs. Furthermore, it can provide an alternative account for why traces of
moved clauses have a nominal distribution. Also, it was argued that Removal of DP shells can also be extended to extraction from DP objects. An old puzzle is that some verbs easily license sub-extraction from their direct objects, while others do not. It was suggested that this arbitrary property of lexical predicates can be encoded as a Remove feature, removing the DP projection and, with it, the phase boundary. Future research should determine whether the Remove approach can shed further light on the DP/NP debate more generally (see e.g. Bošković 2008).

References


Bruening, Benjamin & Eman Al Khalaf (to appear). Category Mismatches in Coordination Revisited. Linguistic Inquiry


A Remove-Based Theory of the Complementarity Effect in Breton

Philipp Weisser*

Abstract
In this paper, I discuss the so-called Complementarity Effect of ϕ-agreement in the Celtic language Breton: Verbal ϕ-agreement is only ever overt if the agreement controller (i.e. the subject) is covert. I argue that this peculiar phenomenon provides an argument that the covertness of the subject should be analysed as being due to syntactic removal rather than due to the presence of a phonologically empty element in the subject position. Theories of pro-drop in terms of an empty pro have no handle to derive why verbs only agree with elements that are phonologically null. In an alternative theory of pro-drop which makes use of the syntactic operation REMOVE, it can be modelled that removal of a pronoun from the subject position leaves the ϕ-features of the pronoun unattached and in accordance with the reintegration property of the operation, these features will be reattached on a functional head that eventually will be realized as a part of the verb. Under these assumptions, the presence of ϕ-features on the verb is an immediate consequence of removal of the pronoun and therefore straightforwardly models the complementary distribution of verbal agreement and its controller.

1. Introduction: Pro-Drop and Remove

Theories of syntax differ crucially as to whether they assume the existence of Dark Matter, i.e. syntactic elements that seem to have some syntactic effect but themselves cannot be observed directly. One of the poster children for the existence of Dark Matter in the syntax is the phenomenon of pro-drop: Some languages entertain the possibility to leave a subset of arguments of a clause phonologically unexpressed. But even though these arguments are unexpressed, they still seem to induce some syntactic effects. Given the right syntactic configurations, they trigger agreement on the verb or other elements of the clause and potentially license anaphors, etc.

* For helpful discussion of the Breton facts and the theoretical underpinnings of the concept that is REMOVE, I would like to thank Johannes Hein, Anke Himmelreich, Gereon Müller and Andrew Murphy. All remaining errors are my own.

Structure Removal, 295–322
Andrew Murphy (ed.)
Linguistische Arbeits Berichte 94, Universität Leipzig 2019
According to the theories which allow for the existence of Dark Matter, these effects are usually derived by positing a phonologically empty element in the syntactic position where usually the respective arguments are located. The element, typically referred to as pro, has the appropriate morphosyntactic features to trigger agreement and license anaphors, etc but does not have phonological features.

Being one of the prototypical instances of Dark Matter in the syntax, the phenomenon of pro-drop, in principle, lends itself to a derivation that fits into the general spirit of this volume. The operation that is Structure Removal is specifically designed to derive Dark Matter effects where, representationally speaking, seemingly invisible elements seem to trigger syntactic processes.

Despite this obvious overlap between the theoretical operation that is REMOVE and the empirical phenomenon of pro-drop, no attempts have, so far, been undertaken to propose a coherent theory of pro-drop making use of the operation REMOVE. Presumably, the reason for this lies in the very nature of the phenomenon itself. As noted above, the standard theory of pro-drop, at least within the current Minimalist framework, is the assumption of a phonologically empty element in argument position. And needless to say, distinguishing between an empty element in argument position and a removed element in argument position is not an easy task.

The main contribution of this short paper is to provide one small argument that a REMOVE-based theory of pro-drop may help us explain some facts about pro-drop that otherwise are highly problematic. The empirical domain of the paper will be what is known as the Complementarity Effect of $\phi$-agreement found in many Celtic languages. $\phi$-agreement on the verb or on prepositions in these languages depends on the covertness of its trigger. Only dropped arguments can trigger agreement. Given the current minimalist assumptions about the architecture of grammar, it seems highly implausible that the phonological shape of a pronoun (overt vs covert) can be held responsible for a difference in syntactic behavior. This suggests that difference between an overt and a covert pronoun is to be located in the syntax. I will propose an alternative account according to which it is the syntactic operation REMOVE that conditions $\phi$-agreement in the syntax.

The paper is structured as follows: In Section 2, I will introduce the complementarity effect. I will focus my discussion on the language Breton as the effect has been investigated in this language in great detail. In Section 3, I will outline the dilemma for theories of the complementarity effect already anticipated by
Stump (1984). I will also outline the current approach proposed by Jouitteau & Rezac (2006) and discuss some of its shortcomings. In Section 4, I will then outline an alternative in the spirit of this volume making use of the operation REMOVE. Section 5 discusses some additional facts mentioned by Jouitteau & Rezac (2006) and sketches a solution to these issues. Section 6 concludes.

2. The Complementarity Effect

Agreement in Celtic languages has been the study of some time since it displays various intricate restrictions concerning (a) the syntactic or linear position of the agreement controller with respect to the agreement target and/or (b) whether the target is actually phonologically overt or not. In this paper, I will discuss agreement in the Celtic language Breton where only restrictions of the latter type are at play.

As other Celtic languages, Breton exhibits ϕ-agreement on verbs with the subject of the respective clause. Notably, this kind of agreement is subject to what Stump (1984) calls the Complementarity Principle given in (1):

(1) Complementarity Principle (Stump 1984:292):
Within a clause, overt argument noun phrases never appear with concurring personal affixes.

The crucial word of this principle is overt. In other words, we see ϕ-agreement only if the trigger of agreement is phonologically null. Consider the examples in (2). Here, the subject is unexpressed and therefore the verb displays a form alternation indexing the ϕ-features of the unexpressed subject. In (2a), the form lennan unambiguously indexes a first person singular subject whereas the form lenmont in (2b) indexes a third person plural subject. The present tense paradigm of verbal agreement of the verb lenn- is given in (3). It shows that there is a one-to-one mapping between function and form. Each set of ϕ-features receives a unique morphological form.1

---

1 I adapted the glosses as well as part of the morphological segmentation from the relevant sources to match for reasons of readability. The glosses follow the standard of the Leipzig Glossing Rules. The gloss PTCL refers to a special preverbal particle, referred to as rannig in Breton (which is why Rezac (2004) and Jouitteau & Rezac (2006) gloss it is R). In accordance with its morphophonological behavior I have used the ‘=’-symbol to indicate that it is cliticized to the verb.
(2)  
   a. Levriou a=lennan  
       books PTCL=read.1SG  
       ‘I read books.’  
   b. Levriou a=lennont  
       books PTCL=read.3PL  
       ‘They read books.’  
      (Stump 1984)

(3)  
   Present tense forms of lenn- (‘read’) (Stump 1984):

   1.SG. lennan  
   2.SG. lennez  
   3.SG. lenn  
   1.PL. lennomp  
   2.PL. lennit  
   3.PL. lennont

   Compare now the corresponding counterparts with an overt pronoun in (4).  
The sets of \( \phi \)-features involved are identical to the examples in (2) but there is  
no agreement on the verb. The verb appears in the form of the third singular  
present tense.

(4)  
   a. Me a=lenn levriou  
       1SG PTCL=read books  
       ‘I read books’  
   b. Int a=lenn levriou  
       3PL PTCL=read books  
       ‘They read books’  
      (Stump 1984:291)

   The pattern extends of course to full NPs which do not trigger agreement on  
the verb either. In (5), the subject is plural but still the verb shows up in the  
form of the third person singular.

(5)  
   a. Levriou a=lenn ar vugale  
       books PTCL=read the children  
       ‘The children read books.’  
   b. *Levriou a=lennont ar vugale  
       books PTCL=read.3PL the children  
       ‘The children read books.’  
      (Stump 1984:292)
The examples in (6) show that at least in Breton, the complementarity effect is independent from the word order in the respective clauses. Regardless of whether the subject is pre- or postverbal, it will never trigger agreement if it is overt.

(6) a. Ar vugale a=lenn levriou the children PTCL=read books ‘The children read books.’
   b. *Ar vugale a=lennont levriou the children PTCL=read.3PL books ‘The children read books.’ (Stump 1984:292)

As for postverbal pronouns, we find that overt pronouns are not available independent of whether we have actual φ-agreement or not.²

(7) a. *Levriou a=lenn me books PTCL=read I ‘I read books.’
   b. *Levriou a=lennan me books PTCL=read I ‘I read books.’ (Stump 1984:302)

Verbal agreement is unaffected by the thematic role of the subject. The Complementarity Principle holds for transitive subjects (as in all previous examples) as well as for subjects of passives or unaccusatives:

(8) a. Dec’h e=veze gwelet ar merc’hed. yesterday PTCL=was.3SG seen the girls ‘The girls were seen yesterday.’

²Stump (1984:302) mentions that emphatic pronouns can occur postverbally and interestingly, they do seem to trigger agreement in that configuration.

(i) Levriou a=lennan=me /*lenn=me books PTCL=read.1SG=EMPH.1SG ‘I read books.’

But he argues that, in that cases, the emphatic pronouns are not the actual arguments of the verb but rather elements in an A’-position which are merely associated with an empty pronoun. I will follow this argumentation assuming that these cases are no real exceptions to the Complementarity Principle.
b. Alies e=kouezhe an deliòu
often PTCL=fall.3SG the leaves
‘The leaves fell often.’  
(Jouitteau & Rezac 2006)

The Complementarity Principle in (1) does not make reference to the agreement target being a verb. And indeed we find that the pattern extends to inflected prepositions in Breton. Some prepositions in Breton show agreement with their complement if the complement is covert. The examples in (9) show that the preposition gan(t)- ‘with’ covaries with the $\phi$-features of its complement. Its full paradigm is given in (10). Again, we see that every set of $\phi$-features receives a unique form. Note also that unlike verbs, inflected PPs also distinguish gender in the third person singular.

(9) a. Ul levr brezhonek a zo ganin
a book Breton PTCL is with.1SG
‘I have a Breton book.’
b. Ul levr brezhonek a zo ganto
a book Breton PTCL is with.3PL
‘They have a Breton book.’

(10) \textit{Forms of gan(t)- (‘with’)} (Stump 1984):

\begin{tabular}{ll}
1.SG & ganin \\
2.SG & ganit \\
3.SG.MASC & gantan \\
3.SG.FEM & ganti \\
1.PL & ganeomp \\
2.PL & ganeoc’h \\
3.PL & ganto \\
\end{tabular}

But, as with verbs, the inflection vanishes as soon as the agreement controller is overt as in (11), where it is a full NP. Interestingly, with prepositions, it is not the third person singular form that appears but the bare stem of the preposition.

(11) a. Ul levr brezhonek a zo gant Yannig
a book Breton PTCL is with Yannig
‘Yannig has a Breton book.’
b. *Ul levr brezhonek a zo gantan Yannig
a book Breton PTCL is with.3SG Yannig
‘Yannig has a Breton book.’
So, it seems that the Complementarity Principle in (1) is a pretty robust generalization about the syntax of Breton. There are some residual cases which cast doubt on the validity of the principle such as preverbal subjects in negative clauses but Stump (1984) argues convincingly that these are in fact no real counterexamples as these are presumably left dislocated topics (as they are emphatic). Thus they do not invalidate the principle in (1) as they are, by assumption outside the clause.

3. Analyses of the Complementarity Effect

In this section, I will address some previous accounts to deriving the Complementarity Effect. I will briefly address the two camps already anticipated by Stump (1984), namely the incorporation approach as well as the agreement approach and list some of their respective shortcomings. I will then spend a little more time discussing a recent approach by Jouitteau & Rezac (2006) who propose a hybrid approach trying to combine the properties of both types.

3.1. The Incorporation approach

The first approach is the incorporation approach. Discussions of this approach are found in Anderson (1982) and Stump (1984) of which only the former actually argues for it. According to the incorporation approach, there is no verbal subject agreement in Breton. What looks like exponents of agreement features on the verb is actually an exponent of the pronoun itself. The pronoun incorporates into the verb and leads to distinct verbal affixes depending on its feature specification. The verb is in (or has moved to) a clause-initial position and a postverbal pronoun subsequently incorporates into it.

\[ (12) \quad [IP \ V [VP \ Pron_{Subj} \ t_V \ Obj]] \]

The straightforward advantage of this account is that it nicely derives the complementary distribution of the subject and the subject agreement on the verb. If they are one and the same thing, the Complementarity Effect falls out as completely natural. However, there are a number of quite severe problems with such an account. First, it is far from clear what this movement step is and how it relates to general theories of movement. If we take the account literally as incorporation, then we find quite clearly that it does not exhibit
any of the hallmark properties of incorporation as identified by Baker (1988) and many others. It is unclear in the first place whether incorporation of transitive subjects is attested to begin with. Only a handful of languages have been claimed to exhibit subject incorporation and in many cases, only very few subject-verb combinations are allowed. This kind of incorporation here, however, is independent of the verb and completely productive. Further, incorporation of pronouns is impossible even in object incorporation languages. Incorporation is typically restricted to non-specific indefinites.

So, we conclude that this kind of incorporation cannot be of the Bakerian type. An alternative would be to view it as some kind of clitic movement as evidenced by some Romance languages. In these languages, typically objects (but sometimes also subjects) cliticize to the main verb of the clause and form a syntactic and phonological constituent with it. However, a closer look at the morphological footprint of the operation at hand reveals that it does not at all look like cliticization of the Romance type. It looks much more like agreement. Even in languages with what Zwicky (1977) calls special clitics, we typically see some morphophonological similarities between the clitic and its non-clitic counterpart. But in the case at hand, there are no similarities of that sort. Compare the forms in the following table:

(13)  *Comparison between pronouns and verbal agreement exponents:*

<table>
<thead>
<tr>
<th></th>
<th>Verbal Agr.</th>
<th>Pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG.</td>
<td>-an</td>
<td>me</td>
</tr>
<tr>
<td>2.SG.</td>
<td>-ez</td>
<td>te</td>
</tr>
<tr>
<td>3.SG.Masc/Fem</td>
<td>-∅</td>
<td>eñ/hi</td>
</tr>
<tr>
<td>1.PL</td>
<td>-omp</td>
<td>ni</td>
</tr>
<tr>
<td>2.PL</td>
<td>-it</td>
<td>c’hwi</td>
</tr>
<tr>
<td>3.PL</td>
<td>-ont</td>
<td>int</td>
</tr>
</tbody>
</table>

(Stump 1984)

The same point of criticism applies to incorporation approaches where the incorporation is more of a superficial morphophonological operation applying in a postsyntactic, prosody-sensitive component (Adger 2000, Ackema & Neeleman 2003). If the process of incorporation were simply prosodic amalgamation, then we would expect there to be phonological overlap between the paradigms.

I thus take the incorporation approach to be insufficient to derive the
complementarity effect in Breton. It has a straightforward way of deriving the effect itself but suffers from several severe problems.

3.2. The Agreement approach

The alternative to an incorporation approach is to say that the form alternations on the verb in Breton are due to agreement with some sort of empty element pro in subject position. The crucial problem for such an approach is the very nature of the Complementarity Effect. Why is it that only the empty element can agree but overt pronouns or full NPs cannot? It seems that in one way or another, it has to be stated that the agreement process itself makes reference to the phonological overtness of the pronoun. This is explicitly stated in the account proposed in Stump (1984) who adopts the language-specific constraint for Breton in (14):

(14) The argument position encoded by AGR must be occupied by a null element.

(Stump 1984:316)

Another version is given in Jouitteau & Rezac (2006):

(15) ϕ-PHON constraint (Jouitteau & Rezac 2006):
ϕ-Agree is limited to phonologically empty goals.

Apart from the fact that (14) and (15) are simply stipulations which restate the observed facts in more theoretical terms, it is clear that a formulation along those lines is highly problematic for the current assumptions about the architecture of grammar. Largely independent of framework-internal specifics, most frameworks adopt the famous Principle of Phonology-free Syntax (Zwicky & Pullum 1986, Miller et al. 1997):

In the grammar of a natural language, rules of syntax make no reference to phonology.

If the ability to control agreement were simply a matter of being phonologically null, then this would be a strong violation of (16). Thus, since most theories

---

3See also the similar treatment of McCloskey & Hale (1984) for Irish.
adhere to the principle in (16), I take it that the difference between the overt and
the covert pronoun must be an issue of syntax. In Section 4, I will argue that the
overt and the covert pronoun structures are derived from the same underlying
source but the two derivations differ as to whether the pronoun undergoes
syntactic removal. Before I will come to present an alternative approach, I will
briefly discuss the theory by Jouitteau & Rezac (2006) which tries to combine
the agreement and the incorporation analysis.

3.3. Hybrid Account

In the previous subsections, we have seen that both the incorporation approach
and the agreement approach suffer from severe problems. Jouitteau & Rezac
(2006) propose a hybrid account which makes use of both, incorporation and
agreement claiming to overcome the problems of the individual approaches.
Their approach features both an agreement probe on the verb (or more accurately
on T) and an incorporation process by which a pronoun adjoins to T. I will go
through the relevant structures one by one. They assume that the position of
the verb is in T, which also bears a $\phi$-probe in order to collect the features of
the subject of the clause. The subject itself has evacuated the vP landing in an
intermediate specifier of a projection they call FP.\(^4\)

(17)

\[
\begin{array}{c}
\text{TP} \\
T_{\phi}\downarrow \\
\text{FP} \\
\text{NP}_{\text{Subj}} \\
F' \\
F \\
vP \\
\text{v'} \\
v\text{VP} \\
v'_{\text{Subj}} \\
\end{array}
\]

\(^4\)Jouitteau & Rezac (2006) discuss whether FP corresponds to the projection AgrP used in
McCloskey (1996) for postverbal subjects in Irish but ultimately decide to remain neutral about
this question. Crucially, they need the additional assumption of a derived subject position
for various structures involving raising or copula predicates. Also, Rezac (2004) shows that
vP-fronting in Breton leaves the subject behind. He therefore concludes that the subject evacuated
the vP before the vP has undergone remnant movement.
Building on the work by Jouitteau (2005), they argue extensively that the vP has nominal properties in Breton which they model by assuming the presence of a feature [+D] on vP and FP. This renders the FP, so they claim, opaque to $\phi$-agreement from T. T probes for the features of the subject but only finds the FP and due to Relativized Minimality, it cannot probe deeper. Thus, it is stuck with the default features on FP itself.

The underlying idea according to Jouitteau & Rezac (2006) is to have the Complementarity Effect reduce to locality effect induced by Relativized Minimality. Agreement with subjects is blocked in Breton because they are themselves embedded into a projection with nominal properties. In order to derive the actual agreement with covert pronouns, they additionally assume an incorporation rule of the type discussed in Section 3.1. $pro$ can move out of FP and adjoin to T, thereby handing its features over to T.

---

$^5$From the discussion they offer, it does not become entirely clear why FP bears the [+D]-feature as well. While their arguments that the vP has nominal properties are largely convincing, they do not offer any insights whether the same holds for FP and if so why.
Unfortunately, Jouitteau & Rezac (2006) are not particularly explicit about the properties of this movement step but from the discussion, it becomes clear that they imagine this step to be somewhat parallel to cliticization movement in Romance. But clearly, many questions about this movement step remain unanswered. They admit that they do not really know what the trigger for this movement step is as they only note that it cannot be for reasons of case assignment or $\phi$-feature checking. Further, it remains unclear why agreement is blocked by the [+D]-feature on the maximal projection of FP but movement of the subject pronoun is not. Finally, it must be said that the approach suffers from the same drawbacks as the other incorporation approaches discussed in Section 3.1. If this were actual cliticization, we would expect there to be some morphophonological overlap between the respective paradigms. As noted above, the verbal paradigm looks very much like a run-of-the-mill agreement paradigm and not like the result of cliticization.

To conclude, the hybrid approach tries to incorporate both approaches to derive the Complementarity Effect. And, in a sense, trying to reduce the absence of verbal agreement with regular subjects to a locality effect due to the nominal properties of v is - given the evidence for it - a nice touch. But ultimately, the actual agreement is nonetheless derived by means of a very obscure incorporation operation that neither fits the established footprints for incorporation nor cliticization. In the next section, I will present an approach that straightforwardly derives the agreement on the verb as the immediate consequence of the absence of a subject. The underlying idea is that the removal of the subject itself from the syntactic derivation is responsible for the transfer of the agreement features onto T. This is the only possibility to get the $\phi$-features of the subject onto T. Hence, removing the subject is the only possibility to ever get verbal agreement.

4. A REMOVE-based alternative

4.1. The Reintegration Property of REMOVE

The operation Structure Removal (Müller 2017, Müller & Murphy 2019, Müller 2018) is specifically designed to derive Dark Matter effects in the syntax where a seemingly invisible element X has some syntactic effects but cannot be observed directly. The logic behind the operation is that syntactic structure is built up as usual and X is merged into the syntactic structure as every other
A Remove-Based Theory of the Complementarity Effect in Breton

At some point X undergoes structure removal, an operation that is triggered by a nearby functional head, which removes X from the syntactic derivation altogether. The result of structure removal is that not only is X ultimately invisible on the surface but it is also invisible to subsequent syntactic processes. In this respect, syntactic structure removal is crucially different from other deletion operations such as impoverishment or obliteration (see e.g. Nevins 2012). Both, impoverishment and obliteration are used to delete morphosyntactic objects towards the end of the derivation so as to not have them appear in the surface string.

Crucially, the operation REMOVE is an extremely local operation as it is subject to the Strict Cycle Condition and interacts with syntactic structure building via Merge. A head Y can only trigger REMOVE of its complement or its specifier or the respective heads of the complement or the specifier. In the case at hand, we are dealing with subjects, thus we will be concerned only with REMOVE targeting specifiers.

A further property of REMOVE is what I will call the reintegration property. If a REMOVE-feature [-Y-] on a head X targets the head of its complement Y which itself is merged with an element ZP, then REMOVE will delete Y including all of its projections and reintegrate ZP as the new complement of X. This is illustrated in (20) and (21):

(20) Before Removal of Y:

Before Removal of Y:

\[
\begin{array}{c}
\text{XP} \\
\text{X[-Y-]} \\
\text{YP} \\
\text{Y} \\
\text{ZP} \\
\end{array}
\]

After Removal of Y:

After Removal of Y:

\[
\begin{array}{c}
\text{XP} \\
\text{X} \\
\text{ZP} \\
\end{array}
\]

If Y has multiple arguments (or possibly even adjuncts), reintegration will apply such that the respective c-command relations and the respective phrase structural status of the elements involved will be maintained. An element formerly known as a specifier of Y will end up as the specifier of X:
Based on this process of reintegration which maintains the phrase structural status of syntactic elements, I would like to suggest that the same process can under certain circumstances also apply to morphosyntactic elements adjoined to the deleted element itself. Consider the situation in (24). The head Y which is to be removed has an adjunct H (presumably created by means of head-movement of H to Y). The question is, of course, what happens to the element H when Y is removed. In accordance with the reintegration property sketched above, I would like to argue that H is reintegrated maintaining its phrase structural status as an adjunct to a head as in (25):

(24) **Before Removal of Y:**

```
XP
X[-y-] YP
WP Y
WP Y' ZP
```

(25) **After Removal of Y:**

```
WP X' ZP
XP X
```

The result of this kind of REMOVE-operation is that X obtains features of Y without actual agreement but only if X removes Y from the derivation. The connection to the Complementarity Effect in Breton is straightforward: Only if the relevant functional head in the verbal spine removes its specifier from the derivation does it obtain its ϕ-features. This essentially means that only removed items will be able to control form alternations on the verb.

4.2. The Derivations

In this section, I will apply the general idea developed in the previous subsection to the concrete derivations in order to derive the basic facts about Breton verbal
agreement. As for the general clause structure of Breton, I follow Jouitteau & Rezac (2006) in assuming a derived subject position outside of the vP. Whether this is SpecAgrP as assumed by McCloskey (1996) for Irish, SpecFP as in Jouitteau & Rezac (2006) or SpecTP as in Jouitteau (2007) is not really relevant for the purposes of this paper.

For the sake of concreteness and comparability, I will adopt the structure used above in Section 3.3, but it should be clear that the structures are all compatible as long as they provide for a derived position outside of the vP that hosts all sorts of subjects (recall from Section 2 that the Complementarity Effect holds for transitive, unergative and unaccusative subjects alike). I will also follow the structures above in labelling the projection hosting the derived subjects FP but as Jouitteau & Rezac (2006), I remain agnostic as to whether this projection can be identified with other positions.

As for the position of the verb, I assume that it moves to a high position in the clause, which hosts the preverbal particle e or a in all the examples above.6 Again, there is some debate as to whether this position is C⁰ (as proposed by Schafer 1995), T⁰ (as proposed by Rezac (2004)) or Fin⁰ (as proposed by Jouitteau (2007)). But crucially, as Breton is a V2-language in matrix clauses, whatever the exact final landing site of the verb is, we must ensure that there is exactly one specifier in the left periphery. For the sake of concreteness, I thus adopt the following clause structure for an example with a postverbal NP-subject and some non-subject XP in the preverbal position:

---

6The particle, typically referred to as the rannig-verb, i.e. small part of the verb alternates depending on the syntactic category of its specifier (see Anderson (1981), Urien (1999), Rezac (2004), Jouitteau (2007), Weisser (2019)) but unambiguously belongs to the verbal spine as it undergoes morphophonological processes triggered by other verbal elements. It e.g. competes for the position with other verbal clitics such as the object clitics and negation.
In order to derive the Complementarity Effect, a few preliminary assumptions need to be made. I assume that pronouns are at least minimally complex in the sense that they consist of a D-like element to which the respective $\phi$-features are adjoined. Whether this structure is base-generated as such or whether it is the result of head movement, does not play a role for now. Further, I assume that even though the D-element shares a number of properties with determiners or determiner phrases (at least in terms of syntactic distribution), it is still featurally sufficiently distinct. I will code this by using a little index $pro$. Thus, pronouns in this theory are represented as in (27). Note that overt and covert pronouns are featurally completely identical. The difference will arise as to whether they are affected by REMOVE or not.

With (27), we have all the background assumptions in place to model the Complementarity Effect. In order to do that, I assume that the functional head $F^o$ has a feature that allows it to remove pronominal arguments from its specifier: $[-D_{pro}]$. If the subject is a pronoun and has moved to SpecFP, then REMOVE will apply, removing the pronoun from F’s specifier. This, by assumption, will leave the adjoined $\phi$-head in (27) unattached (since only D is removed) and as discussed in the previous subsection, the reintegration property of REMOVE will reattach the $\phi$-head while maintaining its phrase structural status as an adjunct to a head. As F is the only available head in the
current cycle, it will attach the $\phi$-head to F. This is illustrated in (28) and (29). In the structure in (28), the pronoun has moved to SpecFP. The head of the current cycle bears a feature that allows it to remove D. This process leaves the adjoined head $\phi$ unattached and consequently, $\phi$ is then attached to F. In the resulting structure in (29), F has obtained the features of its specifier by removing it from the derivation.

(28) **Before Removal:**

(29) **After Removal:**

In the subsequent cycles, F will undergo head movement all the way to Fin carrying the adjoined $\phi$-head along. It is only in this configuration that the complex head including the verb, can obtain the subject features. Since none of the functional heads in Breton have a proper agreement probe, REMOVE of a subject pronoun is the only option for the verb to show form alternations depending on the subject features.

We thus see how, in the case of a pronominal subject, the $\phi$-features end up on the verbal complex. Importantly, since SpecFP is, by assumption, a position dedicated to subjects, the verb can only ever show subject agreement. Also, it is important that the feature on F is relativized to affect only pronominal subjects. Full NP subjects cannot be removed by F. I would like to attribute this fact to the recoverability property of REMOVE. It seems plausible to assume that full NPs cannot be recovered semantically and therefore cannot be removed. In that case, I assume that the respective feature on F is simply deleted without effect.

A final note concerns preverbal pronouns. Postverbal pronouns are obligatorily affected by REMOVE and therefore deleted from the derivation. In order for postverbal pronouns to arrive safely in preverbal position (i.e. SpecFinP), we would need to assume that they have a way of avoiding removal. One possibility would be to assume that elements which move to the preverbal position bear at least information structural features such as [topic] or [focus] as the preverbal position in Breton is, as in other V2-languages, associated with features of this type (see Schafer 1995, Rezac 2004). As these features are presumably located on
D itself, `DELETE` would also delete non-recoverable material. As with full NPs above, I would like to submit that pronouns bearing additional information structural features cannot be deleted. Again, the respective [-D-]-feature on F is deleted without effect.\(^7\)

### 5. Extending the analysis

In this section, I discuss some additional data that need further consideration. First, I discuss the cases of inflected prepositions which also seem to obey the Complementarity Principle (Section 5.1). I will then devote a short subsection to the question of different types of defaults which Jouitteau & Rezac (2006) take as strong evidence for their approach. I will argue that the conclusions they draw are not forced and may receive an alternative explanation. Finally, I will discuss the only verb of the Breton language *eus* (`have`) where the Complementarity Principle breaks down (Section 5.3). *eus* is inflected regardless of whether its subject is covert or overt.

#### 5.1. Inflected Prepositions

In this short subsection, I will briefly discuss the pattern of inflected prepositions. Building on data from Jouitteau & Rezac (2006), it is shown that for those, an analysis in terms of morphophonological amalgamation is much more plausible. Ultimately, this shows in my opinion that deletion operations can apply at different stages of the derivation.

One of the arguments against a morphophonological amalgamation approach which derive the complementarity effect by means of cliticization of the subject pronoun to the verb was that we find absolutely no overlap between the morphophonological form of the pronouns and the verbal inflection throughout the whole paradigm in (13). With inflected prepositions however, the situation is at least somewhat different. Compare the forms of the respective rows in (30).

---

\(^7\) Another possibility how to make sure that preverbal pronouns avoid being removed would be to assume that they manage to get there without making an intermediate stop-over in SpecFP. At this point, I do not see, however, how this can be avoided without look-ahead-like mechanisms.
Pronouns and agreement exponents of gan- ('with') (Stump 1984):

<table>
<thead>
<tr>
<th>Prepositional Agr.</th>
<th>Pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG.</td>
<td>-in</td>
</tr>
<tr>
<td>2.SG.</td>
<td>-it</td>
</tr>
<tr>
<td>3.SG.Masc/Fem</td>
<td>-tān/-ti</td>
</tr>
<tr>
<td>1.PL.</td>
<td>-omp</td>
</tr>
<tr>
<td>2.PL</td>
<td>-c’h</td>
</tr>
<tr>
<td>3.PL</td>
<td>-(n)to</td>
</tr>
</tbody>
</table>

With the exception of the first person pronouns, there is a substantial overlap in morphophonological terms (and even those all include a nasal). In the second person, both forms include a /t/ (in the singular) and a /c’h/ (in the plural). In the third persons, the singular has a straightforward overlap and, in the plural as well if we assume that the nasal consonant of the pronoun merges with the stem-final nasal of the preposition.

Morphophonological similarities are suggestive but do not constitute a knock-down argument against a REMOVE-based approach to inflected prepositions in Breton. However, Jouitteau & Rezac (2006) give a strong argument that there is fundamental difference between the treatment of inflected verbs in Breton and inflected prepositions. The argument involves conjunction with pronouns as the first conjunct of (a) the subject of verbs and (b) the complement of prepositions. Consider the conjunction of subjects in (31). In (31a), we see a conjoined subject with a full NP as the first conjunct. In line with the Complementarity Principle, we do not see agreement. The verb surfaces in the default. In (31b), we see that it is not possible to have an empty pro as the first conjunct of the complex subject.

(31)  a. Dec’h e=erruas [Nolwenn hag ar gorrien]
      yesterday PtCl=arrived.3SG Nolwenn and the dwarves
      ‘Yesterday arrived Nolwenn and the dwarves.’

  b. *Dec’h e=erruas [pro hag ar gorrien]
      yesterday PtCl=arrived.3SG 3SG and the dwarves
      ‘Yesterday arrived she and the dwarves.’

        (Jouitteau & Rezac 2006)

Compare the pair in (31) with the pattern of inflected prepositions in (32). Here, it is possible to have the first conjunct incorporate into the preposition.
Philipp Weisser

(32)  Chom a=rae [etrezi [pro hag ar gorrien]]
      stay PTCL=did.3SG between-3SG.FEM 3SG and the dwarves
      ‘S/He stayed between her and the dwarves.’

(Jouitteau & Rezac 2006)

The minimal pair in (31) follows straightforwardly under a REMOVE-based account. As mentioned above, REMOVE is designed as a strictly local operation. A functional head X can only remove (the head of) its specifier or its complement but nothing embedded into the specifier or the complement. Thus, it follows that, under standard asymmetric &P-accounts to coordination structures, it follows that both conjuncts are too deeply embedded into the &P to be affected by REMOVE:

(33)  **Removal Impossible:**

```
      FP
     / \  
    &P  F'
   / \  
 D_{pro} &' F[-D_{pro}-] vP
   / \  
 & DP t&P ...
```

But since REMOVE is a strictly local operation, we have no immediate way of accounting for the pattern of inflected PPs in (32). If the Complementarity Effect with prepositions were also due to syntactic REMOVE, then (32) would be a huge problem.

I thus follow the proposal by Jouitteau & Rezac (2006) and argue that inflected prepositions in Breton do not involve *Structure Removal* but are to be derived by means of postsyntactic amalgamation along the lines of Ackema & Neeleman (2003, 2004). The inflected PP in (34a) thus derives by means of morphophonological cliticization from a source as in (34b):

(34)  a.  [etrezi [∅ hag ar gorrien]]
      between.3SG.FEM and the dwarves
  b.   etrez [hi hag ar gorrien]

---

8 It also follows under the assumptions by Jouitteau & Rezac (2006) if the movement step that is cliticization into the verb obeys the Coordinate Structure Constraint (cf. Ross 1967).

Before concluding, I would like to remark that the pattern of inflected prepositions shows that there is a fine line between cyclic removal of structure and more surfacy morphological rules that are better analyzed in terms of context-sensitive spell-out. In particular, I would like to suggest that empirical problems such as the Complementarity Effect in Breton, which comes in various guises shows that some deletions operations apply in the syntax (i.e. REMOVE) but some may apply in the morphosyntax (i.e. OBLITERATION or IMPOVERISHMENT) or even later in the (morpho)phonology. I think that such patterns substantiate the claims made in Nevins (2012) that deletion is a general operation which is available at various points of the derivation including syntax proper. I would like to extend his typology of deletion processes to include a genuinely syntactic operation: REMOVE.

5.2. Two types of defaults

In this subsection, I very briefly address an issue raised by Jouitteau & Rezac (2006). They claim that it is a strong argument for their theory that verbs and prepositions show two different types of defaults in the presence of an overt subject/complement. Verbs show 3.sg default agreement but prepositions do not. They show the bare stem:

\[
\begin{align*}
(35) & \quad \text{a. \ Levriouc a=lenn-Ø ar vugale} \\
 & \quad \text{books \ PTCL=\text{read}-3SG \ the \ children} \\
 & \quad \text{‘The children read books.’} \\
 & \quad \text{b. \ Ul levr brezhonek a=zo gant (/*gantañ) Yannig.} \\
 & \quad \text{a \ book Breton \ PTCL=BE with \ with.3SG.masc \ Yannig} \\
 & \quad \text{‘Yannig has a Breton book.’}
\end{align*}
\]

(Stump 1984)

According to their theory, this follows because in the case of the verbs, T probes for agreement features but finds merely the vP. And since the vP has nominal properties, it values the probe on T with default features 3sg. In contrast, a preposition does not probe for features at all and thus remains a bare stem if nothing cliticizes to it. In a sense, it can be said that the 3sg is a syntactic default which is inserted as the result of failed agreement and the bare stem is a morphological default which surfaces in the absence of cliticization.

I want to point out that while the story nicely accounts for the pattern, it is far from the only conclusion that can be drawn from the facts. Note that we cannot
really say whether verbs show 3SG-agreement or the bare stem since the forms are both identical. So, we can simply state that both verbs and prepositions occur in the bare stem in the absence of \( \phi \)-features.

But, more concretely, even if we found that verbs and prepositions exhibit different default patterns, nothing forces us to assume that these correspond to a syntactic default on the one hand and a morphological default on the other. It could simply be the case that prepositions have a fully specified 3SG-exponent (presumably because they also express gender) whereas verbs do not:

\[
\begin{align*}
\textbf{(36)} & \quad \text{Verbal VIs:} & \quad \textbf{(37)} & \quad \text{Prepositional VIs:} \\
& a. \quad [V [\phi \ 1SG]] \leftrightarrow /-\text{an}/ & a. \quad [P [\phi \ 1SG]] \leftrightarrow /-\text{in}/ \\
& b. \quad [V [\phi \ 2SG]] \leftrightarrow /-\text{ez}/ & b. \quad [P [\phi \ 2SG]] \leftrightarrow /-\text{it}/ \\
& c. \quad [V [\phi \ 1PL]] \leftrightarrow /-\text{omp}/ & c. \quad [P [\phi \ 3SG.M]] \leftrightarrow /-\text{an}/ \\
& d. \quad [V [\phi \ 2PL]] \leftrightarrow /-\text{it}/ & d. \quad [P [\phi \ 3SG.F]] \leftrightarrow /-\text{i}/ \\
& e. \quad [V [\phi \ 3PL]] \leftrightarrow /-\text{ont}/ & e. \quad [P [\phi \ 1PL]] \leftrightarrow /-\text{omp}/ \\
& f. \quad [V [\phi \ ]] \leftrightarrow /-\emptyset/ & f. \quad [P [\phi \ 2PL]] \leftrightarrow /-\text{c’}\text{h}/ \\
& & g. \quad [P [\phi \ 3PL]] \leftrightarrow /-\text{o}/ \\
& & h. \quad [P [\phi \ ]] \leftrightarrow /-\emptyset/ \\
\end{align*}
\]

I thus take it that the different defaults for verbs and prepositions can be accounted for in various ways including a straightforward morphological one and thus do not distinguish between different analyses.

5.3. Agreement with the verb eus (‘have’)

The final pattern I want to mention is the verb eus (‘have’) which constitutes the only actual counterexample against the Complementarity Principle in Breton as it shows agreement regardless of whether its subject is phonologically overt or covert:\(^{10}\)

\[
\begin{align*}
(38) & \quad a. \quad \text{Bremañ o=deus} \quad \text{Azenor ha} \quad \text{Iona un ti now} \quad \text{PTCL=have.3PL} \quad \text{Azenor and Iona a} \quad \text{house} \\
& \quad \text{‘Azenor and Iona have a house now.’}
\end{align*}
\]

\(^{10}\)Jouitteau & Rezac (2006) note that the verb eus (‘have’) can also function as an auxiliary. Its behavior with respect to the Complementarity Principle is however unaffected. Even as an auxiliary, it will always show agreement.
b. Breman ο=deus un ti
now PTCL=have.3PL a house
'They have a house now.' (Jouitteau & Rezac 2006)

Since it can be shown that this verb derives from a combination of the copula and a preposition, Jouitteau & Rezac (2006) adopt the prepositional analysis of Schafer (1994) in which the subject of 'have' is located in the specifier of a PP. The preposition itself will undergo head movement into a copula higher up in the tree.

It is not entirely clear how we accommodate these facts in our REMOVE-based account. Comparing the agreement morphology on the verb eus('have'), however, suggests that we are dealing with a rather impoverished agreement paradigm that it is quite different from the typical agreement paradigm of verbs in Breton.¹¹

(39) *Agreement of -eus:*

<table>
<thead>
<tr>
<th></th>
<th>Spoken Breton</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>meus</td>
</tr>
<tr>
<td>2.SG</td>
<td>teus</td>
</tr>
<tr>
<td>3.SG.Fem</td>
<td>neus/deus</td>
</tr>
<tr>
<td>3.SG.Masc</td>
<td>deus</td>
</tr>
<tr>
<td>1.PL</td>
<td>meump</td>
</tr>
<tr>
<td>2.PL</td>
<td>peus</td>
</tr>
<tr>
<td>3.PL</td>
<td>neunt</td>
</tr>
</tbody>
</table>

Jouitteau & Rezac (2006)

Given the facts that this kind of agreement is arbitrary both in terms of its morphological exponency and also inasmuch as it ignores the Complementarity Principle, I would tentatively like to suggest that this preposition is special inasmuch as it is the only element in the Breton language that bears a ϕ-feature probe. Thus, it is able to collect the features of the subject before the subject moves to the postverbal subject position SpecFP where it is potentially removed. I will readily admit that this is a mere restatement of the observed facts but I

¹¹The table in (i) only gives the versions in Spoken Breton. Jouitteau & Rezac (2006) also give the versions in written Breton as well as various dialect forms. As far as I can see, the other forms do not add a whole lot of additional information.
assume such a stipulation will be necessary one way or another to derive the exceptional character of this construction.12

6. Conclusion

In this paper, I put forward an argument that null subject constructions in the Celtic language Breton should be derived by means of syntactic structure removal. The argument is based on the fact that agreement in Breton is subject to the Complementarity Principle which states that only covert pronouns can control agreement on the verb. It has been shown that both types of theories are empirically and theoretically problematic for the facts in Breton. If the complementarity effect in Breton were due to a process of incorporation, this process would be fundamentally different from other incorporation/cliticization processes. If we were to derive the complementarity effects by means of actual \( \phi \)-agreement, we would face the problem that agreement seems to make reference to the phonological properties of the agreement controller; an assumption which violates the Principle of Phonology-free Syntax. Thus, both existing analyses face severe problems deriving the Breton facts.

I proposed an alternative analysis in terms of Structure Removal. I argued that pro-drop of postverbal subjects is to be derived by removing it from the derivation. But due to the reintegration property of REMOVE, this process leads to a situation where the \( \phi \)-features of the removed subject reattach to the functional head which eventually will be realized as part of the verbal spine. Crucially, this is the only option for the verb to obtain the features of the subject. Since none of the typical functional heads in Breton bear a \( \phi \)-feature probe, they will not agree unless REMOVE has taken place. Under this approach, the Complementarity Effect is straightforwardly derived. \( \phi \)-agreement is an immediate consequence of removing the subject. Alternative approaches cannot employ any similarly direct implementation without violating the Principle of Phonology-free Syntax.

12Jouitteau & Rezac (2006) assume that the PP that hosts the subject in this construction is - unlike the vP - transparent for \( \phi \)-agreement of T. Thus, the case involving have emerge as the default case in a way as they are the only configuration where Breton exhibits successful \( \phi \)-agreement with the subject. But it seems to me that it is equally stipulative to assume that this PP is exceptionally transparent for agreement (usually the arguments of PPs of course do not agree). Further, nothing is said about why the exponents of the agreement features in this case are so radically different from other verbs.
One final advantage of the present analysis, I would like to mention is that it provides a straightforward explanation for the question why the empty subject is structurally confined to the postverbal position. In typical pro-theories in which the empty pronoun is an active syntactic element, we would need additional machinery (i.e. licensing, etc.) to make sure that pro does not occur in the position before the finite verb leading to verb-initial matrix clauses, something that is unattested in Breton. In the present account, it is clear why the preverbal position cannot be occupied by the empty pronoun. The impression of an empty pronoun is created by syntactic removal of the subject in postverbal position. Thus, since the pronoun is literally removed (and not merely phonologically empty) it is clear why it cannot move to the preverbal position subsequently. This is a clear case of a syntactic opacity effect which poses a strong argument for the syntactic nature of REMOVE.

I further discussed various empirical issues in Breton including the issue of inflected prepositions. I argued (or rather, followed the argumentation in Jouitteau & Rezac 2006), that inflected prepositions which, at first sight, also obey Complementarity effects, are to be derived using more PF-like operations involving context-sensitive spell-out. This discussion showed that syntactic structure removal in a sense fits into Nevins’ (2012) typology of deletion operations applying at different points of the derivation.

References


Implementing Minimalist Syntax and Remove

Johannes Englisch

Abstract
This article presents a computational implementation of Minimalist Syntax. The program implements a fragment of German grammar, building syntactic structure from a numeration and applying syntactic, such as Move, Agree, and Structure Removal. The article will also apply this system to two specific analyses: complex pre-fields and embedded clauses, which showcase the capabilities, but also the limitations of the program.

1. Introduction

This article describes a computer program that implements a fragment of German grammar within Minimalist Syntax (Chomsky 1995). This program simulates the application of Merge, Movement, and Agreement to generate syntactic structure. In addition to these rules, the implemented system also includes the concept of structure removal (Müller 2017).

The program imitates the generation of syntactic structures from numeration to final tree. This stands in contrast to, for instance, parsing algorithms, which extract syntactic structure from a string of surface symbols.¹ The program starts by taking out a vocabulary item from the numeration and using it as a starting point, from which it builds structure from the bottom up. And whenever there are multiple ways a derivation could progress, the program branches out and pursues all possible options.

Note that Minimalism itself does not require a specific starting point for a derivation: The rule system just generates all possible combinations of all items in the numeration. However, quite a few of these derivations result in the same tree structure. In fact, for every leaf node in a tree there is a derivation that ‘started with this node’.² Spurious derivations like these do not pose a problem from a theoretical standpoint, as they all apply the same operations to the

---

¹See Harkema (2001) for a parser based on Stabler (1997)’s formalisation of Minimalist syntax.
²Also, if you take structure removal into account, then there are derivations starting with an element that does not even show up in the final tree.
same elements and produce the same dependencies – just in a slightly different order. However, they do matter for computational implementations, since recalculating the same structure multiple times is a waste of both memory and computing time.

The article is structured as follows. Section 2 outlines what will and will not be part of the program and Section 3 illustrates how the program is structured from a bird’s-eye view. Then, Section 4 describes the progression of a syntactic derivation from numeration to final tree. After this the article goes more into detail on how the program is implemented, with Section 5 showing how syntactic structure is represented with the program, Section 6 outlining the definitions of the actual syntactic rules, and Section 7 addressing some theoretical decisions made in the system. With the theoretical bits out of the way, Section 8 shows how the program behaves when applied to actual data. Finally, Section 9 concludes the article.

2. Scope of the program

As stated above, the program aims to model a fragment of German grammar. This fragment generates sentence structures and recreates common phenomena like wh movement, case assignment, and V2 word order. This generation process restricts itself to the purely syntactical part of generative grammar, meaning that the system does not attempt to implement variable binding, covert movement, clitic formation, or other phenomena living at the interface to phonology, morphology, or semantics.

The program provides an implementation to the standard rules of Minimalist Syntax: Merge, Movement, Head Movement, and Agreement. In addition the program includes Müller (2017)’s structure removal process, which ‘unbuilds’ the syntactic structures created by Merge and Move.

Note that Minimalism does not define rules specifically tailored to any particular language. Instead the rules are kept very general and the difference between languages lies mainly in the feature specification of the vocabulary items in the lexicon. For instance, the V2 word order in German results from the C head bearing a feature that attracts the finite verb via head movement and another feature that triggers topicalisation, rather than from a language-specific movement rule or word order restriction.
3. Structure of the Program

As the diagram in (1) illustrates, the program is structured into three layers. The lowest layer contains the primitive rules of Minimalist Syntax and is responsible for the creation and manipulation of syntactic tree structure. The second layer lives in the realm of derivations, where numerations are created, derivations are performed, and tree structures are judged for their technical wellformedness. On top of all of that, there is the user interface. This layer parses the input provided by the user in the Yaml format\(^3\) and displays the resulting derivations in a graphical user interface.

(1) Structure of the program:

Note that this structure isolates the rules of syntax from the rest of the system. A syntactic rule is merely a function mapping trees to trees. This means that their application is local to the structures directly involved in the process. Rules cannot refer to other elements in the work space or the numeration, nor can they access previous or future steps of the derivations, let alone different paths the derivations could have taken.

4. From numeration to derivation

Every computer program known to man turns some input into some output. The input of this program is a numeration of lexical items and a feature specification for the item the derivation should start with. The output consists a list of all derivations with a report of the technical wellformedness of the resulting tree attached to them.

\(^3\)http://yam1.org/
The program starts by looking through the numeration for any items that qualify as a starting point. These items are put into an agenda, together their own copy of the numeration. Then the program takes the tree–numeration pairs out of the agenda one by one and tries to apply the rules of syntax to them. The results of successful rule applications are pushed back onto the agenda to keep the derivation going, including a reference to the previous state of the derivation, so the process can be retraced later. If none of the rules are successful the tree–numeration pair is added to a list of final structures. This process repeats until the agenda is empty.

After finishing all derivations the program checks the technical well-formedness their final structures. It considers the final state of a derivation grammatical if:

- There are no elements left in the numeration,
- There are no unsatisfied operation triggers left in the structure, and
- There are no unvalued features left in the structure.

5. **Representation of syntactic structure**

Before we can see how syntactic structure is manipulated, we need to take a look at how these structures are represented in the first place. This section looks at the representation of phrases (Section 5.1), heads (Section 5.2), and whole derivations (Section 5.3).

5.1. **Syntactic phrases**

Syntactic structure is represented in terms of *phrases*. Every phrase consists of a head and a list of arguments as shown in (2a). This representation can be converted back and forth to a binary syntax tree, as illustrated in (2b). Note that $X'$ and XP nodes of a tree are not represented directly in the data. These nodes are created during the conversion based on the head. This gives the system a simple account for projection: Because all dominant nodes are made from the same head, any changes to that head happen to the whole projection line. However, this makes it impossible for syntactic rules to target only parts of a projection. This directly affects the way head movement works in the program (see Section 6.3).
5.2. Syntactic heads

Heads like in (3a) are represented as shown in (3b). A head consists of three parts: A phonological representation, a set of morpho-syntactic features and an ordered list of operation-triggering features. Morpho-syntactic features are a set of key–value pairs, although the value may be empty, which is indicated with the placeholder symbol $e$. Operation-triggering features are uninterpretable features, which require the application of syntactic rules. They are tuples of the type of operation, a feature the target must carry, and – if necessary – the feature value the target must match.

(3) a. ‘do’ [cat:T, case:nom, $\phi$: $\square$, $\bullet$cat:v$\bullet$ $\epsilon$case$\epsilon$ $\epsilon$case$\epsilon$ $\epsilon$phi$\epsilon$ $\epsilon$cat:D$\bullet$]

b. **Representation of ‘do’**:

{ phon: ‘do’,
  triggers: [ (p-select, ‘cat’, ‘v’),
              (agree, ‘case’),
              (agree, ‘phi’),
              (p-select, ‘cat’, ‘D’) ] }

5.3. Derivations

In this program a derivation is merely a list of discrete steps, as shown in (4). Each step consists of the current numeration and the syntactic structure built so
far. Again, this structure is the point that allows the user to look at intermediate steps of a derivation and observe the application of rules, whose effects are not directly visible in the final structure – i.e. structure removal.

(4) **Representation of a derivation:**

\[
[ \{
  \text{numeration: \{} book, read, \},
  \text{tree: } a ,
\},
\{ \text{numeration: \{} read, \},
  \text{tree: } [ a \text{ book} ] ,
\},
\{ \text{numeration: } \{} ,
  \text{tree: } [ \text{read [ a book] } ] \} ]
\]

6. **Implementation of syntactic rules**

This section looks at the implementation of Merge (Section 6.1), movement (Section 6.2), head movement (Section 6.3), agreement (Section 6.4), and structure removal (Section 6.5). As mentioned above, these rules are merely functions that take a tree – or two, in the case of Merge – as arguments and return a transformed tree structure.

6.1. **Merge**

Merge builds structure by combining two smaller trees into a bigger one. It is caused by a selection feature \(\bullet F:v\bullet\), which is deleted in the process. In the actual implementation, Merge adds the target to the argument list of the phrase that triggers the operation. This process is illustrated in (5).

(5) a. **Before Merge:**

\[
\{ \text{head: } \{ \text{phon: 'read',}
  \text{features: {'cat': 'V'}},
  \text{triggers: [ (p-select, 'cat', 'D') ] },
  \text{args: } [ ] \} \}
\]
Implementing Minimalist Syntax and Remove

b. After Merge:
   { head: { phon: ‘read’,
             features: {'cat': ‘V’},
             triggers: [ ] },
   args: [ { head: { phon: ‘it’,
                    features: {'cat': ‘D’},
                    triggers: [ ] },
            args: [ ] } ] }

6.2. Phrasal movement

Phrasal movement merges a tree with its own subtree and replaces the starting position of the subtree with emptiness e. This process is illustrated in (6). To reflect their conceptual similarities, Move and Merge are both triggered by selection features. This gives rise to situations where a derivation could both merge and move. In these cases the system favours Merge.

(6) a. Before Move:
   { head: { phon: ‘x’,
             features: {'cat': ‘X’},
             triggers: [ (p-select, ‘cat’, ‘Y’) ] },
   args: [ { head: { phon: ‘y’, features: {'cat': ‘Y’}, triggers: [ ] },
             args: [ ] },
          { head: { phon: ‘z’, features: {'cat': ‘Z’}, triggers: [ ] },
             args: [ ] } ] }

b. After Move:
   { head: { phon: ‘x’,
             features: {'cat': ‘X’},
             triggers: [ (p-select, ‘cat’, ‘Y’) ] },
   args: [ e,
           { head: { phon: ‘z’, features: {'cat': ‘Z’}, triggers: [ ] },
             args: [ ] },
           { head: { phon: ‘y’, features: {'cat': ‘Y’}, triggers: [ ] },
             args: [ ] } ] }

This raises the question how the algorithm finds a target for movement. It starts with the complement – i.e. the first member of the argument list – of the top-most phrase and checks if it meets the requirements of the selection feature. If that fails the algorithm descends down into the complement, checks its arguments from highest to lowest, and keeps doing so recursively until
it either finds a suitable target or reaches the bottom of the tree. Note that, at the present state of the system, movement is not restricted by any locality constraints.

6.3. Head movement

Head movement, on the other hand, merges the head of a tree with the head of its own subtree. It is triggered by a head selection feature \([\bullet \mathbf{F^0} \bullet]\) and moves the head of a lower phrase onto the top-most head. However, remember that any change of the head affects the whole projection line. This means that removing the head with all its features, would render the remaining phrase inaccessible for further agreement or movement operations. The program avoids this effect by moving only the phonological representation of the head, leaving the features in place, as shown in (7). Also note that this implementation of head movement obeys the Head Movement Constraint (Travis 1984), according to which a head may move only to the next dominating phrase, but not higher.

(7)  a. Before head movement:

\[
\{ \text{head: } \{ \text{phon: } 'x', \\
\text{features: } \{ 'cat': 'X' \}, \\
\text{triggers: } [ \text{(h-select 'cat' 'X')} ] \}, \\
\text{args: } [ \{ \text{head: } \{ \text{phon: } 'y', \\
\text{features: } \{ 'cat': 'Y' \}, \\
\text{triggers: } [ ] \}, \\
\text{args: } [ ] \} ] \}
\]

b. After head movement:

\[
\{ \text{head: } \{ \text{phon: } ('x','y'), \\
\text{features: } \{ 'cat': 'X' \}, \\
\text{triggers: } [ ] \}, \\
\text{args: } [ \{ \text{head: } \{ \text{phon: } e, \\
\text{features: } \{ 'cat': 'Y' \}, \\
\text{triggers: } [ ] \}, \\
\text{args: } [ ] \} ] \}
\]

6.4. Agreement

Agreement ensures that two heads share the same value for a feature. There are two main implementations of agreement. In Feature checking the numeration
Implementing Minimalist Syntax and Remove provides fully specified heads. Agreement compares two heads with respect to a given feature and crashes the derivation if their values differ. In Feature valuation the heads in the numeration may contain features without a value. Agreement fills these gaps by copying values from other heads. When the derivation is over, any unvalued feature left renders the syntactic structure ungrammatical. This system uses feature valuation, mainly because this allows the program to trim number of derivations.

Take, for instance, a sentence with two DPs bearing different number features. In feature checking the numeration contains two determiners: a plural one and a singular one. As (8) shows, the grammar generates two structures, one of which crashes because the determiners were merged with ‘incompatible’ nouns.

(8) Feature checking:

a. dass das Kind die Bücher liest
   → Success

b. *dass die Kind das Bücher liest
   → Crash

In feature valuation the numeration contains two instances of one general determiner. This means that that structure building can only produce the one derivation shown in (9). Merge attaches the two instances to the nouns and agreement fills their number features from their context. The program does not need to compute a second, failing derivation.

(9) Feature valuation:

dass das/die Kind das/die Bücher liest
   → Success

Note, however, that feature valuation struggles with cases like case concord. While feature checking can establish agreement locally within a DP, feature valuation requires a way for the case assigner to agree with multiple heads. For this the system would either have to add more probes or a mechanism for mass-valuing DPs.

Agreement is triggered by a probe feature [*F*]. The head bearing a probe
can only look downwards, when searching for a target. The search algorithm mirrors that of Move, the difference being that Agree also looks at its own specifiers for targets. Once agreement finds a target it can copy the feature value in either direction – the example in (10) illustrates downwards valuation. Because of this a T head can both value its ϕ features from a lower target and copy its case features downwards to assign case. Note that this also means that a case assigner must itself carry case.

\[(10)\]

\[a.\quad \text{Before agreement:} \]
{ head: { phon: ‘x’,
  features: {‘cat’: ‘X’, ‘F’: ‘val’},
  triggers: [ (agree, ‘F’) ] },
 args: [ { head: { phon: ‘y’,
    features: {‘cat’: ‘Y’, ‘F’: e},
    triggers: [ ] },
      args: [ ] } ] }

\[b.\quad \text{After agreement:} \]
{ head: { phon: ‘x’,
  features: {‘cat’: ‘X’, ‘F’: ‘val’},
  triggers: [ ] },
 args: [ { head: { phon: ‘y’,
    features: {‘cat’: ‘Y’, ‘F’: ‘val’},
    triggers: [ ] },
      args: [ ] } ] }

6.5. Structure removal

As the name suggests, Remove deletes the argument of a head. Like movement the operation is split into a phrasal and a head version. In the case of phrasal Remove, a head bearing a Remove feature (−F−) looks among its arguments for the highest instance of the feature F and deletes the argument. Note that the entire argument is removed, as shown in (11).
Implementing Minimalist Syntax and Remove

(11) a. Before phrase removal:
{ head: { phon: ‘x’,
  features: {‘cat’: ‘X’},
  triggers: [ (p-remove, ‘cat’, ‘Y’) ] },
 args: [ { head: { phon: ‘y’,
    features: {‘cat’: ‘Y’},
    triggers: [ ] },
        args: [ ] } ] }

b. After phrase removal:
{ head: { phon: ‘x’,
  features: {‘cat’: ‘X’},
  triggers: [ ] },
 args: [ ] }

Head Removal, on the other hand, deletes only the head of an argument. It is triggered by a Head Remove feature (−F₀−). Not that, unlike head movement, head removal also affects the dominating X’ and XP nodes. This means that the whole projection line of the argument disappears, and any elements attached to that projection line need to be reintegrated into the structure. It does so by attaching these elements as arguments of the head that triggered head removal, in place of the removed element.

(12) a. Before head removal:
{ head: { phon: ‘x’,
  features: {‘cat’: ‘X’},
  triggers: [ (h-remove, ‘cat’, ‘Y’) ] },
 args: [ { head: { phon: ‘y’,
    features: {‘cat’: ‘Y’},
    triggers: [ ] },
        args: [ ] } ] },

args: [ { head: { phon: ‘a’,
    features: {‘cat’: ‘A’},
    triggers: [ ] },
        args: [ ] } ] ] ]}
Due to its destructive nature, structure removal can change how future rules treat the arguments of a phrase. On one hand, this happens when removal applies to a complement. In the case of phrase removal the whole complement disappears and the lowest specifier takes its place, as one can see in (13). This also means that from this point on all syntactic rules will treat that element as a complement.

\[
(13) \quad \begin{array}{c}
\text{XP} \\
\text{WP} \quad \text{XP} \\
\text{ZP} \quad \text{WP} \quad \text{XP} \\
\quad \quad \text{ZP} \quad \text{XP} \\
\quad \quad \quad \text{X}^0[\text{−−}] \quad \text{YP}
\end{array}
\]

On the other hand, head removal can cause a similar effect when deleting a complement and reintegrating its arguments. As the example in (14) illustrates, only the lowest argument attaches to the complement position, while the other arguments land in specifier positions. Again, this means that these elements take on the syntactic properties of specifiers. This gives the system a second way of shifting elements up a tree besides movement.
7. Brief theoretical considerations

This section briefly addresses two topics that have been left out of the discussion thus far: First, Section 7.1 looks at the role of the linear order of elements within the program. Then, Section 7.2 looks at how the system implements – or rather does not implement – c-command.

7.1. A word on linearisation

The program linearises the structure post-syntactically, meaning that above rules are entirely unaffected by the linear order of any elements. On a technical level this is done using rules which specify the ordering of a head with its complement or its specifiers. Which rule is applied to a phrase depends on the lexical category on the head. The example in (15) shows two examples of such rules in German.

(15) a. Verb phrase
   Rule: { cmpl: left,
           spec: left }
   Tree: VP
        /\   V'
       /   \
      DP   DP

   b. Complementiser phrase
   Rule: { cmpl: right,
           spec: left }
   Tree: CP
        /\   C'
       /   \
      XP   C'
7.2. The role of c-command

At the moment there is no definition of c-command in the program. The reason for that is that it is not really needed for the given data set. In movement a landing site c-commands its starting position, because of the way movement attaches a moved element as a specifier at the top. Agreement, on the other hand, looks for the highest target that is dominated by the root node. However, this is likely to be a property of the specific data set covered by the program rather than a general principle. Some of the rules left out from the grammar fragment, like variable binding, may still be sensitive to c-command.

8. Running the program on actual analyses

This section contains two example analyses to illustrate the program more practically. Section 8.1 re-implements the remove-based analysis of double prefields by Müller (2017) to show a case, where the program progresses as expected. On the flip side, Section 8.2 goes through a derivation involving embedded clauses to show some limitations of the system. After that, Section 8.3 looks at some of the possible causes for said limitations.

8.1. A success story: double prefields

As mentioned above, the first example analyses double prefields using structure removal. This is a direct adaptation of an analysis by Müller (2017), which serves as empirical evidence for head removal. The object of the study are sentences like in (16). In this case we have a German V2 structure, except that the verb seems to be positioned in the third position. To be more accurate, Müller (2017:21–24) notes that the elements in the prefield show both signs of being multiple constituents as well as one single unit.

(16) [DP Den Fahrer] [PP zur Dopingkontrolle] begleitete ein Chaperon

‘A chaperon accompanied the rider to the doping test’

The world of structure removal answers to conflicting evidence by taking advantage of the derivational nature of the system: You can have your cake first
Implementing Minimalist Syntax and Remove and eat it later. In other words, the elements in the prefield start off as part of one constituent, which gets destroyed at a later step of the derivation; after the V2 structure has been build.

The analysis achieves this by first creating a V2 structure and then dismantling parts of it using head removal. Such a derivation is illustrated in (17): First the verb is head-moved out of the VP into the C head. Then the whole remaining verb phrase is moved to Spec-CP via topicalisation, creating a V2 structure. Finally, head removal destroys the VP projection of the topicalised verb phrase and the two arguments are reintegrated as specifiers of the CP.

\[
(17) \quad \begin{align*}
a. & \text{ chaperon [VP rider to.the.test accompany]} \quad \mid \text{h-move} \\
b. & \text{ accompany [chaperon [VP rider to.the.test } t_{V_0}] \quad \mid \text{topicalise} \\
c. & \text{ [VP rider to.the.test } t_{V_0}] \text{ [accompany [chaperon } t_{VP}]] \quad \mid \text{remove} \\
d. & \text{ rider to.the.test [accompany [chaperon } t_{VP}]]
\end{align*}
\]

The concrete implementation performs the derivation based the numeration listed in (18). Note that the two instances of the determiner the are collapsed into one entry in the numeration.\(^4\)

\[
(18) \quad \text{Numeration:} \\
\begin{align*}
a. & \quad 1\times \emptyset [\text{cat:C, } \bullet \text{cat:T} \gg \bullet \text{cat:T}^0 \gg \bullet \text{top:t} \gg - \text{top:t}^0 - ] \\
b. & \quad 1\times \emptyset [\text{cat:T, case:nom, } \bullet \text{cat:v} \gg \bullet \text{cat:v}^0 \gg *\text{case}*] \\
c. & \quad 1\times \emptyset [\text{cat:v, case:acc, } \bullet \text{cat:V} \gg *\text{case}* \gg \bullet \text{cat:D} \gg \bullet \text{cat:V}^0 \bullet ] \\
d. & \quad 1\times \text{begleiten [cat:V, top:t, } \bullet \text{cat:P} \gg \bullet \text{cat:D}\bullet ] \\
e. & \quad 1\times \text{zu [cat:P, case:dat, } \bullet \text{cat:D} \gg *\text{case}*] \\
f. & \quad 2\times \text{das [cat:D, case:} \square, \bullet \text{cat:N}\bullet ] \\
g. & \quad 1\times \text{Dopingkontrolle [cat:N]} \\
h. & \quad 1\times \text{Fahrer [cat:N]} \\
i. & \quad 1\times \text{ein [cat:D, case:} \square, \bullet \text{cat:N}\bullet ] \\
j. & \quad 1\times \text{Chaperon [cat:N]}
\end{align*}
\]

Since the program is unable to take semantics into account, it will blindly generate all possible combinations of arguments and determiners. This results in 18 sentences such as the ones in (19), all of which show the desired double-prefield construction.

\(^4\) Also, note that for the sake of clarity this analysis omits some non-essential syntactic processes such as ϕ agreement or the contraction of the preposition zu with the determiner.
8.2. A problematic story: V2 and embedded clauses

The second example derives the seemingly simple example in (20). The sentence consists of an embedded verb-final sentence with a complementiser and a matrix clause with a verb-second word order. It poses a challenge, since the material in the numeration needs to be correctly merged and moved across two clauses. The numeration is shown in (21).

(20) Tobi weiß [ dass Silke schläf ]
Tobi knows that Silke sleeps

(21) Numeration:

For this numeration the program generates four distinct derivations as in (22), which are all judged to be well-formed on a technical level. However, they all showcase some challenges for the system.

(22) a. *dass Silke [ Tobi schläf ] weiß
b. *Tobi weiß Silke, [ dass schläf ]
c. *dass Tobi [ schläf Silke ] weiß
d. Tobi weiß [ dass Silke schläf ]

The first ungrammatical example is the result of the C heads being merged in the wrong order. The phonologically empty C head of the matrix clause
happened to get merged first, ending up in the embedded clause. This leads to a V2 clause embedded within the middle field of a verb-final clause.

(23) *dass Silke [ Tobi₁ schläft t₁ ] weiß
    that Silke Tobi sleeps knows

A possible repair mechanism for this sentence is extraposition. However, whether or not extraposition yields a grammatical structure depends on the matrix verb, as shown in (24).

(24) a. *dass Silke weiß, [ Tobi schläft ]
    that Silke knows Tobi sleeps

b. dass Silke glaubt, [ Tobi würde schlafen ]
    that Silke believes Tobi would sleeps
    ‘that Silke thinks that Tobi would sleep’

The second successful derivation shown in (25) is also a false positive. In this case the DP with the [topic] feature is merged within the embedded clause and then topicalised to the specifier of the matrix CP. This arises due to the lack of locality constraints within the implemented system.

(25) *Tobi₁ weiß Silke [ dass t₁ schläft ]
    Tobi knows Silke that sleeps

The example in (26) is the result of the double duty selection features need to fulfil. This derivation merges the V2-triggering C head and the DP without the [topic] feature in the embedded clause. This means that the C head topicalisation lacks a target to topicalise. However, since selection features trigger both Move and Merge, the derivation turns to the numeration and merges the other DP. After this the topicalised argument gets moved into the matrix vP.

(26) *dass Tobi₁ [ t₁ schläft Silke ] weiß
    that Tobi sleeps Silke knows

The last derivation found by the system is shown in (27). On the surface this example looks perfect. The verb-final clause is embedded within the V2 structure and all the arguments are merged and moved in the right places.

(27) Tobi weiß t₁ [ dass Silke schläft ]
    Tobi knows that Silke sleeps
However, this result seems quite fragile. Consider a variant of the same numeration, where the T heads bear a \([•D•]\) feature causing EPP movement. Such a configuration results in the derivation shown in (28). Just like in (26), the selection feature on the embedded T head can trigger both Merge and Move. And while T has a valid target for movement, it also has a valid target for Merge, resulting in the matrix subject to be merged too early. In a later step the subject moves into the matrix clause, resulting in a structure which is string-invariant to (27). In other words the stucture looks correct on the surface, but is structurally wrong.

\[(28)\quad *_{\text{CP}} \text{Tobi}_i \text{ weiß} \quad [\text{TP} t_1 [\text{vP} t_1 [\text{CP} dass [\text{TP} t_1 [\text{vP} \text{Silke schläf}t ]]]]]]

8.3. Analysing the undesired effects

Now, let us take a look at the potential causes for the behaviour of the program in section 8.2. As mentioned above, the omission of any locality constraints allows for unrestrained long-distant movement. Note, however, that this is a property of the implementation rather than the grammatical formalism.

Another source for error lies in the tight connection of Merge and Move and their triggers. Potentially a \([•F•]\) feature can always cause either Merge or Move. The system resolves this by choosing Merge over Move, when a conflict arises. However, this ties the application of Move to the current state of the numeration:

- Movement can be blocked by elements in the numeration intended to be merged in other clauses.
- A derivation failing to merge a movement target in time for the operation can always be repaired by base-generating directly from the numeration.

This problem is amplified by the fact that Minimalism does not implement the concept of ‘clause boundaries’ on a technical level. Clauses are usually CPs, but the rule system in itself treats C like any other category.\(^5\) This turns the

\(^5\)This includes the Phase Impenetrability Condition, which defines phases in terms of phrases, not clauses (CP and VP in the case of Chomsky (2000:106–107), or any and every phrase in Müller (2003)).
Implementing Minimalist Syntax and Remove

numeration into one big bag containing all elements of all clauses, each of which could potentially bleed or feed movement operations throughout the whole derivation.

9. Conclusion

In conclusion, this article presented a program which imitates the progression of syntactic derivations within the Minimalist framework. The program succeeds in building phrase structure, including the application of movement and structure removal. However, the system reaches its limits, when the resulting structure spans across one or more clause boundaries, which might be rooted in the fact that neither the rules of grammar, nor the representation of its structures are structured in terms of clauses or clause boundaries.

References


The Fleeting DP in Bulgarian and Macedonian: The View From Left-Branch Extraction

Jelena Stojković

Abstract
This paper is concerned with restrictions on Left-Branch Extraction (LBE) in Bulgarian and Macedonian, the only two Slavic languages with articles. The data indicate that the availability of LBE is related to the number of modifiers, and thus require a revision of the generalisation that all languages with articles disallow LBE. Following Martinović (2019), Calabrese & Pescarini (2014), a solution based on the order of operations is proposed, connecting the availability of LBE to the idea that projections can be removed via syntactic operations such as Exfoliation (Pesetsky 2019). The application of Exfoliation depends on the relative timing of article placement, which in turn is performed via a post-syntactic operation of Generalised Lowering (Embick & Noyer 2001). Exfoliation and subsequently LBE are fed by an early application of Lowering, assumed to be able to interleave narrow syntax (Martinović 2019), but if applied late, Lowering counter-feeds these two operations.

1. Introduction

There has been a debate in recent years over whether the DP is projected universally in all languages, or whether there is parametric variation among languages with respect to its existence. Bošković (2005, 2008, 2014c, et seq.) has argued for the latter option, claiming that there is a fundamental difference between languages with and without the definite article. By recourse to a variety of syntactic phenomena, one of them being Left-Branch Extraction (henceforth: LBE), he argues that the differences arise from the fact that the...
former languages have a DP (like English in 1a), and the latter do not (like Serbo-Croatian in 1b-c).

(1) Left-Branch Extraction (Bošković 2005)
   a. *Whose \textit{i} John likes \textit{t}i car?
   b. \textit{Čija} \textit{se} dopadaju Petku [NP \textit{t}i kola] ?
      whose REFL.PRO like Petko car
      ‘Whose car does Petko like?’
   c. Lijepe \textit{i} je vidio [NP \textit{t}i kuće].
      beautiful AUX saw houses
      ‘He saw beautiful houses’ (Serbo-Croatian)

The generalisation is that languages with articles block extractions of left branches, while languages without articles allow for them to be extracted. In order to explain this difference, existing approaches to LBE focus on the concept of phase: assuming that DP is a phase, LBE is blocked due to the Phase Impenetrability Condition (Chomsky 2000, 2001). In languages without articles the DP projection is assumed to be absent, and NP is the highest projection in the extended domain, therefore possible restrictions on LBE would not be a consequence of PIC.

This paper joins the discussion on this phenomenon, focusing on data from Macedonian and Bulgarian (DP languages/languages with articles). The data show that LBE in these languages is allowed, but only if the extracted left-branch is the only modifier of the noun. In cases with two modifiers or more, LBE is blocked. These data are problematic for the existing accounts based on the assumption that the availability of LBE is dependent exclusively on the absence of (definite) articles / the DP projection.

I follow previous accounts in assuming that the DP projection does in fact block LBE in languages with affixal articles such as Bulgarian and Macedonian. However, in the cases where LBE is allowed, I claim that the DP is originally projected, but later syntactically removed via Exfoliation (Pesetsky 2019). Exfoliation can apply only if the principle of Recoverability of Deletion (Chomsky 1981) is not violated. Following Martinović (2019, 2017) in assuming that the post-syntactic operation of Generalised Lowering (similar to Embick & Noyer 2001; here in charge of article placement) may interleave syntax, I argue that

with two genitives, focus marking, island sensitivity, negative concord, focus movement, the sequence of Tense etc.
Exfoliation is fed by an early application of Generalised Lowering, but a late application of Generalised Lowering counter-feeds Exfoliation. Success of Exfoliation to apply leads to the application of LBE.

The paper is organised as follows: section 2 presents the crucial data on LBE in Bulgarian and Macedonian. Previous accounts of LBE, based on the concept of phase and its extensions are discussed in section 3. The core proposal is presented in section 4: subsection 4.1 is concerned with adjustments to the original proposal for article placement from Embick & Noyer (2001), relying on a different structure of the DP in Bulgarian and Macedonian. Additional assumptions are discussed in subsection 4.3, for Exfoliation, the mechanism assumed to feed LBE in this case, and subsection 4.2, for interleaving syntax and post-syntax. I show in section 5 that with these assumptions the asymmetry between cases of one modifier and two modifiers can be straightforwardly derived as a case of feeding and counter-feeding, and offer an outlook to intra- and inter-language variation in subsection 5.3. Section 6 summarises and concludes.

2. Left-Branch Extraction in Bulgarian and Macedonian

This paper looks at cases of extraction from DPs as direct objects. Languages without articles, like the majority of Slavic languages, normally allow LBE, as shown at the beginning in (1) for Serbo-Croatian. The data to be discussed here contradict previous generalisations on the ungrammaticality of LBE in languages with articles, as opposed to languages without them.

Bulgarian and Macedonian, the only two Slavic languages with articles,²

²Macedonian and Bulgarian show a suffixed definite article, always found on the leftmost constituent of the noun phrase. The article morpheme inflects for number, and gender in the singular (Leafgren 2011, Friedman 2001). The agreement mechanism is not the main focus of this paper, but see e.g. Koev (2011) on this matter.

(i)    kniga-ta       book-DEF
       interesna-ta kniga interesting-DEF book
       ADJ-DEF ... (ADJ) N (general pattern)

(ii)   Pet-te golemi crveni šapki padnaa od polica-ta.
       five-DEF.PL big.PL red.PL hat.PL fell from shelf-DEF.F.SG
       ‘The five big red hats fell from the shelf’ (Macedonian)

The article also shows allomorphy, influenced by morphological and phonological factors
were first reported in Uriagereka (1988) as not allowing LBE. The examples recurring in the literature since are given in (2) for Bulgarian and in (3) for Macedonian. In Bošković (2005) this property is strictly related to the fact that these languages have overt definite articles (i.e. project a DP).

(2) a. *Kakva$_i$ prodade Petko [ t$_i$ kola ] ?
   what kind sold   Petko    car
   ‘What kind of a car did Petko sell?’

b. *Čija$_i$ haresva Petko [ t$_i$ kola ] ?
   whose likes     Petko    car
   ‘Whose car does Petko like?’

c. *Nova-ta$_i$ prodade Petko [ t$_i$ kola ]
   new-DEF.FSG sold    Petko    car
   ‘The new car, Petko sold.’ (Bulgarian; Bošković 2005:3)

(3) a. *Kakva$_i$ prodade Petko [ t$_i$ kola ] ?
   what kind sold   Petko    car
   ‘What kind of a car did Petko sell?’

b. *Čija$_i$ ja bendisuva Petko [ t$_i$ kola ] ?
   whose it likes     Petko    car
   ‘Whose car does Petko like?’

c. *Nova-ta$_i$ ja prodade Petko [ t$_i$ kola ]
   new-DEF.FSG it sold   Petko    car
   ‘The new car, Petko sold.’ (Macedonian; Bošković 2005:3–4)

However, Stanković (2013) has come up with counter-evidence, showing that, among other phenomena, LBE is allowed in Macedonian from a definite NP (4). His data come from a corpus research of literary Macedonian and from a grammaticality judgement survey with 20 native speakers.³

(4) Crvenite$_i$ gi kupi [DP t$_i$ čevli ] ?

---
³LaTerza (2014) reports on an acceptability judgement study based on an online survey she conducted with 140 native Macedonian speakers, which has shown that LBE from definite NPs is not acceptable in this language. I return to this in subsection 5.3.
red.\text{def} \text{them} \text{bought} \text{shoes}

‘You bought the red shoes?’ \hspace{1em} \text{(Macedonian; Stanković 2013:11–12)}

The data that the analysis in the present paper relies on come from grammaticality judgements given by 11 native speakers of these two South-Slavic languages: 5 native Macedonian speakers and 6 native Bulgarian speakers, all non-linguists. The speakers come from west and central North Macedonia, and eastern and north-eastern Bulgaria.\textsuperscript{4} The participants were presented with a short text, a conversation between two people on a specific topic (an example from Bulgarian is given below in (5)) and requested to judge all of the sentences in the conversation.

(5) Sara i Marija pijat kafe. V \text{\text{\text{\v{a}g\v{a}}}la} na stajata sa obuvkite na
Sara and Marija drink coffee in corner of room are shoes of
Marija. Sara zabeljazva edni \text{\v{c}erveni obuvki}.
Marija Sara notice \text{one red shoes}
‘Sara and Marija are having coffee. In the corner of the room are Marija’s
shoes. Sara notices a pair of red shoes.’
S: \text{Obuvkite sa tvoi?}
shoes are your
‘Are the shoes yours?’
M: Ne, ne, Ana gi kupi i zabravi tuk onija den.
no no Ana them bought and forgot here that day
‘No, no, Ana bought them and forgot them the other day.’
S: \text{\v{c}ervenite kupi obuvki? Az ne gi haresvam.}
red bought shoes I not them like
‘She bought the red shoes? I don’t like them.’

According to the speakers’ judgements, Left-Branch Extraction is allowed in a definite context in Bulgarian, see (6), in the sentence equivalent to Stanković’s examples for Macedonian (4).\textsuperscript{5}

\textsuperscript{4}Stanković (2013, 2015) also reports that the Timok-Lužnica dialect of Serbo-Croatian also allows LBE in the same contexts. This dialect is structurally (and geographically) close to Macedonian and Bulgarian in morpho-syntax, by exhibiting clitic doubling, having an affixal definite article, lacking case inflection etc.

\textsuperscript{5}In the morpho-syntactic context that the data in this paper illustrate, clitic doubling is obligatory in Macedonian, but optional in Bulgarian. More on the matter of clitic doubling can be found in Franks \& Rudin (2005).
Furthermore, a quantifier in the same position, i.e. as the only modifier of a noun, can also be extracted in Bulgarian and Macedonian. Judgements from the consulted speakers for extraction of a quantifier are in line with the ones for adjectives.

However, LBE is not always allowed in these languages: if a quantifier and an adjective co-occur in a sentence, in the presence of an article, none of the modifiers can be extracted, as shown in (8) for Macedonian, and in (9) for Bulgarian.

(8) Gi prativ [DP šest-te crni / crni-te šest torbi ]
them sent six-DEF.PL black black-DEF six bags
'I sent the six black bags.' (Macedonian)

a. *Crni_i gi prativ [DP šest-te t_i torbi ]
black them sent six-DEF bags
b. *Crni-te_i gi prativ [DP t_i šest torbi ]
black-DEF them sent six bags
c. *Šest-te_i gi prativ [DP t_i crni torbi ]
six-DEF them sent black bags
d. *Šest_i gi prativ [DP crni-te t_i torbi ]
six them sent black-DEF bags

(9) Pratih [DP šest-te černi / černi-te šest čanti ]
sent six-DEF.PL black black-DEF six bags
'I sent the six black bags.' (Bulgarian)

a. *Černi_i pratih [DP šest-te t_i čanti ]
black sent six-DEF bags
b. *Černi-teᵢ pratiḥ [DP tᵢ šest čanti ]  
black-DEF sent six bags  
c. *Šest-teᵢ pratiḥ [DP tᵢ černi čanti ]  
six-DEF sent black bags  
d. *Šestᵢ pratiḥ [DP černi-te tᵢ čanti ]  
six sent black-DEF bags

The same holds for cases where there are two adjectival modifiers of the NP – no LBE is allowed, regardless of the adjective’s position within the NP and the definite article.

(10) Gi skršiv [DP mali-te beli / beli-te mali čaši ]  
them broke small-DEF white white-DEF small cups  
‘I broke the small white cups.’  (Macedonian)  
a. *Mali-teᵢ gi skršiv [DP tᵢ beli čaši ]  
small-DEF them broke white cups  
b. *Mali giᵢ skršiv [DP beli-te tᵢ čaši ]  
small them broke white-DEF cups  
c. *Beli-teᵢ gi skrši [DP tᵢ mali čaši ]  
white-DEF them broke small cups  
d. *Beli giᵢ skrši [DP mali-te tᵢ čaši ]  
white them broke small-DEF cups

(11) Sčupih [DP Malki-te beli / beli-te Malki čaški ]  
broke small-DEF white white-DEF small cups  
‘I broke the small white cups.’  (Bulgarian)  
a. *Malki-teᵢ sčupi [DP tᵢ beli čaški ]  
small-DEF broke white cups  
b. *Malki sčupi [DP beli-te tᵢ čaški ]  
small broke white-DEF cups  
c. *Beli-teᵢ sčupi [DP tᵢ Malki čaški ]  
white-DEF broke small cups  
d. *Beli sčupi [DP Malki-te tᵢ čaški ]  
white broke small-DEF cups

The data in (10) and (11) are reminiscent of the pattern found in Serbo-Croatian. Bošković (2005) notices that LBE of a single modifier of a noun is allowed,
and extraction of one modifier from a configuration with two adjectives is disallowed in Serbo-Croatian, an NP-language.

(12) *Visoke_t je vidio [NP t_ljep_t djevojke ]
    tall aux saw beautiful girls
    ‘He saw tall beautiful girls.’
    (Bošković 2005:12)

Given what we have seen so far, Bulgarian and Macedonian, DP-languages, seem to behave exactly like Serbo-Croatian, an NP-language. Additionally, in Serbo-Croatian, multiple LBE, where both adjectival modifiers of the noun are fronted, is also disallowed (13-a), but can be performed if one of them is a demonstrative (13-b).

(13) a. *Visoke_t ljep_t je on vidio [NP t_ljep_t djevojke ]
    tall beautiful aux he saw girls
    ‘He saw tall beautiful girls’
    (Bošković 2005:12)

    b. Onu_t staru_t prodaje [NP t_ljep_t djevojke ]
    that old aux sells house
    ‘He/She sells that old house.’
    (Bošković 2016:21)

This is not the case in Bulgarian and Macedonian: according to the speakers consulted, multiple LBE is disallowed, in configurations with two adjectives (14-a), in configurations with an adjective and a demonstrative (14-b), and in configurations with a quantifier and an adjective (14-c).

(14) a. *Malki-te_t žol_t prodava [DP t_ljep_t djevojke ]
    small-def yellow aux sells cats
    ‘He/She sells the small yellow cats.’
    (Bulgarian)

    b. *Tezi_t malki_t prodava [DP t_ljep_t djevojke ]
    these small aux sells cats
    ‘He/She sells these small cats.”

    c. *Tri-te_t malki_t prodava [DP t_ljep_t djevojke ]
    three-def small aux sells cats
    ‘He/She sells the three small cats’

The data from Bulgarian and Macedonian with LBE from NPs modified with a

---

For similarities of these two languages with Serbo-Croatian with respect to binding, see LaTerza (2016, 2014).

I return to the issue this set of data represents in subsection 5.3.
quantifier and an adjective or with two adjectives (8)–(11), are in line with the findings of Uriagereka (1988), Bošković (2005, 2008). On the other hand, the data in (4)–(7), with one adjective or quantifier, show that in these languages LBE is available in a restricted set of cases. The summary of the data, given below in Table 1, strongly suggests that the availability of LBE depends on the number of modifiers: if there is one, LBE can be performed, if there are two, it cannot, regardless of whether only one of them or both are to be extracted.

Table 1: The availability of LBE in Bulgarian and Macedonian

<table>
<thead>
<tr>
<th>Structure</th>
<th>LBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>[DP AP NP]</td>
<td>✓</td>
</tr>
<tr>
<td>[DP QP NP]</td>
<td>✓</td>
</tr>
<tr>
<td>[DP QP AP NP]</td>
<td>×</td>
</tr>
<tr>
<td>[DP AP AP NP]</td>
<td>×</td>
</tr>
</tbody>
</table>

3. Previous Discussions on Phase Status

The ban on extraction from the left periphery of the NP is formalised as Left Branch Condition by Ross (1967/1986), for ruling out sentences as in (15-a) for English. However, Ross himself has noticed that the constraint on this and similar extractions does not hold cross-linguistically. Namely, numerous studies have shown that Latin and many Slavic languages (illustrated in (15-b) for Russian), also Korean, Literary Finnish (15-c), Mohawk etc. allow extraction of adjectives (and demonstratives, possessives, a.o.) from NP.

(15) a. *Beautifuli he saw [ t_i houses ]
   b. Doroguju/Etu_i on videl [ t_i mašinu ]
      expensive/that he saw car.ACC
      ‘He saw an expensive/this car.’             (Russian; Bošković 2012:2)
   c. Punaisen_i ostin [ t_i auton ]
      red.ACC    buy.PST.1SG    car.ACC
      ‘I bought a red car.’                    (Literary Finnish; Bošković 2012:2)

A widely accepted explanation comes from Bošković (2005), who explores the availability of LBE, joined with other syntactic and semantic phenomena, with respect to the presence of articles. If the presence of articles equals existence of the DP layer in the structure, and if the DP is a phase, then the ban on LBE in
DP languages is a consequence of the *Phase Impenetrability Condition* (16), illustrated in (17):

\[(16) \quad \text{Phase Impenetrability Condition (Chomsky 2000:108)}
\]

In a phase $\alpha$ with head $H$, the domain of $H$ is not accessible to operations outside $\alpha$; only $H$ and its edge are accessible to such operations.

\[(17) \quad \left[ \begin{array}{c}
\text{XP ...} \ [\text{DP} \ [\text{NP} \ t_{\text{XP}} \ N] ]
\end{array} \right]
\]

As the definition of PIC states, the phase head and its edge are still accessible to operations triggered outside the extended domain, therefore the modifier could simply move to the specifier of $D$ and its extraction would be allowed. To exclude this, Bošković (2005) invokes a version of *anti-locality* (18), the ban on movements that are too short (Abels 2003, Grohmann 2003, 2011, Erlewine 2016), which blocks the alternative to violating PIC. What anti-locality can be assumed to state for present purposes is that a movement step must cross at least one maximal projection other than the one in which it is immediately contained. Therefore, extremely local Spec-to-Spec movement, as illustrated in (19), violates anti-locality.

\[(18) \quad \text{Spec-to-Spec Anti-Locality (Erlewine 2016:431,458)}
\]

$\text{A}$-movement of a phrase from theSpecifier of YP must cross a maximal projection other than YP.

\[(19) \quad \left[ \begin{array}{c}
\text{...} \ [\text{DP} \ \text{XP} \ D \ [\text{NP} \ t_{\text{XP}} \ N] ]
\end{array} \right]
\]

Given that languages without articles are assumed to not have a DP phase, and the PIC is thus not active in the extended domain of a noun, LBE is predicted to be generally available. Any restriction on its availability in an NP-language is assumed to arise from independent reasons.

For deriving the ungrammaticality of LBE in Serbo-Croatian when the noun is modified with two adjectives (20), Bošković (2005:25–26) relies on two important conceptions. First, he assumes that adjectives in Serbo-Croatian are adjuncts. That being said, when multiple adjectives are present, they are assumed to be equidistant from the noun, in terms of Chomsky (1995, 2000). In such cases a revision of the concept of Lethal Ambiguity, put forward in
McGinnis 1998, is invoked: due to the impossibility to determine a hierarchical relation between the two modifiers and decide which one to extract, LBE fails, which gives rise to ungrammaticality of (20) in Serbo-Croatian.

(20) a. *Visoke je on vidio lijepa djevojke.  
    tall be.3sg he saw beautiful girls

b. *Lijepe je on vidio visoke djevojke.  
    beautiful be.3sg he saw tall girls

(Bošković 2005:12)

However, when one of the modifiers is a quantifier, LBE is allowed in Serbo-Croatian (21), which is derived by assuming that quantifiers are heads which take nouns as their complements, additionally projecting an intermediate FP, whose presence in the structure in turn renders all modifiers of NP available for extraction, without violating PIC and anti-locality.

(21) Visokih je video pet devojaka.  
    tall be.3sg saw five girls
    ‘He saw five tall girls.’

(22) [ ... [QP AP [Q' Q [FP [F' F [NP tAP N ] ] ] ] ] ] (✓ anti-locality, ✓ PIC)

For cases like (13-b), similar to Fernandez-Salgueiro (2006) for multiple wh-movement in Serbo-Croatian, Bošković (2016) proposes that Multiple LBE is in fact focus movement, where the demonstrative / higher Spec of NP moves first to SpecFocP.9 The lower Spec of NP can be moved in the next step, and be

---

8 Recall that this is not the case in Bulgarian and Macedonian, cf. (9) and (8) – the presence of a quantifier does not change the speakers’ judgements regarding the ungrammaticality of LBE. This might be because quantifiers have a different status in Bulgarian and Macedonian than in Serbo-Croatian. In the case of LBE quantifiers in Bulgarian and Macedonian seem to fully share the fate of adjectives.

9 According to Bošković (2005), demonstratives are adjectives in Serbo-Croatian, which creates a problem for the analysis from Bošković (2016), as Murphy (2017) points out. Bošković (2014a:59) claims the difference to be a semantic effect: single referent of the focalised element and deicticity. However, for the analysis to work, the difference would have to be structural; the adjectives in (13-a) would have to be equidistant, while the demonstrative in (13-b) would have to be structurally higher than the adjective. A solution that Murphy (2017) proposes is to treat the demonstrative as an argument in the specifier of DP.
tucked in as the lower SpecFocP.

\[
(23) \quad [\text{FocP} \ AP_1 \ AP_2 \ [\text{Foc} \ Foc \ \ldots \ [\text{NP} \ t_1 \ t_2 \ [\text{N} \ N \ ] \ ]] \]
\]

If these analyses are applied to the data from section 2, one encounters several problems. First, Bulgarian and Macedonian have definite articles, which puts them into the group of DP-languages.\(^{10}\) Thus LBE should not be possible in a definite context with a single modifier, as in (4), (6), and (7), and in cases with two adjectives, as in (10) and (11). Second, should one adopt the view that quantifiers are heads taking nouns as their complements, and that adjectives are adjuncts to NP, as Bošković proposes, LBE in the sentences with a quantifier and an adjective modifying the noun, like (8), (9), should be grammatical, since the AP could make an intermediate stop in SpecDP, and subsequently move to its designated position in the structure, incorrectly deriving a part of the data.

\[
(24) \quad \ldots \ [\text{DP} \ ___ \ D \ [\text{QP} \ Q \ [\text{NP} \ AP \ [\text{N} \ N \ ] \ ]] \] \quad (\checkmark \ \text{anti-locality}, \checkmark \ \text{PIC})
\]

In sum, the account that would derive the data correctly needs to predict availability of LBE in the cases with one adjective or a quantifier, but crucially rule out the same operation in the context of two adjectives, and in the context of an adjective and a quantifier.

4. Theoretical Model

If one adopts the generalisation that LBE is only possible when there is no DP in the structure, the pattern presented in section 2 implies that the DP is

\(^{10}\)Talić (2015, 2017), looking at adverb extraction from adjectival phrases, proposes that languages with affixal articles such as Bulgarian (also Icelandic and Romanian) are NP languages, but still project a DP within the noun phrase in order to achieve the semantic interpretation of definiteness. This should be the case because the existence of articles as vocabulary items blocks the possibility of covert type-shifting, available in languages without articles (the Blocking Principle of Chierchia 1998). The proposed analysis accounts for the generalisation Talić operates with, that Bulgarian, Icelandic, and Romanian disallow LBE, but do allow Adverb Extraction from a predicative AP. Given the data on LBE in Bulgarian (section 2), one would be forced to stipulate that covert type-shifting is available with one modifier even though the article is present, but unavailable with multiple modifiers.
not projected in the first place when there is one modifier. However, this is a problematic statement, since the article is obviously present. In the same context, just with one more modifier, LBE is disallowed, meaning that the DP is present.

Therefore, I will take a more radical solution and propose that the DP was initially there, but that it was deleted at a later point, which made performing LBE possible. For this purpose I employ a syntactic operation of Exfoliation, proposed by Pesetsky (2019) (although the original idea that parts of syntactic structures can disappear in the course of the derivation comes from Ross’ 1967 Tree Pruning). The application of this operation depends on the relative timing of article placement, which in turn can be performed at different stages of derivation, depending on the number of modifiers. Thus, I will first revise the assumptions on the structure of the DPs in Bulgarian and Macedonian and the operation in charge of article placement. The interaction of article placement with Exfoliation will derive the restrictions on LBE in Bulgarian and Macedonian.

4.1. Revisiting The Structure of Bulgarian and Macedonian DPs

The traditional approach to the structure of the DP was proposed by Abney (1987), where adjectives are heads taking NPs as complements. Embick & Noyer (2001) assume the same structure for their proposal of the post-syntactic operation of Lowering, a version of Marantz’s (1988) Morphological Merger.

Embick & Noyer (2001) differentiate Lowering, which operates in terms of hierarchical structure, from Local Dislocation, assumed to refer to linear adjacency. In the rest of the subsection I will show that in order to derive the restrictions on LBE in Bulgarian and Macedonian there is no need for employing two different operations. Rather, the relative timing of a single operation, determined by independent factors, will give rise to cases in which both Exfoliation and LBE are either allowed or blocked.

The article is the head of the DP (a ‘head’ clitic in Embick & Noyer 2001), but the affixation is not done via head-movement, rather via a post-syntactic operation of Lowering: the article targets the highest head in its domain and adjoins to it, as illustrated in (25).

\[(\text{DP} \text{ D}^0 \ldots [\text{YP} \ldots \text{Y}^0 \ldots] \rightarrow \text{DP} \ldots [\text{YP} \ldots [\text{Y}, \text{Y}^0 + \text{D}^0] \ldots]\]

Due to its originally proposed timing after all syntactic operations, but before
Vocabulary Insertion, Lowering is assumed to be sensitive to the hierarchical structure, which allows the lowering head to be blind to adjuncts and thus be able to skip them. For example, in (26), the PP *ot doktora* is assumed to be an adjunct of the AP *predpisano*, thus the article skips it when Lowering applies and adjoins to A.

(26) *ot doktora predpisano-to mi lekarstvo*
from doctor prescribed-def me medicine
‘the medicine prescribed to me by the doctor’

(Bulgarian; Schürcks & Wunderlich 2003)

Applied at this stage of the derivation, Lowering cannot refer to later information, such as e.g. word order, or phonological information such as number of syllables, which are available only after Vocabulary Insertion.

However, going back to Corver (1990), but also to more recent studies (Cinque 1994, 2010, Pysz 2008, Franks 2015, 2017, Talić 2017, Stateva 2002, Dimitrova-Vulchanova & Giusti 1996, *inter alia*) it has been claimed that (Slavic) adjectives behave more like specifiers (27), rather than heads. I adopt this NP-over-AP structure. Since quantifiers exhibit the same type of behaviour with respect to the article placement and LBE, I assume they are also projected as specifiers of NP, as in (28).

(27) DP
    D
    NP
  AP
  N'
  N

(28) DP
    D
    NP
  QP
  N'
  N

Note that, given how Lowering is originally assumed to be performed, illustrated earlier in (25), with the assumed structures in (27)–(28), the D would never be able to lower to any constituent other than the noun, which is incorrect in Bulgarian and Macedonian. Therefore, I adopt the proposals of Schürcks & Wunderlich (2003), Corniles cu (2016), Corniles cu & Nicolae (2011), Ortmann & Popescu (2000) that the article targets the [+N] feature on the constituents within the DP, depicting the natural class of nouns, adjectives and quantifiers in
Bulgarian and Macedonian. For this assumption to be incorporated properly, Lowering should be defined slightly more generally than in Embick & Noyer (2001).

(29)  **Generalised Lowering:**

A lexical item $\alpha$ is attached to the right of a lexical item $\beta$ iff:

a. $\alpha$ has a feature $[\text{F}]$, and $\beta$ has a matching feature $[\text{F}]$.

b. There is a $\Gamma$ that contains $\beta$ but does not contain $\alpha$, and there is no $\Gamma'$ that contains $\alpha$ but does not contain $\beta$.

c. There is no $\beta'$ that is closer to $\alpha$ than $\beta$.

Thus, according to (29), Lowering is a feature-driven operation, subject to three requirements: feature matching, containment, and minimality. First, $\alpha$, as the host of the feature triggering Lowering, and $\beta$, as the goal for Lowering, must share a feature. For present purposes, the trigger feature can be assumed to be $[\text{+N}]$, a lexical property of D. The targeted feature $[\text{+N}]$ shows up on heads in the edge domain of NP, that is, nouns, adjectives, and quantifiers. The goal of Lowering, $\beta$, is contained in a possibly complex linguistic expression $\Gamma(')$. The requirement of containment, for hierarchical structures, is based on c-command; in linear structures, $\beta$ is a member of the string, but $\alpha$ (the host of the Lowering feature) is not. Containment is not reflexive. Minimality is determined in terms of closeness: hierarchical, if the structure is available, or linear, if only a string of lexical items is available.

Defined this way, Generalised Lowering captures standard Lowering and Local Dislocation as a single operation. The relevant difference between the

---

11The class of lexical categories targeted for D-Lowering is assumed to vary from language to language. For example, in Romanian, the article fails to lower onto a quantifier (Cornilescu 2016, Cornilescu & Nicolae 2011), but in e.g. Danish, it only ever lowers onto a noun (Hankamer & Mikkelsen 2018, Schoorlemmer 2012). In other cases in Danish the article is a free morpheme.

12Any theory of Lowering would have to have some means of defining what can the D lower to. This is here resolved by assuming that certain lexical categories are underlyingly specified with a lowering feature, while others are not. See Hankamer & Mikkelsen (2018) for a solution based on sisterhood.

Additionally, demonstratives and possessives, given that they undergo concord with the noun they modify, could also be observed as $[\text{+N}]$ categories. However, while possessives can host the article morpheme, demonstratives cannot. An additional issue is the fact that some possessive constructions demand definiteness, and that cases of double definiteness, with both the demonstrative and the definite article, show up in colloquial speech. For more on these issues, see Schürcks & Wunderlich (2003), Franks & Rudin (2005).
two operations arises as an epiphenomenon, i.e. it is a consequence of whether Generalised Lowering precedes or follows spell-out in the post-syntax.

In the cases under consideration, with one modifier, the modifier and the noun would both be eligible goals in terms of features, and the article would dock onto the adjective (30) or the quantifier (31), as the closest head at the edge of NP. If there is no modifier, the article would dock onto the noun (32).\(^\text{13}\)

![Diagram](image)

Recall that whenever in Bulgarian and Macedonian two constituents modify the noun, their linear order is not fixed, but it is the leftmost one to which the article will be suffixed. The assumption about the respective statuses of adjectives and quantifiers in the structure, illustrated in (27) and (28), implies that NPs in Bulgarian and Macedonian can have multiple specifiers. Considering the fact that their order is not fixed and that the article can be found on either of the modifiers, whichever is placed to the leftmost edge, I assume that the specifiers

\(^{13}\text{According to the original proposal of the Minimal Link Condition (Chomsky 1995), given that there is no c-command relation between NP and AP, that is, NP and QP, the noun and its modifier would be equidistant to D. I adopt a version of the Minimal Link Condition that subsumes c-command and dominance, as proposed in Kitahara (1997), Müller (1998), Rackowski & Richards (2005), inter alia.}\)
are equidistant from D (Ura 1996, Chomsky 1995, 2000). In the structure in (33) both XP and YP can be understood as either AP or QP.\(^{14}\)

(33) 

\[
\begin{array}{c}
\text{DP} \\
\quad \text{NP} \\
\quad \quad \text{D} [+N] \\
\quad \quad \quad \text{XP} [+N] \\
\quad \quad \quad \quad \text{N'} \\
\quad \quad \quad \quad \text{YP} [+N] \\
\quad \quad \quad \quad \quad \text{N'} \\
\quad \quad \quad \quad \quad \quad \text{N} [+N]
\end{array}
\]

The previous assumptions about the nature of Lowering and the structure of the DP bring the derivation into a dilemma. If minimality is determined on the basis of a tree, and the two specifiers of NP are equally close to \(\alpha\) (although the head N is not), the D head finds two [+N] goals at the edge of the NP, as in (33). Thus Lowering faces Lethal Ambiguity (McGinnis 1998). The Principle of Lethal Ambiguity states that two elements which are equidistant from a target \(K\) are lethally ambiguous for attraction by \(K\) if they are featurally non-distinct. I ascribe the same property to Lowering – the D head fails to find a unique goal because the two specifiers have the same [+N] feature that the D

\(^{14}\)Note that I deviate slightly from the original idea of equidistance, where equidistant constituents are assumed to be equal goals for a higher head. A potential evidence for equidistance would be a DP in Bulgarian and Macedonian where the D could be found on either of the two modifiers, the outer one or the inner one. I have located a couple of such examples via Google search, however, my informants have either rejected them as ungrammatical or judged the constructions as ‘suspicious’ (thus the question mark).

(i) \(\text{?(Te) se razdeljat na dve sami-te trenirovki.}\) 
   they refl.pro divide on two alone-def.pl trainings
   ‘They are divided into two separate trainings.’ (source: https://trenirai.bg/2012/11/)

(ii) \(\text{?Otnovo be pumnala dolga červena-ta kosa.}\)
    again was released long red-def.sg hair
    ‘She had released again the long red hair.’ (Bulgarian)
    (source: https://www.rulit.me/books/knizharnichka-na-ostrova-read-414306-57.html)
targets. Therefore, Lowering cannot apply. The placement of the article is then postponed.

When the phase is spelled-out and flattened (Uriagereka 1999), the article is still not properly placed, i.e. the feature [+\(N\)] is still not checked. At this stage of the derivation, minimality is determined on the basis of a bare string of lexical items, and the leftmost one is invariably closer to \(\alpha\) than all following items in the string. Generalised Lowering thus gets another chance to apply.\(^{15}\) It follows that now Generalised Lowering cannot skip constituents; thus, intervening elements cannot be ignored. In the case of Bulgarian and Macedonian, the [+\(N\)] constituent, i.e. noun, adjective, or quantifier, would receive the article if it is the leftmost.

Evidence that in the structures with multiple specifiers linear adjacency rather than hierarchy is the relevant relation comes from data with intervening constituents of a different category. Namely, in structures like (34) in Bulgarian, where the adjective is modified with an adverb, the article can never be found on the adjective, regardless of whether the AP is linearised as the leftmost or the QP, cf. (34-a) vs. (34-b), (34-c). If the article placement were performed via early Generalised Lowering, when the hierarchical structure is still visible, (34-c) would be expected to be grammatical. Instead, the D is realised as a full demonstrative pronoun, as in (34-d).

\[(34)\]
\[
\begin{align*}
\text{a.} & \quad \text{dve-te mnogo hubavi momičeta} \\
& \quad \text{two-DEF very beautiful girls} \\
& \quad (i) \quad [Q+D \text{Adv Adj N}] \\
\text{b.} & \quad *\text{dve mnogo hubavi-te momičeta} \\
& \quad \text{two very beautiful-DEF girl} \\
& \quad (i) \quad *[Q \text{Adv Adj+D N}] \\
\text{c.} & \quad *\text{mnogo hubavi-te dve momičeta} \\
& \quad \text{very beautiful-DEF two girls} \\
& \quad (i) \quad *[\text{Adv Adj+D Q N}] 
\end{align*}
\]

\(^{15}\)In this case Generalised Lowering can without consequences be called Local Dislocation, since it is subject to the identical restriction as Local Dislocation in Embick & Noyer 2001, to operate only in terms of linear adjacency. I am not assuming two different operations, because for the purposes of this account their definitions would differ only in the way minimality is determined.
d. tija mnogo hubavi dve momičeta
   th(es)e very beautiful two girls
   (i) [ D Adv Adj Q N ]

(Bulgarian)

Note that the adverb modifying the adjective is not an obstacle for the Generalised Lowering to apply early if the QP is absent from the example like the ones above (35). I take this as evidence that, when applied early, while the hierarchical structure is visible, Generalised Lowering skips adverbs because they lack the [+N] feature, as in (35-a), but if referring to linear order, available only after spell-out of the phase, Generalised Lowering cannot apply if an adverb is the leftmost constituent, as in (34-b-i)–(34-d-i).

(35) mnogo hubavi-te momičeta
   very beautiful-DEF girls
   a. [ Adv Adj+D N ]

To summarise, in this section I have revised the assumptions on the structure of the DP in Bulgarian and Macedonian, and the assumptions on operations that regulate the placement of the article. First of all, both adjectives and quantifiers are projected as specifiers of N. When only one of them is present in the structure, Generalised Lowering can apply early, because it sees the only specifier as the highest [+N] constituent (higher than the noun). In case there are two specifiers, Lowering encounters ambiguity and its application is postponed until minimality can be respected together with Lethal Ambiguity in another way. When the link is established, this time in terms of linear adjacency, Generalised Lowering applies, affixing the article to the linearly closest [+N] constituent.

4.2. Interleaving Syntax and Post-Syntax

Generalised Lowering is a PF operation, and it is usually assumed that all PF operations must follow all syntax, cf. Chomsky (2000), also The Late Lowering Hypothesis:

(36) *The Late Lowering Hypothesis* (Embick & Noyer 2001:567)
   All Lowering in Morphology follows all movement in syntax. Lowering can never remove an environment for syntactic movement.

Some recent studies have argued that post-syntactic processes might interleave
syntactic. For some authors, these would be the processes that refer to the hierarchical structure, which would be then fed back into narrow syntax and parts of the structure would potentially still be accessible to higher operations. Such post-syntactic processes could feed syntactic: for example, this was argued for object cliticisation in Northern Italian Friulian dialect by Calabrese & Pescarini (2014), and for Extraordinary Left-Branch Extraction by Radkevich (2010). The most recent arguments for interleaving come from Martinović’s (2019, 2017) work on Wolof.

Martinović (2019) focuses on the cases when the past tense morpheme -oon is present in the clause: it is then suffixed onto the verb, presumably already (head-)moved to C (37-a). The same holds independently for the negation morpheme -ul in (37-b). However, when both negation and past are present in the structure, the negation morpheme is suffixed, but the past morpheme oon is stranded below C and realised as a free constituent woon (37-c).

(37)  

(a) Xale yi lekk-oon-na=ñu jën.  
    child the.pl eat-pst-c=scl.3pl fish  
    ‘The children had eaten fish.’  
(b) Xale yi lekk-u(l)-∅=ñu jën  
    child the.pl eat-neg-c=scl.3pl fish  
    ‘The children didn’t eat fish.’  
(c) Xale yi lekk-u(l)-∅=ñu woon jën.  
    child the.pl eat-NEG-C=scl.3pl PST fish  
    ‘The children hadn’t eaten fish.’  

(Wolof; Martinović 2019:17)

Martinović (2019) interprets the data in (37) as evidence that syntactic movement can bleed post-syntactic Lowering. She argues that the verb in Wolof head-moves though every inflectional head. The past morpheme is lowered onto the verb only if the the negation is not present in the structure, as illustrated in (38), and the TP is spelled-out as the complement of CP. If both TP and NegP are present in the structure, at the moment of spell-out of the phasal complement the verb has already head-moved to SpecNegP, i.e. out of the TP domain. Lowering of the past morpheme therefore fails, as illustrated in (39). Only the negation morpheme can be affixed onto the verbal root, and the past morpheme is consequently realised as a free constituent.
The central claim I adopt from these proposals is that some post-syntactic operations can be applied to a particular domain, which is then fed back into narrow syntax, with syntactic operations targeting elements in the (partially) spelled out domain. Applied to the cases of LBE in Bulgarian and Macedonian, the object DP in the domain of a higher phase is spelled out for affixation of the article morpheme. Bulgarian/Macedonian LBE thus can be observed as evidence that post-syntax can feed a later syntactic operation in the case of the single modifier, but in the cases of multiple modifiers it counter-feeds syntax.¹⁶

4.3. Deleting Layers

It was mentioned earlier that in order for LBE to be performed in Bulgarian and Macedonian it is necessary for the DP layer to be present for article placement, but at the same time, the DP should be absent, in order to permit LBE. I propose that this is precisely what happens – the DP was deleted after article placement, via a syntactic operation. It is an old idea that parts of syntactic structure can disappear, going back to Ross’s (1967) *Tree Pruning* and the proposal from Heycock & Kroch (1994) that deletion of a trace of a head results in the disappearance of the category that dominates it, and the entire projection of the head. More recently, Stepanov (2012) has argued that the projection of a head can disappear after head-movement, and the residual material can be re-associated. Müller (2018, 2017) proposes the operation of Remove as the mirror image of Merge, subject to the same restrictions that apply to Merge. Pesetsky (2019) proposes that infinitives and reduced clauses are derived via ‘deletion of the C layer of the clause’, triggered by the interaction between a higher probe (R₁, R₂ or ⌐A) and an embedded subject (40). Probing of an embedded subject by a clause-external probe triggers Exfoliation of the C layer, as in (41), transforming a full CP into

---

¹⁶ Note that these assumptions are not contrary to the core conception of the Late Lowering Hypothesis in (36): it is the failure of Lowering to apply that blocks Exfoliation and LBE, operations which apply in the narrow syntax, not its application.
an infinitive, that is, a TP. Since TP is not a phase, the subject in Spec,TP can be
moved to a higher position in the structure without violating PIC (42).

\[
\begin{align*}
(40) & \quad \left[ V \left[ CP \left[ TP \right. \left. \text{ Subj } \left[ T' \left[ T \left[ IP \left[ I' \left[ vP \right] \right] \right] \right] \right] \right] \right] \right] \hspace{1cm} (\not\times \text{ PIC}, \not\times \text{ anti-locality}) \\
(41) & \quad \left[ V \left[ CP \left[ TP \right. \left. \text{ Subj } \left[ T' \left[ T \left[ IP \left[ I' \left[ vP \right] \right] \right] \right] \right] \right] \right] \right] \hspace{1cm} (\text{Exfoliation}) \\
(42) & \quad \left[ V \left[ TP \text{ Subj } \left[ T' \left[ T \left[ IP \left[ I' \left[ vP \right] \right] \right] \right] \right] \right] \right] \hspace{1cm} \text{(Subject extraction allowed)}
\end{align*}
\]

Thus, given that WP and YP are adjacent phases, and $\gamma P$ is not a phase, probe
$\pi$ with an EPP property can locate a goal $\gamma$ across a YP boundary, even if $\gamma$
does not occupy the edge of that YP, but, given PIC, $\gamma$ can move to $\pi$ only if $\gamma$
occupies the edge of YP. Applying Exfoliation as a repair operation removes the
phase boundary and renders $\gamma$ available for movement.

(43) \textit{Structural Description:} \\
\[
\left[ WP \ldots \beta \left[ NP \ldots \left[ \gamma P \ldots \gamma \ldots \right] \right] \right]
\]

Crucial for Exfoliation are its restrictions; namely, Exfoliation is a Last Resort
operation, applied to make a movement operation triggered by a higher head
applicable without violating PIC and anti-locality.\textsuperscript{17} For the purposes of this
analysis, Exfoliation is employed to delete the DP, hence get rid of the phase
boundary and thus enable LBE.

The question emerges how can one assume that the DP is absent from
the structure, when the article is evidently present, if one recalls data in (4),
(6) and (7), where the single modifier extracted. I argue that this is in fact
possible, because the D head is already gone from its original projection (i.e.
early Lowering has already adjoined it) when Exfoliation is supposed to be
applied. Exfoliation simply removes the residue of the DP and renders the
remaining structure accessible. In the case when D is affixed to the linearly
closest constituent, which happens much later in the derivation, Exfoliation is
blocked from applying.

\textsuperscript{17}Unlike Exfoliation, Remove is feature-driven, follows the Strict Cycle Condition, and can
apply recursively. While Exfoliation only applies downwards, Remove can also apply to an
m-commanded goal. Remove is not applicable to the present case of LBE, because, as shown in
section 2, LBE in Macedonian and Bulgarian is not made available by the presence of a lexical
category bearing the removal feature, rather by the number of modifiers of the noun.
For this reason, Exfoliation must be constrained by Recoverability of Deletion (Chomsky 1981).

Informally, this constraint states that a node can only be deleted if there is no semantic information that would otherwise not be recoverable from the remaining structure.

In the cases discussed in this paper the DP phase is the barrier for LBE, and therefore is targeted as goal for Exfoliation. In the case where minimality is determined hierarchically, and Generalised Lowering applied early (44), Exfoliation can also be applied because the Recoverability of Deletion is satisfied, i.e. the semantic information of definiteness is recoverable from the remaining structure (45).

\[
\text{(44) \quad \ldots [DP D [NP XP [N'] N ] ] ]}
\]

\[
\text{(45) \quad \ldots [DP [NP [XP X+D ] [N' N ] ] ] } (\checkmark \text{ Recoverability of Deletion})
\]

In other words, Recoverability of Deletion has nothing to say regarding deletion of a DP from which the D head is already gone, as in (45), but it would be violated if the displacement had not yet occurred. I have argued that this is the case when the noun has two modifiers. According to the proposal from subsection 4.1, due to the equidistance of modifiers, minimality cannot be established in terms of hierarchical relations, and Generalised Lowering cannot be applied early (46). When Exfoliation is in order to apply, the DP is still targeted as a barrier that needs to be removed from the structure. However, should Exfoliation apply, the information about definiteness of the object noun would be unrecoverable from the structure (47). In order to avoid violating Recoverability of Deletion, Exfoliation fails to apply.

\[18\text{The concept of deletion under recoverability (cf. Chomsky & Lasnik 1977) is also formalised as a constraint by Pesetsky (1998), in an OT-approach to the pronunciation of complementizers, and subsequently in Ackema & Neeleman (2004). The imposing of the Recoverability constraint is not per se incompatible with the original proposal in Pesetsky (1999), although it may seem so at first glance. Namely, for the purposes of deriving infinitivisation, Pesetsky makes no explicit mentions of recoverability, since the CP has to be fully removed in order for the infinitive to be created, and not left behind anywhere in the structure, as the D in the present case. However, the constraint refers to the semantic information that would, if its syntactic node is deleted, not be recoverable. In the case of D, this information would be definiteness, but there is no information of equivalent semantic relevance contained in a declarative complementizer, which brings us to a conclusion that the Recoverability of Deletion would have nothing to say against deleting a complementizer, as in Pesetsky’s original proposal of Exfoliation.}\]
In sum, the successful application of Exfoliation, the operation which deletes the phase boundary from the structure, depends on the successful application of Generalised Lowering. In the cases where Lowering has already applied, Exfoliation removes the residue of the DP projection. However, in cases where Lowering cannot establish a hierarchical link between the D and its goal, and thus must be postponed, Exfoliation fails so that the information on definiteness can be preserved.

**5. Exfoliation, Lowering, and Left-Branch Extraction**

The data presented in section 2 can be derived using the three operations motivated in the previous sections: Exfoliation, Left-Branch Extraction, and Generalised Lowering, ordered as given in Table 2.

<table>
<thead>
<tr>
<th>narrow syntax</th>
<th>post-syntax</th>
<th>narrow syntax</th>
<th>post-syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merge (DP, WP)</td>
<td>Generalised Lowering</td>
<td>Exfoliation</td>
<td>LBE</td>
</tr>
</tbody>
</table>

Table 2: Order of operations

Adopting a cyclic spell-out model (Uriagereka 1988, 1999), I assume the structure is built incrementally, and the completed cycles (phases) are subject to transfer. Transfer involves (among others) the spell-out of the complement of a phase head (Chomsky 2001, 2000). The object DP is regularly merged as the complement of VP, and entirely spelled out as a complement of the vP phase. The purpose of Generalised Lowering is to attach the article to its host. I propose that Generalised Lowering can be triggered at the moment of spell-out of the DP, to satisfy the feature [+N] on D. The spelled-out domain, fed back into narrow syntax, is still accessible, since its hierarchical structure has not yet been flattened. Exfoliation applies to resolve the dilemma between PIC and anti-locality, by deleting the DP layer. LBE can apply. After all syntax is
completed, Vocabulary Insertion and Linearisation (reflecting the c-command relations) apply.

The proposed order of operations does not make a difference between the cases when LBE is possible, with one specifier, and when LBE is blocked, with multiple specifiers. In the rest of the text I show how can this difference be derived, using the proposed order of operations.

5.1. One Specifier – LBE

In the cases with one specifier, the order of operations applied is such that Generalised Lowering precedes Exfoliation and LBE. Generalised Lowering feeds Exfoliation, which in turn feeds LBE.

(48) Generalised Lowering $\gg$ Exfoliation $\gg$ LBE

An adjective or a quantifier (here marked with a provisional XP label) is merged as the specifier of N, which is the complement of the DP. At the moment of Spell-Out of the DP, which is in turn a complement of a higher phase, Lowering is triggered by D as a phase head. D targets the XP with the [+N] feature as the hierarchically closest constituent, and adjoins to it, cf. (50).

(49) \[ vP \]
\[ ... \]
\[ VP \]
\[ V \]
\[ DP \]
\[ D_{[+N]} \]
\[ NP \]
\[ XP_{[+N]} \]
\[ N' \]
\[ N_{[+N]} \]

(50) \[ vP \]
\[ ... \]
\[ VP \]
\[ V \]
\[ DP \]
\[ D_{[+N]} \]
\[ NP \]
\[ XP_{[+N]} \]
\[ N' \]
\[ N_{[+N]} \]

The entire DP, together with higher parts of the phasal complement, is fed back into narrow syntax. At this point of the derivation LBE cannot be performed.
because the DP, as a phase, renders the structure below inaccessible to higher operations, due to the dilemma imposed by PIC and anti-locality. As a last resort, Exfoliation applies, by targeting the DP layer as the phasal boundary. The information of definiteness is retrievable from the rest of the structure, since the D head has lowered to the modifier, therefore the DP shell can be deleted, as shown in (51). Now LBE may be performed (since the PIC does not hold for this domain any longer) by moving the XP to SpecCP (52), together with the article.\textsuperscript{19} All of this is exactly as argued by Pesetsky (2019) for CP deletion in infinitives.

\textsuperscript{19With Pesetsky (2019), I assume here that the phase status is also completely gone with the projection that held it, as in Chomsky’s original conception of phases. Another way of implementing Exfoliation of DP would be to say that the phase status is afterwards assigned to NP, as the highest in the extended domain (assuming the contextual approach to phases, cf. Bošković 2014\textit{b}, Bošković 2014\textit{c} et seq.). The specifier, as the phase edge, would still be accessible for extraction, however, extraction from the right would be banned, such as the ban on extracting the genitive complement of a noun in Serbo-Croatian (i).

(i) *[ Ovog studenta, sam pronašla [NP knjigu, ti ]
\begin{tabular}{ll}
this.GEN.SG & student.GEN.SG \\
NP & AUX found \\
knjigu & book.ACC.SG
\end{tabular}

‘I found this student’s book.’ (Serbo-Croatian)

Bulgarian and Macedonian have no such NP–under–NP constructions, and the meaning above would be expressed most probably with a PP, which seem to be extractable, cf. (26), also LaTerza (2014:ch.3.4.2). Although extractability of NP-complements could be understood as an argument in favour of not having the phase at all after applying Exfoliation, I leave the issue for subsequent research.
In the post-syntax the vocabulary items are inserted and the structure is linearised, giving rise to the configuration as in (53), with the adjective/quantifier extracted together with the article.

(53) Vocabulary Insertion and Linearisation

5.2. More Specifiers – No LBE

In the cases when the object NP has two specifiers, Generalised Lowering can be successfully applied only after spell-out of the CP as a whole. The late application of Lowering creates a counter-feeding order of operations.

(54) \( \times \) Generalised Lowering \( \gg \) Exfoliation \( \gg \) LBE \( \gg \) check Generalised Lowering

In the second type of cases an adjective and a quantifier, or two adjectives, are merged as multiple specifiers of NP. The specifiers are marked with XP and YP in (55) et seq. When this DP is spelled out (55), Generalised Lowering could also be triggered by the [+N] feature on D. However, in (56) Lowering cannot apply because the two specifiers are equidistant from the D head and both carry the [+N] feature. In other words, minimality cannot be determined based on hierarchical relations. Due to Lethal Ambiguity (McGinnis 1998), Lowering cannot apply (56).
In the next step, Exfoliation would need to be applied, but it is blocked for one crucial reason: Lowering has not applied, and the D head is still in its original projection. Deleting the DP now would mean losing the information about definiteness of the object NP, thus violating Recoverability of Deletion. Exfoliation fails to apply (57).

Since Exfoliation has failed to apply, the DP layer remains in the structure and performs its role of a phase: due to PIC and anti-locality LBE is blocked. This is illustrated in (58).

(57)
Note that after all these operations have failed, the article is still in its original position, meaning that the feature $[\uparrow +N]$ has not been checked. Since Lowering has not applied earlier (56), the morpheme is forced to wait until after Linearisation. Once linear adjacency is established as a link between the D and its goal, the article will be associated with the immediately adjacent constituent (assuming the constituent is specified with a $[+N]$ feature), as shown in (59).

(59) Vocabulary Insertion, Linearisation,

$$\left[ \begin{array}{l} C \quad \ldots \quad D \quad X[+N] \quad Y[+N] \quad N[+N] \end{array} \right]$$

Generalised Lowering

In sum, LBE in Bulgarian and Macedonian is dependent on the outcome of Exfoliation, which is in turn dependent on the outcome of Generalised Lowering in the first cycle of spell-out – if Lowering applies while the hierarchical structure is still visible, so can Exfoliation and LBE. If Lowering does not apply, Exfoliation is blocked, and subsequently, LBE. Thus, article placement can feed Exfoliation
and, then, LBE, if it is performed by early Generalised Lowering, or counter-feed these operations, if done via late applied Generalised Lowering.

5.3. Outlook – Deriving Variation

The interest in the LBE data is motivated by my own observation that speakers of Bulgarian and Macedonian I am exposed to use the constructions with LBE in the unmarked discourse. LaTerza (2014:ch.3.4.2), however, reports on an online-study based on acceptability judgements she had conducted with 140 speakers of Macedonian. The sentence from LaTerza’s study in (60), with an adjective extracted from a definite object noun phrase, is structurally identical to the sentences in (4), from Stanković (2013) for Macedonian, and (6), my own data from Bulgarian, however the judgements conflict.²⁰

(60) *Vredni-ot Jovana go zapozna student.
    diligent-def Jovana him met student
    ‘Jovana met the diligent student.’ (mean 1.72; LaTerza 2014:200–204)

Marked word orders cross-linguistically have been often reported to give rise to decrease in acceptability, processing difficulties, also reported to be less frequent in corpora and acquired later than unmarked word orders.²¹ It is possible that, had LaTerza provided some appropriate context for the participants in her study, she would get the same results as Stanković and myself (also see footnote 20).

Another way to look at these differences is in terms of inter-speaker variation. In syntactic approaches to word order in German (such as Haider & Rosengren

²⁰In the case of Stanković (2013), his data come from two sources: corpora and judgements of 20 native speakers, where the sentences were placed in appropriate context. LaTerza reports that the participants in her study were presented with sentences without any prior context, and also states (on pages 205 and 209) that the collaborator on the study, I. Stojanovska, has informed her that, given the right context, sentences with LBE are grammatical for her. Speakers who participated in my study have all been presented with sentences with LBE in a conversational context.

²¹Experimental studies (Altmann & Steedman 1988, Steedman 2000, Weber & Neu 2003, Hyönä & Hujanen 1997, Lau et al. 2017, Bernardy et al. 2018, inter alia) have shown that appropriate discourse context or even repetition (as in e.g. Nagata 1988) can positively influence acceptability and even neutralise any increase in processing difficulty. Non-canonical order sentences are, according to these studies, expected to involve more processing cost and therefore lower the acceptability grades when no preceding context is provided.
2003, Frey 2006, inter alia), a marked word order is derived from the unmarked word order via syntactic operations. A growing body of proposals account for variation in subject-verb agreement in terms of order of syntactic (and post-syntactic) operations (Murphy & Puškar 2018, Assmann et al. 2015, Müller 2009, Marušič et al. 2015, Bhatt & Walkow 2013, Driemel & Stojković 2019), where different possible orders / different timing of operations give rise to different agreement strategies.

Such an approach can be implemented in the case of the two sets of data in opposition. In one set of data, LBE is allowed only with a single modifier (data from Stanković 2013, myself, I. Stojanovska); in the other, LBE is blocked in all cases (data from LaTerza 2014). The two orders of operations proposed in (48) and (54), which derive LBE of a single modifier, are based on the assumption that Generalised Lowering may interleave operations performed in narrow syntax, but is not necessarily required to. I propose that for the variety represented by LaTerza’s speakers, the order of syntactic and post-syntactic operations is such that Generalised Lowering does not interleave narrow syntax – rather, it is ordered and applied after the entire CP is spelled out.

If the order of operations where Generalised Lowering follows all operations in narrow syntax, the feature [ ↗↑ + N] could be checked only when the linear order is made available. Therefore, Exfoliation, the operation which applies in narrow syntax and could enable LBE, would be blocked due to Recoverability of Deletion, regardless of the number of specifiers. Note that the different timing of Generalised Lowering with respect to operations in narrow syntax does not change anything for its original function of article placement.

A potentially problematic set of data comes from Bašić (2004), who claims that in both Bulgarian and Macedonian LBE is possible, with the article affixed to the noun, as in (62-a), (62-b), and (63). LaTerza (2014) also discusses these examples, claiming that these are not cases of LBE, rather that these adjectives function as “depictive secondary predicates”.

(62)  

(a) Nova ja prodade kola-ta (toj).
   new it sold car-DEF (he)
   ‘He sold the car new.’
The Fleeting DP in Bulgarian and Macedonian

b. Visoki gi haresva momičeta-ta.
   tall them like girls-DEF
   'He likes the girls tall.'  (Bulgarian; Bašić 2004:96)

(63) Nova ja Petko prodade kola-ta.
   new it Petko sold car-DEF
   'Petko sold the car new.'  (Macedonian; Bašić 2004:96)

The answer to why these date are grammatical lies in the position from which the adjectives are extracted. As secondary predicates, they are complements of N, as in (64).

\[
\text{(64) } [\text{VP } V [\text{DP } AP [D' D [NP N t_{AP} ] ] ] } (\checkmark \text{ anti-locality, } \checkmark \text{ PIC})
\]

Extraction from this position would be performed in two steps: in the first step, the AP would move to SpecDP, which does not violate (Comp-to-Spec) anti-locality since the AP crosses a maximal projection – NP. Then, in the second step, without violating the PIC, the AP can be moved out to a higher position. This movement of AP does not interact with Generalised Lowering, since NP is already the higher goal for the article.

Cross-linguistic data show that the generalisations on LBE in Bulgarian and Macedonian, reported in section 2, do not hold for other languages with an affixal definite article. For example, Romanian, a DP language with an affixal article does not allow LBE (Petroj 2018). Many Germanic languages with an affixal definite article (Swedish, Norwegian, Faroese, Danish and Icelandic, cf. Santelmann 1993, Schoorlemmer 2012, Hankamer & Mikkelsen 2018), also disallow LBE (except for some special wh-extraction cases). For these languages, I would assume that Generalised Lowering does not interleave narrow syntax. It follows that the article is lowered later in the derivation, after Exfoliation had been blocked by the Recoverability of Deletion constraint, which renders LBE unavailable.

For example, it was mentioned earlier that the article in Danish only ever lowers to a bare noun. In other cases it is a free morpheme (Hankamer & Mikkelsen 2018, Schoorlemmer 2012). In Icelandic the article always appears on the noun, regardless of whether it is modified or not (Schoorlemmer 2012). None of the languages allow LBE. In the present theory, in Icelandic, only
nouns would bear the feature [+N], which the article targets, and Generalised Lowering, ordered after spell-out of the entire CP, would still counter-feed Exfoliation and LBE. For Danish, it would be necessary to assume that the AP and NP are equidistant from D, which would force Generalised Lowering to wait until linear order is established. This timing of Generalised Lowering would also counter-feed Exfoliation and LBE.

It is well known and discussed that, for speakers of English, sentences with extracted left-branches are unacceptable. I assume the reason why lies in the nature of the article in English. Namely, the affixal article lowers, but the one in English does not, simply because there is no such rule in the morphology. Exfoliation cannot apply for the same reason as above: deleting the intervening DP would result in losing information on definiteness, which could not be recovered from the remaining structure.

The account proposed in this paper does not contradict the prediction that LBE should be generally allowed in languages without articles, and that restrictions on its availability must arise from independent reasons. Such a language is Serbo-Croatian, which allows LBE, except in the case of multiple modifiers. Bošković (2005) explains this by recourse to equidistance of modifiers and Lethal Ambiguity.

However, as shown earlier, Serbo-Croatian does allow multiple LBE when one of the modifiers is a demonstrative, recall (13-b). For these cases only, Bošković (2016) proposes that Multiple LBE is in fact focus movement, where the demonstrative / higher Spec of NP moves first to SpecFocP. The lower Spec of NP can be moved in the next step, and be tucked in as the lower SpecFocP.²²

\[
\begin{array}{c}
\text{(FocP) AP}_1 \text{ AP}_2 \text{ Foc' Foc ... [NP } t_1 \ t_2 \ [N' \ N ] \ ]] \\
\end{array}
\]

According to Bošković (2005), demonstratives are adjectives in Serbo-Croatian, which creates a problem for the analysis from Bošković (2016), as Murphy (2017) points out. Bošković (2014a:59) claims the difference to be a semantic effect: single referent of the focalised element and deicticity. However, for the analysis to work, the difference would have to be structural; the adjectives in

(13-a) would have to be equidistant, while the demonstrative in (13-b) would have to be structurally higher than the adjective. A solution that Murphy (2017) proposes is to treat the demonstrative as an argument in the specifier of DP. If that is the case, the demonstrative would be available for extraction, but the DP phase would still block extraction of the adjective.

\[
\text{(66) } \quad \ldots \ [\text{DP Dem D} [\text{NP AP N}]]
\]

In the configuration as in (66), multiple LBE can be derived if the before mentioned assumptions from Bošković (2016), Murphy (2017) are enriched with Exfoliation. Namely, multiple LBE would apply in two steps. In the first step, the demonstrative is extracted from the specifier of DP (67). Exfoliation applies as a last resort (68), deleting the DP and rendering the adjective available for extraction, which subsequently happens (69).

\[
\text{(67) } \quad \ldots \ [\text{DP Dem D} [\text{NP AP N}]] \quad (\text{LBE } \#1; \checkmark \text{ PIC})
\]

\[
\text{(68) } \quad \text{Dem} \quad [\text{DP } t_{Dem} \text{ D} [\text{NP AP N}]] \quad (\text{Exfoliation}; \checkmark \text{ Recoverability of Deletion})
\]

\[
\text{(69) } \quad \text{Dem} \quad \ldots \ [\text{NP AP N}] \quad (\text{LBE } \#2)
\]

Two remarks have to be made. First, in (68) the DP can be deleted completely because the demonstrative has moved out of the DP, and therefore, the information on definiteness is recoverable. Second, the extraction of the adjective in (69) still requires tucking in, given the superiority effect. The solution I am giving is only a sketch, since the status of demonstratives in NP languages is not completely clear; however, it illustrates once more how the relative timing of Exfoliation can give rise to different extraction effects cross-linguistically.

---

6. Conclusion

Data from Bulgarian and Macedonian challenge the generalisations on differences between languages with and without articles – Left-Branch Extraction is allowed in definite contexts, but iff there is one modifier of the noun. With two modifiers LBE is blocked, including Multiple LBE. This pattern is problematic for phase-based approaches (Bošković 2005, 2016, 2014b) since the DP is required to be both present and absent at the same time – present, for an article to be inserted, absent, in order not to block LBE.

I argue that the availability of LBE is dependent precisely on the presence/absence of the DP layer, which can be syntactically removed via Exfoliation (Pesetsky 2019), but not if the semantic information is not recoverable from the remaining structure. The successful application of Exfoliation is dependent of the relative timing of Generalised Lowering, an operation in charge of article placement. The specific restrictions on availability of LBE in Bulgarian and Macedonian (the number of modifiers) are accounted for in terms of order of operations, in such a way that a post-syntactic operation of Generalised Lowering is assumed to be able to interleave narrow syntax (Martinović 2019). In the case with one modifier, Generalised Lowering is applied at the moment of spell-out of the DP-domain, so Exfoliation and subsequently LBE are thereby fed by Lowering. In the cases with two modifiers Generalised Lowering is applied later, when the linear order of constituents is available, which counter-feeds Exfoliation and, subsequently, LBE.

Therefore, apart from being able to account for the generalisations imposed by the data, an advantage of the account presented here is in that it correctly extends to inter-language and cross-linguistic variation, with minor adjustments.

References


Bošković, Željko (2014b). Now I’m a phase, now I’m not a phase: On variability of phases with extraction and ellipsis. Linguistic Inquiry 45(1). 27–89.


URL: [http://doi.org/10.5334/gjgl.584](http://doi.org/10.5334/gjgl.584)


URL: DOI: http://doi.org/10.5334/gjgl.193


URL: https://doi.org/10.1007/s11049-017-9374-z


Stanković, Branimir (2013). DP or NP? The Case of Serbian Southeastern Dialects and Macedonian Language. Ms., University of Niš.


Feature Deletion by Head Movement – A New Solution to Agreement Asymmetries in Modern Standard Arabic

Anke Himmelreich*

Abstract
In this paper, I present a new approach to agreement asymmetries in Modern Standard Arabic. In Arabic clauses with a VS word order—unlike clauses with SV order—, the verb is not marked for number agreement. I argue against previous approaches that suggest that this lack of number agreement is due to the absence of syntactic Agree in these configuration. I further propose a new account that is based on the assumption that the lack of number agreement arises from feature deletion triggered by head movement in cases where the verb satisfies the EPP property. I show that this approach does not face the same problems as previous approaches and I discuss some implications that the theory has for EPP movement.

1. Introduction
Verb-subject agreement in Arabic marks up to three categories on the verb: person, gender, and number. In Modern Standard Arabic (MSA), number agreement depends on the position (and the form) of the subject: Preverbal subjects agree in number with the verb (1a), while postverbal subjects do not agree in number (1b). Person and gender agreement are not affected by the position (and the form) of the subject.

(1) a. at-ṭaalibaat-u ṭakal-*at/na
   the-students.F.PL-NOM eat.PST-*3F.SG/3F.PL
   ‘The students ate.’

   SV (Benmamoun 2000:121)

*I would like to thank the participants of the syntax colloquium at Leipzig University, the participants of the workshop “Dependencies in Syntactic Covariance” (Leipzig University, 2019) and the audience and reviewers at GLOW 42 (University of Oslo, 2019) for valuable comments and discussion. I am especially grateful to Aya Al-Ghanem, Andrew Murphy, Gereon Müller, Limor Naaman, Nassim Saleh Obeid and Philipp Weisser. All errors are my own.

*I would like to thank the participants of the syntax colloquium at Leipzig University, the participants of the workshop “Dependencies in Syntactic Covariance” (Leipzig University, 2019) and the audience and reviewers at GLOW 42 (University of Oslo, 2019) for valuable comments and discussion. I am especially grateful to Aya Al-Ghanem, Andrew Murphy, Gereon Müller, Limor Naaman, Nassim Saleh Obeid and Philipp Weisser. All errors are my own.

Structure Removal, 385–418
Andrew Murphy (ed.)
LINGUISTISCHE ARBEITSBERICHTE 94, Universität Leipzig 2019
This agreement asymmetry poses quite a puzzle for current syntactic and morphological theories. The reason is that the empirical generalization combines linearity – a concept usually not or only partially associated with syntax – with grammatical function – something that is in most frameworks associated with a certain syntactic position. Put differently, the question for linguistic theories is what kind of theoretical operation agreement asymmetries arise from.

Despite the large number of different analyses proposed so far for the data in (1), the full pattern of agreement asymmetries in MSA has not been derived. The accounts in the literature can be classified according to two assumptions: The first type of account assumes that the agreement asymmetry is due to an early syntactic process: Number agreement only ever applies in certain syntactic configurations. The second type of account assumes that the agreement asymmetry is due to some late morphological process: Number agreement is unrestricted in the syntax, but postsyntactic operations lead to the loss of the number marker in certain configurations.

In this paper, I show that syntactic accounts overgenerate because they depend on the presence of both a postverbal and a preverbal subject position. I discuss structures with the past progressive auxiliary showing that full agreement can occur without there being a postverbal subject position. As for morphological accounts, I show that they undergenerate because they depend on some form of adjacency. I discuss examples which show that the agreement asymmetry does not depend on the adjacency of the subject and the verb. Ultimately, this paper introduces a “hybrid” approach that overcomes both the overgeneration and the undergeneration problem. The new approach is itself entirely syntactic, but shares with the morphological approach that number agreement applies in all configurations, but is ultimately lost in the derivations which result in VS orders.

The account is based on the old idea formulated in Alexiadou & Anagnostopoulou (1998) that verb movement in pro-drop languages can check the EPP-property. In contrast to Alexiadou & Anagnostopoulou (1998), I argue that the agreement asymmetry in MSA is directly linked to this assumption: First, I suggest a head movement operation that results in the deletion of both
the head movement triggering feature on the higher head as well as the goal feature of the lower, moving head. Second, I assume that the EPP-feature in MSA is not a D feature but a number feature. Finally, if the verb checks the EPP-property as part of head movement (leading to a VS word order), the verb will lose its number feature that it has received from an earlier application of agreement with the subject. If the subject checks the EPP-feature (leading to a SV word order), the number feature stays on the verb since it is not involved in head movement.

The paper is organized as follows. In section 2, I introduce the data and empirical generalization about agreement asymmetries in MSA. Section 3 presents the new “hybrid” approach and shows detailed derivations of the examples introduced in section 2. In section 4, I discuss previous analyses and compare them to the new approach. Finally, section 5 provides a conclusion and an outlook on future research.

2. Data

This section summarizes and exemplifies all the empirical observations that any analysis of agreement asymmetries in MSA should be able to account for. The section starts with some examples that illustrate the main points (section 2.1). Afterwards, in section 2.2, I introduce some standard assumptions about clause structure in MSA. Finally, section 2.3 addresses the empirical generalization that can be derived from the data. This generalization will be the starting point for the new approach.

2.1. The Agreement Asymmetry in Modern Standard Arabic

The data in (2) show that verbs do not show number agreement if the subject linearly follows the verb.

(2) a. ?akal-at/*na aṭ-ṭaalaibaat-u¹²
eat.PST-3F.SG/*3F.PL the-students.F.PL-NOM
‘The female students ate.’

VS (Benmamoun 2000:121)
b. wašal-a l-raʔiis-aani ?ila dimashqa
   arrive.PST-3M.SG the-president.M.DU-NOM in Damascus
   yesterday
   ‘The two presidents arrived in Damascus yesterday.’
   VS (Ryding 2005:66)

c. al-ʔawlaad-u raʔ-at bint-un
   the-boys-NOM saw-3SG.F girl-NOM
   ‘The boys, a girl saw them.’
   OVS (Mohammad 2000:50)

d. ?akala at-tuffaaʕίatu al-ʔawlaadu
   eat.PST-3M.SG the-apple the-children
   ‘The children ate the apple.’
   VOS (Benmamoun 2000:132)

In all the examples in (2), the verb has to show up with a number marker for
singular, which is the default marker. It cannot bear the plural marker (2a,c,d)
or the dual marker (2b). Furthermore, (2c) shows that number agreement
cannot target the object instead of the subject. Finally, (2d) illustrates that
default agreement also shows up if the object is scrambled in front of the
subject.

Moving on to cases with preverbal subjects, it can be observed from the
examples in (3) that preverbal subjects require number agreement with the
verb.

---

1 All the examples from the literature that appear in this paper have been checked with at least
one speaker of MSA. If needed, the examples where corrected. Thanks to Nassim Saleh Obeid
for pointing out mistakes. Any remaining errors are my own.

2 The original examples for (2a,c-d) are given in (i).

(i) a. ?akal-at/*na tatʔalibaat-u
   eat.PST-3F.SG/*3F.PL the-students.F.PL-NOM
   ‘The students ate.’
   (Benmamoun 2000:121)

b. ?al-ʔawlaad-u raʔ-at-hum bint-un
   the-boys-NOM saw-3SG.F-them the-girl-NOM
   ‘The boys, a girl saw them.’
   (Mohammad 2000:50)

c. ?akala t-tuffaaʕίatu l-ʔawlaadu
   eat.PST-3M.SG the-apple the-children
   ‘The children ate the apple.’
   (Benmamoun 2000:132)
(3) a. at-ṭaalaibaat-u ʔakal-*at/na\(^4\)
the-students.F.PL-NOM eat.PST-*3F.SG/3F.PL
‘The female students ate.’

SV (Benmamoun 2000:121)

b. ?ayy u at-ṭullaab ʔaraf-uu/*ʔarafa
which.NOM the-students knew-3M.PL/knew.3M.SG
l-ʔijaabat-a?
the-answer-ACC
‘Which students knew the answer?’

\(S_{wh}\) VO (Alotaibi & Borsley 2013:10)

In (3a), we see the counterpart to (2a) with a preverbal subject and obligatory plural agreement. It can furthermore be observed that gender agreement is not affected by the position of the subject (compare (2a) and (3a)). (3b) shows that if the subject is wh-moved, number agreement shows up as well.

Next, cases like (4) show what happens in clauses that have two verbal elements: an auxiliary and a lexical verb.

(4) a. kaanat/*kunna at-ṭaalaibaat-u ya-ʔkulna\(^5\)
be.PST.3F.SG/*be.PST-3F.PL the-students.F.PL-NOM 3-eat.F.PL
‘The female students were eating.’

Aux SV (Benmamoun 2000:121)

b. at-ṭaalaibaat-u *kaanat/kunna ya-ʔkulna
the-students.F.PL-NOM *be.PST.3F.SG/be.PST-3F.PL 3-eat.F.PL
‘The female students were eating.’

\(S Aux\) V (Benmamoun 2000:121)

\(^4\)The original examples for (3a-b) are given in (i).

(i) a. t-ṭaalaibaat-u ʔakal-*at/na
the-students.F.PL-NOM eat.PST-*3F.SG/3F.PL
‘The students ate.’

(Benmamoun 2000:121)

b. ?ayy u tullaab-in ʔaraf-uu/*ʔarafa l-ʔijaabat-a?
which-NOM students-GEN knew-3M.PL/knew.3M.SG the-answer-ACC
‘Which students knew the answer?’

(Alotaibi & Borsley 2013:10)
c. *kaan-at/*kun-na ya-/?kul-*at/*na at-ṭaalibaat-u
   *was-3SG.F/*was-3PL.F 3-eat.*SG.F/*PL.F the-students.F-NOM
   'The female students were eating.'
   *Aux V S (Aya Al-Ghanem, p.c.)

The first observation we can make from (4) is that both the auxiliary and the lexical verb are marked for number agreement. Furthermore, we see that the auxiliary and the lexical verb can differ in number agreement: In (4a), where the subject occurs in between both verbs, the auxiliary bears default singular agreement, while the lexical verb shows plural agreement. In (4b), with the subject in the clause-initial position, both verbs show number agreement. Finally, (4c) shows that the subject cannot appear after the lexical verb in constructions with an auxiliary. We can conclude from the pattern in (4) that both verbs independently agree with the subject and that agreement asymmetries also occur with auxiliaries.

The final set of data concerns the type of the subject. As shown in (5), the verb obligatorily bears number agreement if the subject is a pronoun.

(5) a. *kaanat/kunna ya-/?kulna
   *be.PST.3F.SG/be.PST-3F.PL 3-eat.F.PL
   'They (female) were eating.'
   Aux V (Benmamoun 2000:126)

   b. (hum) qara?-u ad-dars-a
   they read-3.pl.m the-lesson-acc
   'They read the lessons.'
   S_pro V (Soltan 2006)

   c. qara?-u/*qara?-a (hum-u) d-dars-a
   read-3.pl.m/read-3.sg.m they-EV the-lesson-acc
   'They read the lessons.'
   VS_pro (Soltan 2006)

5The original examples for (4a–b) are given in (i).

(i) a. kaanat/*kunna t-ṭaalibaat-u ya-/?kulna
   be.PST.3F.SG/*be.PST-3F.PL the-students.F.PL-NOM 3-eat.F.PL
   'The female students were eating.'
   Aux SV (Benmamoun 2000:121)

   b. t-ṭaalibaat-u *kaanat/kunna ya-/?kulna
   the-students.F.PL-NOM *be.PST.3F.SG/be.PST-3F.PL 3-eat.F.PL
   'The female students were eating.'
   S Aux V (Benmamoun 2000:121)
Feature Deletion by Head Movement

(5a) illustrates that MSA is a pro-drop language. Under the reading of a plural pronominal subject, plural agreement on both the auxiliary and the lexical verb is obligatory. In (5b-c), the pronominal subject is overt, which is possible but requires a special context. (5c) shows that, even with the pronominal subject in postverbal position, number agreement is obligatory. This is a complication for nearly all theories of agreement asymmetries since (5c) is a counterexample to the simple generalization that verb do not show number agreement if they precede the subject.

In the next subsection, some standard assumptions about clause structure in MSA are introduced.

2.2. Basic Clause Structure in Arabic

Having introduced the basic facts about agreement asymmetries, this subsection addresses standard assumptions about clause structure in MSA. In accounts couched in a derivational minimalist framework, it is often assumed that verb-initial word orders come about by movement of the verb to the functional head T. Subject-initial word orders are often assumed to be due to EPP movement of the subject to Spec-TP.6

In order to get a better understanding of what the data in section 2.1 mean for a derivational theory, I adopt the structures in (6) for a verb-initial word order and (7) for a subject-initial word order:

---

6See Aoun et al. (1994) for an analysis where the verb is in C and Wurmbrand & Haddad (2014) for an analysis where the verb may stay in the vP. See Alotaibi & Borsley (2013) for an analysis where subject movement is not linked to EPP movement. This issue is discussed in more detail in section 4.3.
(6) *Clauses with Postverbal Subjects* (adapted from Benmamoun 2000:128):

![Diagram of CP TP vP T+V DP_subj v' v VP t_V]

(7) *Clauses with Preverbal Subjects* (adapted from Benmamoun 2000:129):

![Diagram of CP TP T' DP_subj T+V vP v' v VP t_V]

The structures in (6) and (7) show that verb movement to T is obligatory while subject movement to Spec-TP is optional. This optionality eventually results in different word orders. Relating these assumptions to the data above, the pattern summarized in the table in (8) emerges.
The next subsection briefly discusses the empirical generalization that can be drawn from the table in (8).

2.3. Empirical Generalizations about the Agreement Asymmetry

Benmamoun (2000:128) summarizes the pattern of agreement asymmetries as in (9) (henceforth ‘Benmamoun’s Generalization’):

(9)  Number-Suffix Generalization (Benmamoun’s Generalization)

The number suffix is obligatory whenever the postverbal subject position is phonologically null.

Except for the case of (5c), where an overt pronominal plural subject appears after a verb marked for plural agreement, this generalization is true. Theoretically, however, this generalization poses a challenge as there are at least four different types of phonologically null postverbal subject positions: The absence of a postverbal subject position (10a), a dropped pronominal subject (10b), nullness derived by A-movement (10c), and nullness derived by ˘A-movement (10d).

(10)  a. Null due to absence of a position  

[TP T+Aux [vP Subject ... V __ ]]  

b. Base-generated null  

[TP T+V [vP pro ... ]]
Furthermore, Benmamoun’s Generalization in (9) comprises various factors that affect number agreement in MSA into one generalization. In order to stress the complexity of the agreement asymmetry in MSA, (11) breaks his generalization down into five smaller observations. Additionally, (11) serves as a summary of the data discussed in this section.

\[(11)\]
\begin{enumerate}
\item Number agreement is different from gender and person agreement.
\item Number agreement is dependent on the linear order of subject and verb.
\item Number agreement does not require surface adjacency between subject and verb.
\item Agreement is dependent on the grammatical function of the agreement target: Only subjects can agree.
\item Number agreement is dependent on whether the subject is a full nominal phrase or a pronoun.
\end{enumerate}

In section 3, I develop a new approach to the agreement asymmetry with the goal to account for all the observations in (11).

### 3. Analysis

The main idea of the new “hybrid” approach is that number agreement itself is not constrained by any syntactic configuration. Instead, the feature encoding number agreement on the verb can get lost throughout the syntactic derivation. That is, in all the cases in the table in (8), the verb agrees with the subject in number. The special assumption, however, is that number agreement applies early in the derivation in the vP domain before the verb moves to its final position in the TP. If the verb moves to T, it is able to check the EPP property of T (cf. Alexiadou & Anagnostopoulou (1998). If it does, the subject has to remain in the vP, which results in a VS order on the surface. But this derivation comes at a price: The number feature on the verb is deleted. On the other hand, no deletion applies if the subject checks the EPP property.
This section is structured as follows: In section 3.1 I will lay out the main assumptions of the new analysis. Section 3.2 puts the assumptions together and shows how the basic pattern in (1) is derived. Finally, section 3.3 shows how the more complex cases with auxiliaries, wh-subjects, and pronominal subjects can be accounted for.

3.1. Assumptions

The present account is couched within a derivational minimalist framework (Chomsky 1995 et. seq.). Features that trigger the structure-building operations Merge or Move will appear in bullets below [\(\bullet F \bullet\)], while features triggering Agree will appear with an underlined value [\(F: \underline{\_}\)]. There are four processes that are essential for the new approach: head movement, EPP movement, agreement, and morphological realization of features. I discuss each of these four points below.

3.1.1. Head Movement

Starting with head movement, I assume that it is syntactic movement that forms a complex head (Baker 1988, Chomsky 1993, 1995). Head movement is triggered by a feature \([H F_H]\) on the higher head. The head movement configuration is schematized in (12).

\[
\text{(12) Head Movement:}
\begin{align*}
[XP \ldots X_{[X, H, Y_H, \ldots]} \ [YP \ldots Y_{[Y, \ldots]} \ldots]] & \Rightarrow \\
[XP \ldots [X_{[X, H, Y_H, \ldots]} \ Y_{[Y, \ldots]}] \ [YP \ldots]]
\end{align*}
\]

Before head movement applies, both the higher head \(X\) and the lower head \(Y\) are independent. After \(Y\) undergoes head movement to \(X\), \(X\) and \(Y\) form a complex head in the position of \(Y\).

Crucially, I assume that in these configurations, the movement triggering feature as well as the targeted feature are deleted. This deletion operation is defined in (13).

\[
\text{(13) Complex Head Feature Deletion (CoHFeD):}
\begin{align*}
\text{In a complex head [X Y], if X bears an operation-triggering feature [F] and Y bears a matching feature [F], delete [F] on both X and Y.}
\end{align*}
\]
Going back to our abstract schema in (12), Complex Head Feature Deletion leads to the following structure.

(14) **Application of CoHFeD I:**

\[
[XP \ldots [X[X_{h\,Y_{H},\ldots}]\,Y_{[Y,\ldots]}] \,YP\ldots]] \Rightarrow \\
[XP \ldots [X[X_{h\,Y_{H},\ldots}]\,Y_{[Y,\ldots]}] \,YP\ldots]]
\]

As shown in (14), CoHFeD is responsible for the deletion of the categorial feature of the moved head, which is the feature that has been targeted by head movement. This formalizes the idea that the moved head incorporates into the higher head. In other words, the moved head is impoverished and loses its ability to project further (see also Lahne 2009, Keine 2010 for instances of syntactic impoverishment). This solves a potential labeling conflict between the two heads, by creating an asymmetry. After deletion, only the higher head that triggered movement (X in (14)) can project further. Note further that CoHFeD enforces the deletion of as many features as possible. In this sense, it is an instance of a “Maximize Satisfaction” principle (see Müller 2016, Driemel & Stojković 2017 for the opposite concept of “Minimize Satisfaction”). Thus, if X were to have another operation-triggering feature that Y could check, Y would lose this feature as well. This scenario is shown in (15).

(15) **Application of CoHFeD II:**

\[
[XP \ldots [X[X_{h\,Y_{H},\ldots}]\,Y_{[Y,F,\ldots]}] \,YP\ldots]] \Rightarrow \\
[XP \ldots [X[X_{h\,Y_{H},\ldots}]\,Y_{[Y,F,\ldots]}] \,YP\ldots]]
\]

As we will see below, this is exactly how the number feature on the verb is lost.

### 3.1.2. EPP Movement

I assume that the SV order in MSA comes about by EPP movement. The EPP feature on T in MSA is, however, not a D feature, but a number feature ([•#•]).\(^7\) I argue that it is this property of MSA that allows the verb to check the EPP feature. If the EPP feature were simply a D feature, it would be unclear why the verb could satisfy it, assuming that inflected verbs in general do not have nominal properties (but see Alexiadou & Anagnostopoulou (1998) for this assumption). If the EPP feature can in principle be a \(\phi\)-feature, the proposal

\[^{7}\text{I discuss the implications of this assumption in section 5.}\]
that verbs can satisfy the EPP property becomes more intuitive because both subject DPs and verbs bear $\phi$-features.

Note that surface case does not seem to be connected to EPP movement in MSA, but is, most likely, an independent process. Evidence for this assumption comes from examples with postverbal subjects bearing nominative case (16a) and preverbal subjects bearing accusative case (16b).

(16) a. ?akal-at/*na aṭ-ṭaalibaat-u
eat.PST-3E.SG/*3F.PL the-students.F.PL-NOM
‘The students ate.’

(benmamoun 2000:121)

b. ?inna n-nisaa?-a daxal-na makatib-a-hunna
that the-women-ACC entered-FEM-PL office-PL.ACC-their.FEM
‘that the women entered their office’

(Ackema & Neeleman 2003:726)

3.1.3. Agreement

As is standard in minimalist frameworks, I assume that agreement results from the application of the syntactic operation Agree (Chomsky 2000). In contrast to the standard definition of Agree, I assume that probe features on a head H can find a goal in the m-command domain of H (see e.g. Baker (2008)). Alternatively, Agree is allowed to probe upwards (see Wurmbrand 2012, Zeijlstra 2012, Himmelreich 2017, Bjorkman & Zeijlstra to appear). Importantly, I assume that a valued probe does not delete but is accessible to further operations (Legate 2005, Assmann 2012).

Regarding subject-verb agreement in MSA, I assume that Agree applies in the vP before head movement or EPP movement: v bears a $\phi$-probe that finds matching $\phi$-features on the subject. Importantly, I assume that the object is not accessible to this $\phi$-probe, either because it is inactive due to abstract (not morphological, see page 397) accusative case or because v has a second $\phi$-probe for the object.

3.1.4. Morphological realization

Finally, I assume that functional material (heads and features) is realized postsyntactically. For the sake of concreteness, I assume a standard version of Distributed Morphology (Halle & Marantz 1993). Throughout this paper,
nothing more will be said about the details of morphological realization as the agreement asymmetry is argued to arise in the syntax and does not seem to be influenced by any surface-related process.

3.2. Proposal

After having laid out the special assumptions about head movement, EPP movement, and agreement, this subsection puts the pieces together. The idea of the analysis is the following: Agree between \( v \) and the subject is always carried out in the \( vP \). This step is shown in (17).

\[
\begin{align*}
\text{(17)} & \quad \text{vP} \\
& \quad \text{DP}_{\text{subj}}[\pi : 3, \gamma : f, #:\text{pl}] \\
& \quad \text{v}' \\
& \quad v \\
& \quad V[N] \\
& \quad V[HV, \pi : 3, \gamma : f, #:\text{pl}] \\
& \quad \text{DP}_{\text{obj}} \\
\end{align*}
\]

In (17), the verb has first head-moved to the functional head \( v \). The complex \( V+v \) is built and \( V \) loses its V-feature due to an application of CoHFeD. Afterwards, \( v \) Agrees with the subject in gender, person, and number.

After Agree, both \( v \) and the subject bear a valued number feature that can be targeted by further operation-triggering features. When \( T \) is merged, it has a feature \( [HV_H] \) for head movement of \( v \) and an EPP feature \( [\bullet\#\bullet] \). At this point, there are two options how the derivation could continue: Either head movement applies first or EPP movement applies first.

If head movement is carried out first, CoHFeD enforces the deletion not only of \( [HV_H] \) on \( T \) and \( [v] \) on \( v \), but also of \( [\bullet\#\bullet] \) on \( T \) and \( [\#] \) on \( v \) (see the discussion above (15)). This derivation is shown in (18).
Since head movement of v to T checks the EPP feature in (18), the subject cannot move to Spec-TP anymore. Thus, early head movement results in a VS order. At the same time, because of CoHFeD, v loses its [v]-feature and its number feature. Consequently, the verb in a VS clause does not show number agreement because the absence of a number feature leads to the insertion of the default singular marker. This derivation derives the example in (1b).

The second option for a derivation is that EPP movement is carried out first. If this is the case, the subject, being closer to T than the v head, moves to Spec-TP. At this point, the number feature [•#•] on T is deleted. Head movement can apply afterwards, but only the v-features on v and T will be subject to CoHFeD. Thus, the number feature remains on v and will be realized post-syntactically. This derivation, which is the derivation for (1a) is shown in (19).
Note that in order to make the derivation in (19) compatible with the Strict Cycle Condition (SCC, Chomsky (1973)), one needs to assume a relaxed version of the SCC (cf. Richards (1997)), defining a cycle, for example, as a phrase.

Before moving on to the more complex cases of agreement asymmetries, I would like to briefly mention that scrambling of the object in front of the subject as in (2d), repeated in (20), does not necessarily pose a problem for the analysis. Similar to agreement in the vP, one has to assume that the object cannot be targeted by EPP movement. This might be due to the object being inactive because of abstract case or because of Criterial Freezing (Rizzi 2006).

(20) ?akala at-tuffaasiatu al-?awlaadu
    eat.PST.3M.SG the-apple the-children
    ‘The children ate the apple.’

3.3. Deriving the full pattern

So far, the analysis is able to derive cases where the postverbal subject position is null due to A-movement (see (10c)). In order to account for the other cases, a couple of minor assumptions need to be added. The main idea of the approach, however, will stay intact. The cases derived in this subsection are nullness due to the absence of a postverbal position (10a) in structures with auxiliaries, nullness do to Á-movement (10d), and nullness due to a dropped pronominal subject (10b). This section provides detailed accounts for all of these cases. Additionally,
the exception to Benmamoun’s Generalization (number agreement with a postverbal pronominal subject) is addressed.

3.3.1. Two ϕ-Probes in Past Progressive

Starting with the auxiliary case, I assume that the past progressive auxiliary kaana (‘to be’) realizes the complex head Prog+T (Bjorkman 2011). Prog is a functional head encoding progressive that is merged in between the vP and the TP. Furthermore, I assume that in cases where this additional functional projection ProgP is present, V only moves to v.8

In order to derive the two independent instances of agreement in past progressive configurations, a second ϕ-probe is needed. I assume that Prog introduces this additional ϕ-probe. Also, Prog hosts a kind of EPP feature [•D•] (or [•#•]), which attracts the subject to Spec-ProgP.

Finally, I assume that Prog moves to T. Thus, in all its properties regarding agreement, head movement and EPP movement, Prog has the same properties as v.9

Having established the assumptions about the syntax of the past progressive auxiliary, we can now turn to the derivations of the data in (4), repeated in (21).

(21) a. kaanat/*kunna at-taşalibaat-u ya-ʔkulna be.PST.3F.PL/*be.PST-3F.PL the-students.F.PL-NOM 3-eat.F.PL
   ‘The female students were eating.’

b. at-taşalibaat-u *kaanat/kunna ya-ʔkulna the-students.F.PL-NOM *be.PST.3F.SG/be.PST-3F.PL 3-eat.F.PL
   ‘The female students were eating.’

c. *kaan-at/*kun-na ya-ʔkul-*at/*na at-taşalibaat-u *was-3SG.F/*was-3PL.F 3-eat.*SG.F/*PL.F the-students.F-NOM
   ‘The female students were eating.’

First, (22) shows how Agree between Prog and the subject applies in ProgP.

---

8Note that this implies that the verb does not move to T in present progressive either. See Benmamoun (2000:57), Bjorkman (2011:68) for arguments that this might be true.

9That Prog can move to T means either that there are two types of T (one attracting Prog, the other attracting v) or that the head movement feature on T is a more general feature that is present on both Prog and v. The first option requires an additional assumption that filters out movement of v in the structure [T[H[HV]] [Prog][Prog] [v[v]]], for instance a version of the Head Movement Constraint (Travis 1984) that bans non-local head movement.
Note that at the point shown in (22), v has already Agreed with the subject and now bears plural number. Since this feature cannot be deleted in the TP, the lexical verb must show up with plural agreement. Also, the structure in (22) straightforwardly accounts for the fact that the subject cannot follow the lexical verb in these structures (21c): The subject is simply base-merged higher than the final position of the lexical verb.

After Agree in ProgP has applied, Prog moves to T and checks the EPP feature. This is shown in (23). This derivation is exactly like (18), except that not v but Prog moves: The number feature on Prog is deleted due to CoHFeD and the auxiliary shows singular agreement on the surface. Thus, example (21a) follows without further ado.
Finally, (24) shows the derivation of (21b). Similarly to the derivation in (19), the subject checks the EPP feature and number on Prog is retained.

(24)

To sum up, this section has shown that the present approach to agreement asymmetries can easily be extended to capture the pattern with auxiliaries.
3.3.2. *wh*-Subjects

As shown in (3b), repeated in (25), a subject does not have to be in Spec-TP in order to trigger full agreement on the verb.

(25) ?ay at-tullaab ?araf-uu/*?arafa l-?ijaabat-a?
    which.NOM the-students knew-3M.PL/knew.3M.SG the-answer-ACC
    ‘Which students knew the answer?’

This, however, poses a potential problem for the present account: Assuming that movement to CP is triggered by some movement feature [•F•] on C, the derivation in (26) should be possible, contrary to fact.

(26) a. Head movement of v to T: Deletion of the number feature
    \[ \text{[TP } [V+v[\#] T] [\text{VP DP}_{\text{wh}} ]] \]

b. movement to CP: SV order
    \[ \text{[CP DP}_{\text{wh}} C[\text{\{\{\{\|\}}} \text{[TP } [V+v[\#] T] [\text{VP t}_{\text{DP}} ]] \]

In (26), the order VS and partial agreement is established in the TP, but later movement of the subject in the CP changes the word order to SV.

Assuming that the present account is in principle correct, the following assumptions need to be added in order to avoid the derivation in (26): First, movement to Spec-CP has to go through Spec-TP. This might be due to TP being a phase (cf. among others Sportiche 1989, Takahashi 1994, Agbayani 1998, Bošković 2002, Boeckx 2003, Müller 2004, Boeckx & Grohmann 2007, Chomsky 2005, 2008, Richards 2011, Assmann et al. 2015).

Second, I assume that the theory of edge feature insertion and deletion in Müller (2010, 2011) is correct. The theory is based on two important assumptions: Edge feature insertion on T must apply as long as T still has operation-triggering features and edge features must be discarded right after they are inserted. Consequently, edge feature movement must apply first, if a head has an edge feature.

Finally, I assume that elements that are moved to the specifier of a head H check as many features as possible on H.\(^{10}\)

---

\(^{10}\)This is another instance of the “Maximize Satisfaction” Principle, see the discussion on page 396. Ultimately, it might be possible to define a general constraint for both head movement and phrasal movement. However, maximizing satisfaction leads to feature deletion (CoHFeD) on moved heads, while it does not lead to feature deletion on moved phrases. A way to reconcile
With these assumptions in place, a subject has to undergo movement to Spec-TP, if it has to stay accessible for processes outside of Spec-TP. Movement to Spec-TP, however, will result in full agreement on the verb.

The tree in (27) illustrates why the derivation in (26) is ruled out by the assumptions above.

(27)

If head movement is the first operation to apply in the TP, all operation-triggering features on T are deleted. In this case, no edge feature can be inserted and the subject stays in the vP. Then, however, it will not be accessible to the C head and the derivation crashes.

Thus, if a subject obligatorily has to move to a preverbal position, edge feature insertion must apply before head movement, as shown in (28). This, however, will result in full number agreement on the surface.

---

both movement types might be to invoke the general mechanism of feature checking that is used Minimalist Grammars (Collins & Stabler 2016).
Finally, the last case that needs to be discussed involves pronominal subjects. Essentially, I assume that pronouns in MSA, whether they are null or overt, need to move to Spec-CP. Thus, the derivations for the examples in (5), repeated in (29), are equivalent to the derivation in (28) for wh-subjects.

As for null subjects (29a), I follow McFadden & Sundaresan (2016) and assume that pros must be licensed by an aboutness topic that is in a projection above TP. Thus, pro has to move to Spec-TP because it has to stay visible for the aboutness topic.

As for overt pronouns, I assume that they must move to a focus projection in the C-domain. As before, movement must go through Spec-TP, leading to
full agreement on the verb. Evidence for overt pronouns being related to focus comes from the fact that they must be emphasized and have contrastive focus
(Soltan 2006, Al-Ghanem, p.c.).

A potential problem with this analysis is posed by (29c), where the pronoun follows the verb. Note that this is a problem that occurs in most other analyses as well since the generalization about the agreement asymmetry is violated. Following Ackema & Neeleman (2003), I assume that in contexts with an overt pronoun, the overt pronoun is just a tonic double for a null (preverbal) focused pronoun. A concrete implementation of this idea could be the following: Focused pronouns are complex, consisting of a null pronoun and an overt pronoun, as shown in (30).

$$\begin{align*}
(30) & \quad \text{DP}_{\text{pro}} \\
& \quad \text{pro}_{\text{foc}} \quad \text{DP}_{\text{pro}} \\
\end{align*}$$

Now, either just the null pronoun or the entire complex moves to the C-Domain. If just the null pronoun moves, the overt pronoun is stranded in a postverbal position. Nevertheless, movement of the null pronoun suffices to involve edge feature insertion, which checks the EPP feature. Since the verb does not check the EPP feature on T, it can retain its number feature. This derivation is depicted in (31).

$$\begin{align*}
(31) & \quad \text{TP} \\
& \quad \text{pro}_{\text{foc}}[\ldots,\#:\text{pl}] \\
& \quad \text{4. EF-mov.(Del. [EF], \{••\})} \\
& \quad \text{T} \\
& \quad \text{vP} \\
& \quad \text{V+v[\ldots,\#:\text{pl}]} \\
& \quad \text{T'[T,\text{EF-insertion}} \\
& \quad \text{3. EF-insertion} \\
& \quad \text{v'} \\
& \quad \text{5. HM (Delete [v])} \\
\end{align*}$$
3.4. Interim Summary

This section introduced the new “hybrid” approach, which is syntactic but has the spirit of a morphological approach. It has been shown how this new approach can derive the pattern generalized in section 2.3. The core idea of the approach is that number agreement itself is not constrained, but that the feature encoding number agreement on the verb can get lost throughout the syntactic derivation, namely in cases where the verb checks the EPP feature. If the subject checks the EPP feature, the number feature on the verb survives. The main observations are repeated in (32).

(32) a. Number agreement is different from gender and person agreement.
    b. Number agreement is dependent on the linear order of subject and verb.
    c. Number agreement does not require surface adjacency between subject and verb.
    d. Agreement is dependent on the grammatical function of the agreement target: Only subjects can agree.
    e. Number agreement is dependent on whether the subject is a full nominal phrase or a pronoun.

All the observations in (32) are derived by the approach developed in this section: (32a) follows because the EPP feature in MSA is a number feature. And because the EPP feature is a number feature, number on the verb is subject to CoHFeD in the TP. Gender and person, on the other hand, are not affected. (32b) follows because subjects that end up in Spec-TP (or in Spec-CP) are linearized to the left of the verb. If the subject is in such a position, the number feature on the verb could not have been deleted. (32c) follows because the approach is syntactic and does not make use of the concept of adjacency. (32d) follows from the assumption that the object is inactive both for agreement with the verb and for satisfying the EPP feature.

Finally, (32e) follows since pronouns have to obligatorily move through Spec-TP. With these ideas in mind, we can now turn to the discussion of previous approaches.

11This assumption will receive some discussion in section 5.
4. Previous Analyses

Agreement asymmetries in MSA have received a lot of attention over the last three decades. There are two types of accounts that have been proposed so far: First, there are syntactic analyses that assume that in certain configurations, number agreement does not arise to begin with. Second, there are morphological analyses, which argue that number agreement always applies, but that it is disguised by postsyntactic processes that rely on linear order and adjacency.

In section 3, I have developed a new approach that is in between both types: Like in morphological analyses, number agreement is modeled as a regular syntactic process, which applies early in the derivation, but is manipulated by later processes. And like in syntactic analyses, operations that affect number agreement are entirely syntactic. Morphological processes do not play a role at all. In this section, I will go into some detail about the range of theories proposed so far in the literature and compare them to the present theory with respect to their empirical coverage.

The section is structured as follows: Section 4.1 discusses previous syntactic analyses. Section 4.2 compares the new approach to morphological accounts. Finally in section 4.3, I address a more general question, namely whether preverbal subjects in MSA can be in Spec-TP at all.

4.1. Syntactic Analyses

Syntactic analyses can be further divided into two subtypes: Analyses that assume a derivational relation between SV and VS and analyses that do not assume such a relation. As for the first type, it has been proposed that either VS is derived from SV (Aoun et al. 1994, Wurmbrand & Haddad 2014) or that SV is derived from VS (Kobayashi 2013, Bjorkman & Zeijlstra (2014), Preminger & Polinsky (2015), Fakih (2016)). Each theory ultimately derives the agreement asymmetry from an interaction of movement and agreement processes in such a way that SV yields full agreement, while VS leads to partial agreement.

Theories that do not assume a derivational relation between SV and VS generally suggest that one of the two structures (VS or SV) involves a null pronoun (or a null expletive) which is the actual target of agreement. If pro is assumed to be in the preverbal position Mohammad (1990), it is defective for number agreement. If pro is in a postverbal position (Soltan 2006, Al-Horais
2009, 2012, Alotaibi & Borsley 2013) it is the only possible target for number agreement.

This brief summary suffices to identify a crucial problem of the syntactic approaches proposed so far: All of them rely on the presence of two subject positions: a postverbal one and a preverbal one. In cases where no postverbal position is available, as we have seen in the examples (4) with auxiliaries (see also the discussion in section 3.3.1), these accounts run into a problem: It is not clear how the plural agreement on the lexical verb comes about in these cases because all of these theories locate the single probe for number agreement on the T head. If one wants to maintain this idea, the only way out would be to assume a biclausal structure. While this is in principle possible (even though it is less than desirable), the question would be how the subject can be excluded from following the lexical verb. This would require additional assumptions, in those accounts, while it follows in a fairly straightforward way in the present account.12

4.2. Morphological Analysis

The first type of morphological analysis assumes that the number feature is deleted under adjacency of verb and subject in a VS order: After full syntactic agreement, the number feature is targeted for some deletion process in a VS order if the number feature on the respective verbal head is close enough to the subject (Benmamoun 2000, Ackema & Neeleman 2003).

The second type assumes that number agreement is not a syntactic operation to begin with: Number agreement is special and requires a postsyntactic matching process under adjacency (Walkow 2010).

While morphological analyses can in principle overcome the problem of syntactic accounts, they have another fairly obvious problem: In general, they cannot derive cases where number agreement is deleted (or comes about) and the verb and the subject are not adjacent. (33) illustrates that neither the

12 Note that Kobayashi (2013) does not have the same problem since movement of the subject is not a prerequisite for agreement in this approach. However, the theory has to assume that the number feature is not visible to the verb if the subject is postverbal (or c-commanded by the verb), while number is visible for agreement, if the subject is preverbal (see also Bahloul & Harbert 1992), Harbert & Bahloul (2002). Ultimately, such approaches have to ensure that agreement happens at the right point in the derivation, which is a difficulty, the present approach does not face.
postverbal subject (33a) nor the preverbal subject (33b) have to be adjacent to the verb.

\[(33) \quad \text{a. } \text{?akala at-tuffaafiata al-} \text{-?awlaadu} \\
\text{eat.pst.3m.sg the-apple the-children} \\
\text{‘The children ate the apple.’} \quad \text{(Benmamoun 2000:132)} \\
\text{b. } \text{?al-walad-u lan yasil-a} \\
\text{the-boy-nom will.not come.3m.sg.subj} \\
\text{‘The boy will not come.’} \quad \text{(Mohammad 2000:30)}\]

A second (minor) point is that, in general, it is more difficult to model the importance of the grammatical function under a morphological, surface-oriented approach.

To conclude, all types of previous approaches suffer from problems and difficulties that the present approach does not face.

4.3. EPP Movement in MSA?

One potential counterclaim that can be made against the present approach (as well as against all accounts locating the preverbal subject in Spec-TP) is that the subject in SV structures is not in TP, but actually in a higher position in the C-Domain. Approaches like Soltan (2006) and Alotaibi & Borsley (2013) try to show that a preverbal subject is a topic. Based on examples such as (34), they argue that the preverbal subject can only be definite.\(^{13}\)

\[(34) \quad \text{a. } \text{al-} \text{-?awlaad-u jaa?uu} \\
\text{the-children-nom came.3m.pl} \\
\text{‘The children came.’} \quad \text{(Alotaibi & Borsley 2013:9)} \\
\text{b. } \text{*?awlaad-un jaa?uu} \\
\text{children-nom came.3m.pl} \\
\text{‘Children came.’} \quad \text{(Alotaibi & Borsley 2013:10)}\]

However, the picture is more complex. An indefinite preverbal subject is possible if it is modified by an adjective (35a) or simply conjoined (35b), even if all conjuncts are non-specific. Also, sometimes preverbal subjects are out despite being definite (35c) (e.g. because there is no case and/or no pragmatic

\(^{13}\)Note that the sentence is only slightly marked for some speakers (Al-Ghanem, p.c.).
discrimination between object and subject). It is not obvious how all these configurations fall under the term “topic”.

(35) a. walad-un tawiil-un jaa?a
   boy-NOM tall-NOM came.3M.SG
   ‘A tall boy came.’
   (Mohammad 2000:11)

b. walad-un wa-rajul-un jaa?aa
   boy-NOM and-man-NOM came.3M.DU
   ‘A boy and a man came.’
   (Mohammad 2000:12)

c. *muusaa ?iisaa qaabala
   Musa  Isa  met.3M.SG
   ‘Musa met Isa.’
   (Mohammad 2000:4)

Furthermore, assuming that the complementizers ?anna and ?inna are in C, it is unclear why the subject appears to the right of these complementizers and not to left, as would be expected if the subject were in very high projection in the clause.

   said.3M.SG Ahmed-NOM that  Ali-ACC came.3M.SG
   ‘Ahmed said that Ali came.’
   (Mohammad 2000:19)

   said.3M.SG Ahmed-NOM that came.3M.SG Ali-ACC
   ‘Ahmed said that Ali came.’
   (Mohammad 2000:20)

I conclude that since there seems to be no clear information-structural difference between preverbal and postverbal subjects, subjects in TP cannot be generally excluded. Given this, an approach that is based on EPP movement of the subject seems to be valid.

5. Conclusion and Outlook

In this paper, I have argued for a new approach to the agreement asymmetry in Modern Standard Arabic. The data can be summarized into the five empirical observations repeated in (37).
Feature Deletion by Head Movement

(37)  
a. Number agreement is different from gender and person agreement.  
b. Number agreement is dependent on the linear order of subject and verb.  
c. Number agreement does not require surface adjacency between subject and verb.  
d. Agreement is dependent on the grammatical function of the agreement target: Only subjects can agree.  
e. Number agreement is dependent on whether the subject is a full nominal phrase or a pronoun.

In the present approach, these observations are derived by the following assumptions: Verbs in MSA undergo obligatory head movement to T and can in some cases check the EPP number feature on T. If they do, the subject remains in its base position in the vP and at the same time, the number feature on the verb (which has resulted from previous Agree with the subject) is deleted (because the verb undergoes Complex Head Feature Deletion). If the verb does not check the EPP feature, the subject can move to the preverbal position and the verb retains its number feature.

(37a) is due to number, but not gender or person being the EPP feature in MSA. (37b) follows because every preverbal subject has to move through Spec-TP, which means that the number feature on the verb is not deleted. (37c) naturally falls out from the approach because there is no process that requires linear adjacency. (37d) follows from the assumption that the object is not an active goal either for agreement or for EPP movement. Finally, (37e) is due to the assumption that pronouns, in contrast to full noun phrases, undergo obligatory EPP movement.

Even though, the new approach nicely captures the Arabic data, it makes an important prediction for the EPP property that still needs to be confirmed. The prediction is that there should be at least six different language types with respect to the EPP property. The typology predicted by the present account is summarized in (38).

(38)  
a. [D]: The verb is not able to check the EPP feature and any noun phrase might check the EPP property. There are no agreement asymmetries.  
b. [#]: The verb is able to check the EPP feature and agreement asymmetries in number should occur.
c. $[\pi]$: The verb is able to check the EPP feature and agreement asymmetries in person should occur.
d. $[\gamma]$: The verb is able to check the EPP feature and agreement asymmetries in gender should occur.
e. [case]: The verb is not able to check the EPP feature and only noun phrases with case might check the EPP property. There are no agreement asymmetries.
f. no EPP: The language has no EPP property.

While MSA is a clear case of (38b), it is unclear at this point if the typology is correct. As for (38a) and (38e), these types should be easy to find, as they are not predicted to show agreement asymmetries. The difference between the two types should be which arguments can ultimately check the EPP. Regarding (38c), this type should comprise languages without pro-drop that show an agreement asymmetry in person. If there is a connection between verbs being able to check the EPP and pro-drop, as Alexiadou & Anagnostopoulou (1998) suggest, this type might not exist for independent reasons. (38d), i.e. languages that show an asymmetry in gender, are clearly predicted to exist (see Samek-Lodovici 2002 for potential evidence). And finally, languages without any EPP property should exist. I leave this typology to future research.

Lastly, it is important to mention that the agreement pattern in Arabic is even more complicated than described in this paper once coordination (Aoun et al. 1994) and raising (Wurmbrand & Haddad 2014) is taken into account. It is well known that MSA is a language that has left conjunct agreement and is a backward raising language. I again leave these issues to future research.

**References**


Wurmbrand, Susi (2012). Agree(ment): Looking up or looking down?. Lecture given in Agreement Seminar, MIT.


Zeijlstra, Hedde (2012). There is only one way to agree. The Linguistic Review 29. 491–539.
The Third Construction and Strength of C: A Gradient Harmonic Grammar Approach

Gereon Müller

Abstract
This paper addresses the third construction in German, i.e., sentences that combine clause-internal movement from a control infinitive with extraposition of that infinitive. I argue that conflicting evidence regarding the degree of bi-/mono-clausality of the extraposed infinitive (as evidenced by Santorini & Kroch’s (1991) observation that long-distance scrambling is possible whereas wide scope of negation is not) is best captured by assuming that it qualifies as a CP with a C head that has more strength than the C of a preverbal restructuring infinitive embedded under a control verb, but less strength than the C of a non-restructuring infinitive (or a finite clause). This presupposes an approach to syntax in which a number of different strength assignments to a given type of category (like C) can be postulated, and can have a direct effect on the (non-) application of syntactic operations. I will show that a version of minimalist syntax incorporating the Phase Impenetrability Condition (Chomsky 2001, 2013) that is embedded in a Gradient Harmonic Grammar approach (Smolensky & Goldrick 2016) can account for the variable strength of C in a principled way.

1. A Paradox

The third construction in German involves a combination of scrambling or unstressed pronoun fronting from an infinitive embedded by a restructuring control verb on the one hand, and extraposition of that infinitive on the other hand; see Besten & Rutten (1989), Geilfuß (1991), Santorini & Kroch (1991), Wöllstein-Leisten (2001), Wurmbrand (2001, 2007), Reis & Sternefeld (2004), and Lee-Schoenfeld (2007), among others. A relevant example illustrating the transparency of the extraposed infinitive (Γ₁) for fronting of an unstressed pronoun (ihn₂) is given in (1a); (1b) is a minimally different example of such movement with the restructuring infinitive Γ₁ in situ.

(1) a. dass sie ihn₂ t₁ versucht [Γ₁ PRO t₂ zu küssen ]
    that sheₙᵒᵐ himₐᶜᶜ tries to kiss

---

Struktur REMOVAL, 419–446
Andrew Murphy (ed.)
LINGUISTISCHEN ARBEITSBERICHTE 94, Universität Leipzig 2019
Given that scrambling from a (finite or non-restructuring, non-finite) CP (unlike, say, wh-movement) is impossible in German (see Ross (1967)), the transparency of the extraposed infinitive for this movement operation is often taken to indicate that $\Gamma_1$ is not a CP in either (1a) or (1b). However, there is also conflicting evidence that supports a CP status of $\Gamma_1$ in the third construction. An indirect argument for this is that lower projections in the clausal spine (TPs, vPs, VPs) can otherwise never undergo extraposition in German (see Müller (2017)), with the Ersatzinfinitiv construction an exception that, under closer inspection, proves the rule (see Schmid (2005)). And a very clear and direct argument for a CP status of the extraposed infinitive $\Gamma_1$ is that scope of negation is strictly clause-bound in the third construction, in stark contrast to what is the case with non-extraposed restructuring infinitives. This observation goes back to Santorini & Kroch (1991). This is illustrated in (2a) (with only narrow scope of negation available in the third construction) vs. (2b) (where wide scope of negation is possible with standard restructuring infinitives).

Thus, a paradox arises: The availability of unstressed pronoun fronting and scrambling in the third construction in (1a) and (2a) suggests that $\Gamma_1$ is not a CP; and the unavailability of wide scope of negation in (2a) suggests that $\Gamma_1$ is a CP. It is the main goal of the present study to resolve this paradox in a principled way, by postulating that C is somewhat weaker in the third construction than in non-restructuring (and finite) contexts (so that scrambling and unstressed pronoun fronting from it are possible), but slightly stronger than in standard
restructuring contexts (so that it can undergo extraposition in the first place, and wide scope of negation becomes impossible).

2. Background: Strength in Grammar

It is an old idea in syntactic theory that a functional category X can be strong or weak (see, e.g., Rizzi (1986) and Koster (1986)). On this view, some syntactic operations may require a strong X, and others may require a weak X; yet others are compatible with any X. A more recent application of this general hypothesis involves complementizer-trace effects. Wh-Movement of a subject DP from a declarative clause embedded by a bridge verb is ungrammatical in English if it takes place across a C realized as that (see (3a)), but is possible if C is phonologically zero (see (3b)).

\[(3)\]
\[
\begin{align*}
\text{a. } & [\text{CP Who} \text{ do you think } [\text{CP t'} [\text{C } \emptyset ] t \text{ saw John } ]] ? \\
\text{b. } & [\text{CP Who} \text{ do you think } [\text{CP t'} [\text{C that} ] t \text{ saw John } ]] ?
\end{align*}
\]

To account for this, Chomsky (2013) suggests that “deletion of that [...] might leave only a weakened form of C” (my emphasis); this implies that the non-overt realization of C makes it possible to satisfy a constraint on movement that must be violated if the overt realization of C as that is chosen. Notwithstanding the issue of how such an idea is to be formally implemented, it can be noted that it raises a problem if a post-syntactic morphological realization of (at least) functional categories is adopted, as is the case in Distributed Morphology (see Halle & Marantz (1993)). On the one hand, a complementizer that cannot be assumed to be deleted in the syntax – that is in fact only inserted post-syntactically. On the other hand, if the difference between (3a) and (3b) only arises post-syntactically, how can it be the crucial factor for extraction?

There are many other areas where strength of functional categories has been invoked. A well-known example involves subject pro-drop; see, e.g., (4a) in Spanish vs. (4b) in English.

\[(4)\]
\[
\begin{align*}
\text{a. } & [\text{TP Hemos } [\text{vp pro trabajado todo el día } ]] \\
& \text{have-3.pl worked all the day} \\
\text{b. } & [\text{TP pro₁ Have } [\text{vp t₁ worked all day } ]] \\
& \text{a traditional assumption has been that the strength of T is decisive for allowing pro (see Rizzi (1986)): A strong T licenses pro, a weak T does not. More recently,}
\end{align*}
\]
Chomsky (2015) makes use of essentially the same distinction, when he claims that in some languages, “T is too weak to serve as a label”, and that “Italian T, with rich agreement, can label TP [...] for English, with weak agreement, it cannot”.

A further widespread assumption instantiating the very same idea of strength concerns V-to-T movement; see, e.g. (5a) in English vs. (5b) in French.

\[(5) \quad \begin{align*}
\text{a. John often kisses}_1 \text{ Mary} \\
\text{b. Jean embrasse}_1 \text{ souvent t}_1 \text{ Marie}
\end{align*} \]

In what is arguably still the standard approach (see Pollock (1989), Roberts (1993), Vikner (1997, 2001a,b), Holmberg & Platzack (1995), Rohrbacher (1999)), it is postulated that a strong T licenses V-to-T movement (as in French), whereas a weak T (as in English) does not.

In all these cases, it is typically assumed that strength correlates in one way or another with the extent of morphological realization (with zero realization as the limiting case). However, as observed by Bobaljik (2002), all these analyses face the problem of being incompatible with post-syntactic morphology that I have illustrated for complementizer-trace effects above. For instance, as regards V-to-T movement, properties of the morphological inventory cannot be held responsible for whether such movement can apply in the syntax or not if inflectional morphology is post-syntactic.

I conclude from all this, first, that there is some evidence that functional categories can have different degrees of syntactic strength; and second, that such strength cannot be determined on the basis of morphological realization if this latter information is not yet present in the syntax. Given this state of affairs, it looks as though two ways out suggest themselves naturally. One is to abandon the idea of post-syntactic morphological realization. The other one is to conclude that strength is an abstract inherent property of functional categories that (i) determines whether or not syntactic operations can apply, and that (ii) also determines post-syntactic morphological realization. I will pursue this latter approach in what follows. From this perspective, the task at hand is to show how syntactic building blocks (in the sense of operations, constraints, or rules) can be sensitive to different degrees of strength. Gradient Harmonic Grammar (see Smolensky & Goldrick (2016)) is a new grammatical theory designed to implement effects of this type. The particular minimalist version that I will adopt is laid out in the next section.
I would like to contend that Gradient Harmonic Grammar, which is introduced in Smolensky & Goldrick (2016) mainly on the basis of phonology, offers a new perspective on how to derive three different types of asymmetries as they can be observed with long-distance dependencies in the world’s languages: first, asymmetries between movement types (e.g., movement types that are clause-bound vs. movement types that can apply long-distance); second, asymmetries between types of moved items (e.g., subjects vs. objects, or arguments vs. adjuncts); and third (and most importantly in the present context), asymmetries between types of local domain (e.g., VP typically permits extraction from it, CP often does not – and certain types of CPs will be shown to be different from certain other types of CPs, too). More specifically, the version of Gradient Harmonic Grammar that will be relevant in what follows combines properties of three subtheories: (i) Harmonic Grammar; (ii) Gradient Symbolic Representations; and (iii) Harmonic Serialism. I will address these in turn.

3.1. Harmonic Grammar

Harmonic Grammar (see Smolensky & Legendre (2006), Pater (2016)) is a version of optimality theory (see Prince & Smolensky (1993)) that abandons the strict domination property (according to which no number of violations of lower-ranked constraints can outweigh a single violation of a higher-ranked constraint) and replaces harmony evaluation by constraint ranking with harmony evaluation based on weight assignment to constraints. This makes it possible to derive some (but not all) kinds of cumulative effects in syntax. The central notion of harmony is defined in (6) (see Pater (2009)).

\[
\text{Harmony:} \quad K \\
\mathcal{H} = \sum_{k=1}^{K} s_k w_k \\
\text{where } w_k = \text{weight of a constraint; } s_k = \text{violation score of a candidate}
\]

Thus, the weight of a constraint is multiplied with the violation score of a candidate for that constraint, and all the resulting numbers are added up, thereby determining the harmony score of a candidate. For present purposes,
we can assume that constraints assign negative scores throughout (e.g., −1 if the candidate violates a constraint once), and that constraint weights are always nonnegative (e.g., 2, or 3). Thus, if a candidate violates constraint A (with weight 2.0) once (−1) and constraint B (with weight 3.0) twice (−2), the harmony score of the candidate would be −8 if there were no further constraints in the grammar. Finally, an output qualifies as optimal if it is the candidate with maximal harmony in its candidate set; i.e., if it has the value closest to zero (or the lowest penalty).

3.2. Gradient Harmonic Grammar

Against this background, the main innovation of Gradient Harmonic Grammar is that Smolensky & Goldrick postulate that it is not just the constraints that are assigned weights. Rather, symbols in linguistic representations are also assigned weights; i.e., they are not categorical either. The weights in question are encoded by assigning some real number between 0 and 1. This way, the concept of varying strength of syntactic categories can be formally implemented in the grammar. For example, suppose that some category X can have three different kinds of weights in a given grammar: X:[0.4], X:[0.7], and X:[1.0]. Suppose further that X violates some constraint Γ that is associated with a weight of 2, and that it does so once (−1). Then, the first X will give rise to a −0.4 violation of Γ, yielding a (partial) harmony score of −0.8; the second X induces a −0.7 violation of Γ, which results in a (partial) harmony score of −1.4; and the third X triggers a −1.0 violation of Γ, which produces a (partial) harmony score of −2.0. Of course, there will be constraints counter-acting Γ, which may then imply that the violation of Γ incurred by X is tolerable in an optimal candidate if X has a weight of [0.4] but not tolerable in an optimal candidate if X has a weight of [1.0].

So far, most of the work on Gradient Harmonic Grammar has been in phonology; but cf. Smolensky (2017), Lee (2018), and Müller (2019) for applications in syntax.¹

¹As it turns out, there is a fairly obvious predecessor of Gradient Harmonic Grammar in syntax (not mentioned in Smolensky & Goldrick (2016)), viz., Squishy Grammar, which was developed by Ross (1973a,b, 1975). Ross argues that there is constituent class membership to a degree, and presupposes that instead of standard category symbols like [X], there are weighted category symbols like [αX] (where α ranges over the real numbers in [0,1]). Rules, filters, and other syntactic building blocks are given upper and lower threshold values of α between which
3.3. Harmonic Serialism

In addition to Harmonic Grammar and Gradient Representations, Harmonic Serialism is a third important ingredient of the present approach. Harmonic serialism is a strictly derivational version of optimality theory. (7) illustrates how it works (see McCarthy (2008) and Heck & Müller (2013), for phonology and syntax, respectively).

(7) **Harmonic serialism:**

a. Given some input $I_i$, the candidate set $CS_i = \{O_{i1}, O_{i2}, \ldots, O_{in}\}$ is generated by applying at most one operation to $I_i$.

b. The output $O_{ij}$ with the best constraint profile is selected as optimal.

c. $O_{ij}$ forms the input $I_{ij}$ for the next generation step producing a new candidate set $CS_j = \{O_{ij1}, O_{ij2}, \ldots, O_{ijn}\}$.

d. The output $O_{ijk}$ with the best constraint profile is selected as optimal.

e. Candidate set generation stops (i.e., the derivation converges) when the output of an optimization procedure is identical to the input (i.e., when the constraint profile cannot be improved anymore).

Harmonic Serialism was already identified as a possible alternative to standard parallel optimization in Prince & Smolensky (1993). However, it has been pursued in depth only over the last decade or so (see, e.g., McCarthy (2008, 2016), Torres-Tamarit (2016), Elfner (2016) for phonology; Caballero & Inkelas (2013), Müller (2018) for morphology; and Heck & Müller (2013), Georgi (2012), Assmann et al. (2015), Murphy (2017) for syntax). As shown in McCarthy & Pater (2016) and Murphy (2017), the combination of Harmonic Grammar and Harmonic Serialism is a natural one. As far as syntax is concerned, Harmonic Serialism can be viewed as a version of minimalist approaches employing sequential bottom-up structure-building (see Chomsky (1995, 2001, 2014)) that incorporates optimization procedures (like Merge over Move). The main
empirical arguments here concern phenomena which provide evidence that (i) there is syntactic optimization, but (ii) this optimization can only take into account information that is accessible in an extremely local syntactic domain (from the current root down to the closest phase edge), and it can only distinguish between a finite (and small) number of operations that can in principle be carried out at any given step. In the present context, a Harmonic Serialism perspective ensures that the scores of constraint violations resulting from combining the weights of the constraints and the weights assigned to the linguistic expressions are consistently fairly small and manageable, and are forgotten again once the derivation moves on to the next cycle.

Taken together, the three subtheories can be referred to as Serial Gradient Harmonic Grammar.

4. Proposal

4.1. Constraints and Weights

In the analysis of extraction from CP to be developed below, three constraints turn out to be important. First, there is the Phase Impenetrability Condition (PIC; Chomsky (2001, 2008, 2013)), which demands that all operations involving some item $\alpha_i$ in a phase and some other item outside the phase requires $\alpha_i$ to be in the edge (specifier or head) domain of the phase. In (8), the PIC is formulated as a constraint on heads.

(8) \textit{Phase Impenetrability Condition} (PIC):

For all heads $Y$: $^*Y$ that c-commands $\alpha_i$ of a dependency $\delta$ but does not m-command $\alpha_{i-1}$ of $\delta$.

The PIC in (8) is a strengthened version of Chomsky's original PIC since it acknowledges a potential barrier status of all XPs: Every phrase is a phase. In this respect, it resembles concepts proposed by Riemsdijk (1978), Koster (1978, 1987), Sportiche (1989), and Legendre et al. (2006), among others.

For movement steps leaving a phase, the PIC in (8) thus demands that extraction takes place via the specifier of the phase head. Crucially, I assume that the PIC is an inviolable constraint of the GEN component of the grammar (see Prince & Smolensky (1993)).

\footnote{This follows without further ado if one follows Chomsky in assuming that the PIC is derivable...}
In contrast, the remaining two constraints are violable, and are assigned weights. These are the Merge Condition and the Anti-Locality Condition. The Merge Condition (MC) can be formulated as in (9) (see Chomsky (1995, 2001); and Heck & Müller (2013) for the particular [●F●] notation for features triggering structure-building.)

(9) **Merge Condition** (MC):
For all features [●F●] and XPs with a matching [F]: [●F●] triggers Merge of XP.

(9) presupposes that each head is associated with a set of structure-building features [●F●] which are discharged by individual Merge operations one at a time.3 MC is formulated here as a constraint on two items: structure-building features on the one hand, and XPs with a matching feature on the other. This makes it possible to determine violations of the constraint (with its own weight) relative to the weights of these items (i.e., the attracting feature and the moved item).

The second violable constraint is the Anti-Locality Condition (see Bošković (1997), Abels (2003), Grohmann (2003a, b, 2011), Pesetsky (2016), and Erlewine (2016) for different implementations of this general idea), which is formulated in (10) in a maximally strict way that is made possible by assuming violability.

(10) **Anti-Locality Condition** (AL):
For all heads Y: *Y that c-commands αi of a dependency δ and m-commands αi−1 of δ.

As regards links of movement dependencies, (10) is violated by all heads which c-command a (base or derived) position from which movement takes place and also m-command the landing site of this movement. The prototypical scenarios for this are (i) that movement has taken place from the specifier of some phrase ZP, across ZP’s sister Y, to a specifier of Y, as in [YP αi−1 [Y Y · ZP αi [Z · ... ]]]; or (ii) that movement has taken place from the complement of Y to Y’s specifier, as in [YP αi−1 [Y Y αi ]].4 Given the PIC in (8) as a constraint from cyclic spell-out of the phase head’s complement after completion of the phase; under this assumption, material that is not in the edge domain is literally irrevocably gone after spell-out.

3Alternatively, these features may be assumed to show up as members of a list (rather than a set); while ultimately important, this issue is negligible in the present context.

4Strictly speaking, a third scenario might involve the configuration [YP Y · ZP αi−1 [Z Y · Z ... αi]}.
on all phrase heads, *all* movement violates AL (movement originates either in the complement position of some head Y, or in the specifier position of Y’s complement). Thus, whereas MC is a trigger for movement, AL acts as a potential blocker: If AL cannot be violated in an optimal candidate, the PIC will subsequently ensure that movement cannot take place. Note that unlike a general economy constraint blocking movement (e.g., *Trace, as in Grimshaw (1997), Legendre et al. (2006)), AL has different effects depending on the nature of the head crossed in the course of movement. A head Y with a larger weight (i.e., more strength) will give rise to a more severe violation of AL than a head Y with a lower weight (i.e., less strength).

This approach depends on the availability of edge features that may trigger intermediate movement steps via MC. Following Abels (2012), I assume that intermediate movement steps are brought about by duplicates of criterial features, which can freely be assigned to any head Y. For instance, a feature like [●wh●] that is an inherent property of interrogative C in German can show up on all heads (C, T, V, v, etc.) intervening between the base position and the ultimate landing site SpecC_{wh}.

Summarizing so far, it emerges that weight (i.e., relative strength) plays a role for three different kinds of items that are subject to the constraints MC and AL. First, some Y heads give rise to stronger violations of AL than other Y heads if movement takes place across them. This derives asymmetries between types of local domain. For instance, VP typically permits extraction from it, and vP often does so; but CP in many cases does not. As will be shown below, this also accounts for the difference between restructuring and non-restructuring infinitival C in German, where the former but not the latter permits scrambling and unstressed pronoun fronting to the matrix domain. For concreteness, I will assume the following weights for Y heads involved in AL violations in German:

(11) **Strength of Y:**

a. V: [0.45]
b. C_{[+restr,−fin]}: [0.6]
c. C_{[−restr,−fin]}: [0.8]
d. C_{[−wh,+fin]}: [0.9]
e. C_{[+wh,+fin]}: [1.0]

...]]], where Y also c-commands α_i and m-commands α_{i−1}. However, it is not clear whether this scenario needs to be excluded by modifying AL (e.g., by adopting minimal c-command), given that α_i will never be accessible to Y because of the inviolable PIC (α_i will fail to be c-commanded by Y if it is not even part of the representation anymore at this point; see footnote 2 above).
Thus, V does not bear a lot of weight; consequently, an AL violation induced by movement to SpecV is usually tolerable in German. Similar considerations apply for v and T (where the weights are not shown here). According to (11), C has more weight. More generally, the underlying hypothesis is that the weight increases from bottom to top with functional heads in the clausal spine. Furthermore, all control infinitives in German are assumed to have CP status throughout. Abstracting away from the third construction for now, the infinitival C head comes in two varieties, a non-restructuring version that has nearly the same weight as finite declarative C ([0.8]), and a restructuring version that has less weight ([0.6]). It is a property of restructuring control predicates that they can select either version of non-finite C (whereas other control predicates can only select the non-restructuring version).

Second, some movement-related features [●F●] give rise to stronger violations of MC (i.e., are stronger triggers of movement) than other movement-related features. This derives asymmetries between movement types. For instance, wh-movement can leave a finite CP in German whereas scrambling cannot do so. Concrete weights assigned to structure-building features that trigger movement in German include those in (12); [●wh●] is involved in wh-movement, and [●scr●] is involved in scrambling and unstressed pronoun fronting. Again, the increase in strength corresponds to the relative position of the head(s) bearing the feature in the tree: The landing site of wh-movement is SpecC, the landing site of scrambling is Specv or SpecV.

---

5See, however, Müller (2019), where I argue that the ban on splitting up particularly opaque kinds of idioms by certain kinds of movement can be traced back to an AL violation with movement to SpecV that is fatal in the presence of a moved item with extremely little strength (giving rise to a less severe MC violation if movement does not take place).

6Also, a finite interrogative C has more weight than a finite declarative C ([1.0] vs. [0.9]); this ultimately accounts for wh-islands; see Müller (2019).

7There are in fact several differences between scrambling of non-pronominal items, as in (2b), and unstressed pronoun fronting, as in (1b). Still, to simplify matters I pretend here that [●scr●] covers both movements; a more detailed analysis would postulate two separate features with sufficiently similar weights.

8Topicalization can leave wh-islands in German with objects (but not subjects), whereas wh-movement (or scrambling) cannot do so. In Müller (2019), this is modelled by assuming that the feature [●top●], which triggers topicalization, has more weight than the features triggering wh-movement and scrambling (viz., [0.65] vs. [0.5], [0.2]).
(12) *Strength of \( \bullet F \bullet \): 
   a. \( \bullet \text{scr} \bullet \): [0.2]
   b. \( \bullet \text{wh} \bullet \): [0.5]

Third, some XPs give rise to stronger violations of MC than other XPs if they do not undergo movement. This accounts for asymmetries between moved items (e.g., unmoved objects may induce stronger violations of MC than unmoved subjects, and thus make MC violable less easily in optimal outputs). For German, I assume that an object DP has a weight of [0.9], whereas a subject DP only has a weight of [0.8]. However, I will be exclusively concerned with object DPs in what follows.\(^9\)

With these assumptions in place, let me next illustrate the mechanics of the resulting system on the basis of some data involving extraction from different domains, and by different movement types.

4.2. Two Extraction Asymmetries in German

4.2.1. Asymmetries between Types of Local Domain

Scrambling can target SpecV in the German, either as a final landing site, or as an intermediate escape hatch for further movement to Specv required by the PIC; see (13a) and (13b), respectively.

(13) a. dass sie \( [\text{VP} \left[ \text{DP}_2 \text{ das Buch } \right] \left[V \left[ \text{V} \right] \text{ dem Karl } \right] \left[V' \left[ \text{DP}_1 \right] \text{ the Karl}_{\text{dat}} \right] \right] \text{ gegeben hat} \] \]
given has

b. dass \( [\text{VP} \left[ \text{DP}_2 \text{ das Buch } \right] \left[V \left[ \text{V} \right] \text{ keiner } \right] \left[V' \left[ \text{VP} t'_2 \left[V' \left[ \text{V} \right] \text{ t}_2 \right] \text{ read } \right] \left[V \left[ \text{V} \right] \text{ v } \right] \right] \text{ no-one}_{\text{nom}} \]

read has

However, as noted above, scrambling is clause-bound in German (see Ross (1967)): A finite CP cannot be crossed. From the present, PIC-based perspective,

---

\(^9\)See Müller (2019) for discussion of asymmetries between types of moved items.
this can be taken to indicate that SpecC cannot be targeted as an intermediate landing site by this movement operation; see (14).

(14) *dass sie [DP₂ das Buch] gesagt hat [CP t₂’ [C’ dass ] [TP t₂ sie
gesehen hat ]]
read has

This asymmetry between VP and CP with respect to scrambling follows from the current assumptions about weight assignments. On the one hand, given that what is moved is an object DP ([0.9]), and given that the feature responsible for the (intermediate or final) movement step is [scr●] ([0.2], a relatively weak trigger), there will be a −1.1 violation of MC in both environments if movement does not take place. Assuming MC itself to have a weight of 2.0, this produces a harmony score of −2.2. On the other hand, if movement takes place, an AL violation will be generated. Suppose that the intrinsic weight of AL is 3.0. Then, since V, by assumption, has a weight of [0.45] (see (11)), movement of any item to SpecV gives rise to a −0.45 violation of AL, and thus (abstracting away from other constraint violations that are irrelevant in the present context) to a harmony score of −1.35. Consequently, the output candidate O₂ employing a local scrambling step to SpecV emerges as optimal, and the output candidate O₁ which fails to carry out movement is suboptimal. This is illustrated by the tableau in (15) (where \( \mathcal{H} \) stands for the overall harmony score of a candidate).

(15) **Object scrambling via VP:**

<table>
<thead>
<tr>
<th>I: [VP ... DP_{obj:[0.9]} V_{[0.45]}[scr●]:[0.2]]</th>
<th>MC w = 2.0</th>
<th>AL w = 3.0</th>
<th>( \mathcal{H} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁: [VP ... DP_{obj:[0.9]} V_{[0.45]}[scr●]:[0.2]]</td>
<td>−1.1</td>
<td>−2.2</td>
<td></td>
</tr>
<tr>
<td>( \text{andr} ) O₂: [VP DP_{obj:[0.9]} V_t_{obj} V_{[0.45]}[scr●]:[0.2]]</td>
<td>−0.45</td>
<td>−1.35</td>
<td></td>
</tr>
</tbody>
</table>

In contrast, if object scrambling wants to leave a finite declarative CP, intermediate movement to SpecC, across an intervening C with weight [0.9], produces a much more severe violation of AL: This time there is a −0.9 violation of AL, which ceteris paribus leads to a harmony score of −2.7. The candidate without movement (in the presence of [scr●] and an object DP) has a harmony score

---

\(^{10}\) In contrast, there would be nothing wrong as such with the subsequent movement step to matrix SpecV. Such a step is often excluded by some specific constraints against improper movement (see Müller (2014), Keine (2016) for recent overviews), but in the present approach based on variable weights, such constraints can be dispensed with; cf. 4.2.2 below.
of −2.2, exactly as before; but this MC violation now emerges as optimal, and intermediate scrambling to SpecC is therefore blocked. Ultimately, the PIC then ensures that long-distance scrambling cannot take place from the lower SpecT position in the embedded clause that we can assume to have been reached by prior intermediate scrambling-movement. This competition is shown in (16).

(16) Object scrambling via finite declarative CP:

<table>
<thead>
<tr>
<th>Example</th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\exp O_1$: $[C \stackrel{\text{scram}}{\rightarrow} C]<em>{[0,9]} [\text{DP}</em>{\text{obj}}]<em>{[0,9]} [TP \text{DP}</em>{\text{obj}}]<em>{[0,9]} [TP \text{DP}</em>{\text{obj}}]_{[0,9]}$</td>
<td>$w = 2.0$</td>
<td>$w = 3.0$</td>
<td>$-1.1$</td>
</tr>
<tr>
<td>$O_2$: $[C \stackrel{\text{scram}}{\rightarrow} C]<em>{[0,9]} [\text{DP}</em>{\text{obj}}]<em>{[0,9]} [C^{\prime} C^{\prime}]</em>{[0,9]} [\text{DP}<em>{\text{obj}}]</em>{[0,9]} [TP \text{DP}<em>{\text{obj}}]</em>{[0,9]}$</td>
<td>$w = 3.0$</td>
<td>$w = 3.0$</td>
<td>$-2.2$</td>
</tr>
</tbody>
</table>

Next, if different kinds of Cs ([±finite], [±restructuring], [±wh], etc.) can have different weights, it can be derived that one and the same movement type (e.g., scrambling) may leave CPs with a weak C head (restructuring infinitives) but not CP with a stronger C head (finite clauses or non-restructuring infinitives). A relevant pair of examples illustrating the lexically governed restructuring effect with control infinitives in German is given in (17).

(17) a. dass $[\text{DP}_{\text{obj}} \text{ das Buch }] \text{ keiner } [\text{CP } t_2 \text{ zu lesen }]$ versucht hat
   read tried has

b. *dass $[\text{DP}_{\text{obj}} \text{ das Buch }] \text{ keiner } [\text{CP } t_2 \text{ zu lesen }]$ abgelehnt hat
   read rejected has

By assumption, restructuring C in (17a) has a weight of [0.6], whereas non-restructuring C in (17b) has a weight of [0.8]. Consequently, non-restructuring infinitival C blocks scrambling from it in basically the same way as finite declarative C in (16) (with a suboptimal harmony score of −2.4 if movement applies, violating AL); but with restructuring C, the AL violation incurred by movement is not so severe anymore (the overall harmony score is −1.8), and successfully blocks the candidate that fails to carry out movement (in violation of MC, with a harmony score of −2.2); see (18)."
The present approach makes it possible to uniformly assume a CP status of restructuring infinitives embedded under control verbs. This is arguably conceptually attractive in view of the implicational generalization that there is no control verb that permits restructuring which would not also permit a non-restructuring clausal complement. In approaches where the two complement types have a different categorial status (e.g., vP vs. CP; see Haider (1993, 2010), Wurmbrand (2001)), this state of affairs is purely accidental; in the present approach, it only requires the assumption that there is an unmarked strength of infinitival C items (viz., [0.8]) which can optionally be reduced (and which then is tolerated only by a subset of control predicates). However, there is also empirical evidence for CP in restructuring infinitives embedded by control verbs; see Baker (1988), Sternefeld (1990), Müller & Sternefeld (1995), Sabel (1996), Koopman & Szabolcsi (2000), and Müller (2017). For instance, one argument from the last-mentioned study relies on the generalization that unstressed pronoun fronting to the left edge of vP (which is obligatory in German) must be licensed by a higher C phase head. And whereas such movement is impossible in structures clearly lacking a CP (verb-auxiliary combinations as in (19a), raising environments as in (19b)), it is possible in restructuring contexts embedded by control verbs (as in (19c)).

\((18)\) **Object scrambling via restructuring infinitive CP:**

\[
\begin{array}{ccc}
\text{MC} & \text{AL} & \mathcal{H} \\
\text{w = 2.0} & \text{w = 3.0} \\
\end{array}
\]

\[
\begin{array}{c}
{\text{O}_1: [CP \ C_{[0.6]}][\text{sc}][0.2] \ [TP \ \text{DP}_{\text{obj}}][0.9] \ [T' \ ... \ T]]} \\
{\text{w = 2.0}} \\
{-1.1} \\
{-2.2} \\
\end{array}
\]

\[
\begin{array}{c}
{\text{O}_2: [CP \ \text{DP}_{\text{obj}}][0.9] \ [C' \ C_{[0.6]}][\text{sc}][0.2] \ [TP \ t_{\text{obj}}][T' \ ... \ T]]} \\
{\text{w = 3.0}} \\
{-0.6} \\
{-1.8} \\
\end{array}
\]

The present approach makes it possible to uniformly assume a CP status of restructuring infinitives embedded under control verbs. This is arguably conceptually attractive in view of the implicational generalization that there is no control verb that permits restructuring which would not also permit a non-restructuring clausal complement. In approaches where the two complement types have a different categorial status (e.g., vP vs. CP; see Haider (1993, 2010), Wurmbrand (2001)), this state of affairs is purely accidental; in the present approach, it only requires the assumption that there is an unmarked strength of infinitival C items (viz., [0.8]) which can optionally be reduced (and which then is tolerated only by a subset of control predicates). However, there is also empirical evidence for CP in restructuring infinitives embedded by control verbs; see Baker (1988), Sternefeld (1990), Müller & Sternefeld (1995), Sabel (1996), Koopman & Szabolcsi (2000), and Müller (2017). For instance, one argument from the last-mentioned study relies on the generalization that unstressed pronoun fronting to the left edge of vP (which is obligatory in German) must be licensed by a higher C phase head. And whereas such movement is impossible in structures clearly lacking a CP (verb-auxiliary combinations as in (19a), raising environments as in (19b)), it is possible in restructuring contexts embedded by control verbs (as in (19c)).

\((19)\) a. *dass sie mir\textsubscript{1} schon letzte Woche [v\textsubscript{P} es\textsubscript{2} t\textsubscript{1} t\textsubscript{2} gegeben ]

that she\textsubscript{nom} me\textsubscript{dat} already last week it\textsubscript{acc} given

hat

has

\[\text{ predicates permit restructuring, and which ones do not. For some speakers, (17b) may be possible, but this does not affect the analysis: ablehnen just tolerates a weaker C here.}\]

\[\text{In these examples, mir\textsubscript{1} undergoes fronting to the matrix domain, thereby indicating transparency of the complement of the higher verb; es\textsubscript{1} is fronted string-vacuously in the complement.}\]
From a slightly more general perspective, under present assumptions there can be a lot of variation as far as the transparency of projections in the clausal spine for extraction is concerned (depending on the weights assigned to the heads in the extended projection of V). However, the variation is principled in the sense that it must obey an implicational universal: If an XP \( \alpha \) can undergo \( \Sigma \)-movement across a Y head \( \delta_1 \), and \( \delta_1 \) has more weight than another Y head \( \delta_2 \), then \( \alpha \) can ceteris paribus also undergo \( \Sigma \)-movement across \( \delta_2 \). Given the ancillary assumption that weight increases from bottom to top in the clausal spine, it is then predicted that if a given movement type affecting some particular item can take place across CP, it can also take place across TP; if it can leave TP, it can ceteris paribus leave vP; and similarly for vP and VP. I take this prediction to be correct.

4.2.2. Asymmetries between Movement Types

If a given head Y blocks a movement type triggered by a (intermediate or final) feature \( \Sigma_1 \) because the AL violation incurred by movement has a lower harmony score than the relatively weak MC violation incurred by not moving the item, this does not necessarily mean that Y will also block another movement type triggered by a different feature \( \Sigma_2 \): Not satisfying \( \Sigma_2 \)'s demand by leaving the item in place may give rise to a much more severe violation of MC if \( \Sigma_2 \) has greater strength than \( \Sigma_1 \), and this can then make the AL violation optimal. Such a situation obtains with wh-movement (triggered by [\( \bullet \text{wh} \bullet \)]) vs. scrambling (triggered by [\( \bullet \text{scr} \bullet \)]). Recall from (12) that the former feature is associated with a weight of [0.5] in German, and the latter with a weight of [0.2]. And indeed, for most speakers of German, wh-movement can leave a finite declarative CP where scrambling cannot (for reasons discussed in the previous subsection); see (20a) (with wh-movement) vs. (20b) (= (14)).
As shown in (21), wh-movement of an object DP via VP (as in O₃) is entirely unproblematic; as was the case with scrambling (see (15)), an AL violation is tolerable because the overall harmony score is closer to zero than that of a candidate that does not carry out movement in violation of MC (cf. O₁).

(21) **Object wh-movement via VP:**

<table>
<thead>
<tr>
<th>I: [VP ... DP_{obj}[:0.9] V[:0.45],{\textit{\textbullet}}[:0.5]]</th>
<th>MC</th>
<th>AL</th>
<th>ℋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁: [VP ... DP_{obj}[:0.9] V[:0.45],{\textit{\textbullet}}[:0.5]]</td>
<td>-1.4</td>
<td>-2.8</td>
<td></td>
</tr>
<tr>
<td>O₂: [VP V′ ... t_{obj} V[:0.45],{\textit{\textbullet}}[:0.5]]</td>
<td>-0.45</td>
<td>-1.35</td>
<td></td>
</tr>
</tbody>
</table>

However, things are different when it comes to extraction via CP. As shown in (22), the output candidate that moves the object DP to SpecC (i.e., O₂) now still has a better constraint profile than the candidate that does without such movement (i.e., O₁): The reason is that C’s [\textit{\textbullet}] feature in (22) (with a weight of [0.5]) ceteris paribus gives rise to a much stronger violation of MC if movement does not take place than C’s [\textit{\textbullet}] feature in (16) (with a weight of [0.2]) does.

(22) **Object wh-movement via finite declarative CP:**

<table>
<thead>
<tr>
<th>I: [CP C[:0.9],{\textit{\textbullet}}[:0.5] [TP DP_{obj}[:0.9] [T′ ... T]]]</th>
<th>MC</th>
<th>AL</th>
<th>ℋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁: [CP C[:0.9],{\textit{\textbullet}}[:0.5] [TP DP_{obj}[:0.9] [T′ ... T]]]</td>
<td>-1.4</td>
<td>-2.8</td>
<td></td>
</tr>
<tr>
<td>O₂: [CP DP_{obj}[:0.9] C′ C[:0.9],{\textit{\textbullet}}[:0.5] [TP t_{obj} [T′ ... T]]]</td>
<td>-0.9</td>
<td>-2.7</td>
<td></td>
</tr>
</tbody>
</table>

Again, the approach predicts a lot of variation, but as before, such variation is principled: A second implicational universal can be derived which states that if an XP α can undergo Σ₁-movement across a Y head δ, and Σ₁ has less weight than another movement type Σ₂, then α can also undergo Σ₂-movement across δ, other things being equal. And, also as before, the relative weight of the
features that bring about movement via MC is not arbitrary but corresponds to the relative position of the heads bearing the features in the tree.\textsuperscript{13}

Needless to say, the approach to extraction in German sketched so far needs to be extended in many directions, and with a broader empirical coverage, it must be subject to many further ramifications. However, I will leave it at that here. Instead, I will now turn to the main goal of the present paper, which is to solve the paradox with the third construction outlined in section 1 above.

5. The Third Construction

In many respects, the extraposed infinitival complement in the third construction in German behaves like the non-extraposed restructuring infinitive counterpart analyzed in subsection 4.2.1 above. First, as noted in section 1, the extraposed infinitival complement is transparent for scrambling and unstressed pronoun fronting if it would be transparent for these movement types in the pre-verbal base position – i.e., if the matrix predicate licenses restructuring.\textsuperscript{14} Some relevant examples that document this are given in (23a), (23b) (= (1a)), and (23c).

(23) a. dass das Buch\textsubscript{2} keiner \textsubscript{t1} versucht hat [CP\textsubscript{1} PRO t\textsubscript{2} zu lesen ]
that the book\textsubscript{acc} no-one\textsubscript{nom} tried has to read
b. dass sie \textsubscript{nom} ihn\textsubscript{2} \textsubscript{t1} versucht [CP\textsubscript{1} PRO t\textsubscript{2} zu küsse n ]
that she\textsubscript{nom} him\textsubscript{acc} tries to kiss
c. dass es\textsubscript{2} Fritz \textsubscript{nom} ihr \textsubscript{t1} empfohlen hat [CP\textsubscript{1} PRO im Fritz\textsubscript{acc} her\textsubscript{dat} recommended has on the Zug t\textsubscript{2} zu lesen ]
train to read

As with restructuring infinitives in situ, this might initially be taken to suggest that extraposed restructuring infinitives in the third construction do not have CP status. But as before, there are conceptual and empirical arguments for the presence of a CP shell here. For instance, the third construction provides a

\textsuperscript{13}Concerning variation, it is also worth noting that by slightly increasing the weight of finite declarative C, wh-movement from CP will become impossible. As a matter of fact, such a scenario comes close to the situation in certain Northern varieties of German, which do not easily permit wh-movement from finite declarative clauses headed by a C with that.

\textsuperscript{14}Of course, this holds true virtually by definition – movement from an extraposed restructuring infinitive is the constitutive property of the third construction.
C-licensed landing site (at the left edge of the embedded vP) for unstressed pronoun fronting, just like restructuring infinitives in situ do (cf. (19)); see (24) (where fronting of mir\(_3\) into the matrix domain indicates transparency of the extraposed infinitive, and string-vacuous movement of es\(_2\) indicates the presence of C as a licensor for unstressed pronoun fronting in the infinitive).

\[
\text{(24) dass sie mir\(_1\) schon letzte Woche versucht hat [CP es\(_2\) PRO t\(_1\) that she\(_{nom}\) me\(_{dat}\) already last week tried has it\(_{acc}\) t\(_2\) zu geben ] to give}
\]

However, there are also differences between standard (i.e., pre-verbal) restructuring control infinitives and the third construction. In particular, there is Santorini & Kroch’s (1991) observation that a negation showing up in the extraposed infinitive can never take wide scope; cf. (2), repeated here in (25) (with CP\(_1\) replacing the original \(\Gamma\)\(_1\) as the label of the infinitive, and some other information added).

\[
\text{(25) dass ich seinen neuesten Roman\(_2\) t\(_1\) beschlossen habe [CP\(_1\) PRO t\(_2\) that I his newest novel\(_{acc}\) decided have nicht zu lesen ] not to read}
\]

(only narrow scope)

Thus, we end up with the paradox that extraposed infinitives in restructuring contexts are transparent for scrambling but not transparent for scope of sentential negation. This paradox arguably poses a non-trivial problem for standard approaches.\(^{15}\) From the present perspective, a simple solution suggests itself: The C head of the extraposed infinitive in the third construction has more

\(^{15}\)One might think that directionality could be the relevant factor determining obligatorily narrow scope of negation in the third construction, especially since there is some evidence that pre- vs. postverbal position can play a role for scope assignment in German when focus particles are involved (see Bayer (1996)). However, for the case at hand, this seems unlikely. As shown in (i), a universal quantifier embedded in an extraposed PP can easily take wide scope (as a matter of fact, wide scope of the universal quantifier produces the only reading that is compatible with world knowledge).

\[
\text{(i) dass der Polizist eine Bombe t\(_1\) gefunden hat [PP\(_1\) hinter jedem Haus ] that the policeman\(_{nom}\) a bomb\(_{acc}\) found has behind every house}
\]
strength than the C head of a restructuring infinitive in situ but less strength than the C head of a non-restructuring infinitive (or a finite C). More specifically, I would like to suggest that the C head of an extraposed infinitive in the third construction has a weight of [0.7] (as opposed to [0.8] for a non-restructuring C and [0.6] for a regular restructuring C; cf. (11)).

A first consequence of this weight assignment to non-finite C in the third construction is that it patterns with restructuring C as far as scrambling or unstressed pronoun fronting to the matrix domain is concerned, rather than with non-restructuring (or finite) C. Thus, the outcome of the competition in (26) parallels that of (18) (where the optimal output candidate violates AL by applying the intermediate movement step to SpecC required by the PIC), and not that of (16) (where the optimal output candidate violates MC by not carrying out movement); see (26).

(26) Object scrambling via extraposed infinitive CP in the third construction:

| T
| I: [CP C[0.7][scr][0.2] [TP DPobj[0.9] [T' ... T]]] | MC w = 2.0 | AL w = 3.0 | H |
|---|---|---|---|---|
| O₁: [CP C[0.7][scr][0.2] [TP DPobj[0.9] [T' ... T]]] | −1.1 | −2.2 |
| O₂: [CP DPobj[0.9] [C' C[0.7][scr][0.2] [TP tobj [T' ... T]]]] | −0.7 | −2.1 |

The AL violation incurred by DP movement to SpecC in O₂ is more severe in (26) (−2.1) than it was in the case of restructuring infinitives in situ in (18) (−1.8), but the harmony score is still better than the harmony score of the competing output O₁ where movement fails to apply, and MC (with weight 2.0) gets a combined −1.1 violation incurred by the [scr] feature ([0.2]) and the object DP ([0.9]), yielding a fatal −2.2 overall.

On the other hand, the larger weight of [0.7] for this type of non-finite C can be held responsible for differences to standard restructuring infinitives. First of all, suppose that CP extraposition in German targets the next higher CP domain (a right-peripheral specifier or adjunct) if extraction from the extraposed CP needs to take place. This implies that in order to permit a combination of CP extraposition and extraction from CP, an infinitive must

---

16 See Müller (1998) for arguments to this effect. If there is no extraction from CP, extraposition can also target a lower position, and then participate in VP topicalization. This accounts for the contrast in (i-a) (without extraction from the extraposed infinitive) and (i-b) (without extraposition) vs. (i-c) (with extraction from the extraposed infinitive).

(i) a. [VP₃ t₂ Versucht [CP₂ dem Peter das Buch₂ zu geben ] hat sie nicht t₃ tried the Peter_dat the book_acc to give has she_nom not
have sufficient weight to outweigh the AL violation automatically incurred by all movement across a finite \( C \); as we have seen, the latter has a harmony score of \(-2.7\). Assuming a feature \([\bullet \text{ex} \bullet]\) involved in extraposition to have a strength of \([0.7]\), it is correctly predicted that an infinitival CP with a \( C \) head with strength \([0.7]\) can undergo extraposition to the next higher CP domain, in optimal violation of AL: If movement does not take place, the resulting MC violation leads to a harmony score of \(-2.8\). All of this is shown in (27).

(27)  

**Infinitive extraposition in the third construction:**

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>AL</th>
<th>( \mathcal{H} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: ([\text{CP} \ C_{[0.9]},[\bullet \text{ex} \bullet];[0.7] \ldots \text{CP}<em>{[0.7]} \ V</em>{\text{restr}}])</td>
<td></td>
<td>2.0</td>
<td>3.0</td>
</tr>
<tr>
<td>O: [\text{CP} \ C_{[0.9]},[\bullet \text{ex} \bullet];[0.7] \ldots \text{CP}<em>{[0.7]} \ V</em>{\text{restr}}]\</td>
<td>-1.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[\text{CP} \ C_{[0.9]},[\bullet \text{ex} \bullet];[0.7] \ldots \text{CP}<em>{[0.7]} \ V</em>{\text{restr}}]\</td>
<td>-0.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Under these assumptions, it is clear that if the infinitival CP has a smaller weight of \([0.6]\), it can never be affected by extraposition to the CP domain – in this latter case, the harmony score of \(-2.6\) amassed by the MC-violating output is better than the harmony score of the AL-violating candidate that applies extraposition (which continues to be \(-2.7\)).

Finally, the lack of wide scope for negation in the third construction (and the concurrent availability of wide scope for negation in regular, preverbal restructuring infinitives) can also be tied to the different weights (\([0.7]\) vs. \([0.6]\)). I assume that scope of negation is in general the consequence of an Agree relation between an abstract operator position high in the clause and an overt negative item, which is typically in a much lower position in German (see Stechow (1993) and Zeijlstra (2004), among others). Agree is subject to an Agree Condition (AC; see Heck & Müller (2013)) that requires probe features (\([*F*]\)) to participate in Agree with appropriate goal features (\([F]\)). In the case at hand, there is a probe feature \([*\text{neg}*]\) on the overt negation (\(\text{nicht}\) in (25)), and a goal feature \([\text{neg}]\) in the left periphery of the matrix clause. Suppose furthermore that to bridge the distance in a local way that is compatible with the strict PIC employed here, Agree must take place cyclically (see Legate

\[ \text{VP}_3 \ [\text{CP}_2 \ \text{Dem Peter}_{\text{dat}} \ t_1 \ zu \ geben \] versucht \] hat sie das Buch$_1$ nicht t$_3$

\[ \text{VP}_3 \ t_2 \ Versucht \ [\text{CP}_2 \ \text{dem Peter}_{\text{dat}} \ t_1 \ zu \ geben \ ] \ hat sie das Buch$_1$

\[ \text{VP}_3 \ t_2 \ Versucht \ [\text{CP}_2 \ \text{Dem Peter}_{\text{dat}} \ t_1 \ zu \ geben \] hat sie das Buch$_1$

\[ \text{VP}_3 \ t_2 \ Versucht \ [\text{CP}_2 \ \text{dem Peter}_{\text{dat}} \ t_1 \ zu \ geben \] hat sie das Buch$_1$

\[ \text{VP}_3 \ t_2 \ Versucht \ [\text{CP}_2 \ \text{Dem Peter}_{\text{dat}} \ t_1 \ zu \ geben \] hat sie das Buch$_1$

\[ \text{VP}_3 \ t_2 \ Versucht \ [\text{CP}_2 \ \text{dem Peter}_{\text{dat}} \ t_1 \ zu \ geben \] hat sie das Buch$_1$
Such cyclic Agree will then also give rise to an AL violation for every head that it involves on the path to the ultimate target position in the matrix clause.\textsuperscript{17} On this basis, it can be concluded that the harmony score of an output that does not carry out cyclic Agree for a $[\ast \text{neg}\ast]$ feature across a CP and thereby violates AC must be better than $-2.1$ (so as be optimal vis-a-vis the harmony score of $-2.1$ resulting from AL if cyclic Agree across C applies in the third construction), but worse than $-1.8$ (so as to be suboptimal vis-a-vis the harmony score of $-1.8$ resulting from AL if cyclic Agree across C applies with regular restructuring infinitives). This result is achieved if, e.g., $[\ast \text{neg}\ast]$ has a weight of $[1.0]$, and AC has a weight of $[2.0]$. The competition underlying failed wide scope of negation in the third construction is illustrated in (28).

(28) **Wide scope of negation in the third construction:**

<table>
<thead>
<tr>
<th></th>
<th>AC</th>
<th>AL</th>
<th>$\mathcal{H}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[\ast \text{neg}\ast]$</td>
<td>$w = 2.0$</td>
<td>$w = 3.0$</td>
<td>$-2.1$</td>
</tr>
</tbody>
</table>

Thus, the PIC will block any non-local transmittance of $[\ast \text{neg}\ast]$, and an Agree relation with the target position in the matrix clause cannot be established in the third construction. Of course, with a lower C weight of $[0.6]$ (as in regular restructuring infinitives), the candidate that carries out (intermediate) cyclic Agree with the C head (as required by AC) becomes optimal: Now the violation of AL is less severe (yielding a harmony score of $-1.8$).\textsuperscript{18}

\textsuperscript{17}Strictly speaking, given the definition of AL in (10), this presupposes that if there is a (cyclic) Agree dependency between $[(\ast \text{neg}\ast)]$ on some head Y and $[\ast \text{neg}\ast]$ on an item c-commanded by Y, the former feature must be $m$-commanded by Y to generate an AL violation. Depending on the exact nature of feature insertion in cyclic Agree contexts and the precise definition of $m$-command, this may either follow directly, or it may require a generalization of the concept of $m$-command (e.g., along the lines of Chomsky’s (1995) notion of minimal residue).

\textsuperscript{18}Ultimately, a bit more will have to be said. E.g., it is generally held that narrow scope of negation is in fact impossible in standard restructuring infinitives. This does not yet follow from the analysis; an obvious possibility here might be to assume that a certain strength of C is required to license an interpretable [neg] feature. In this context, it is worth pointing out that the present approach to scope of negation in terms of cyclic Agree is by far not the only one that can be entertained. One could, e.g., assume that AC-driven Agree does not obey the PIC (cf., e.g., Bošković (2007)), and then let the strength differences of the two infinitival C heads (restructuring vs. third construction) interact with a violable intervention constraint.
6. Strength and Morphological Realization

In section 2 above, I concluded that strength is an abstract property of heads that can have two different consequences: First, it determines whether or not syntactic operations can apply, and second, it also determines post-syntactic morphological realization. In the present study of strength of C in German I have focussed on the former issue; to end this paper, let me make a few remarks on the latter one.

In Lee (2018), it is argued that finite declarative C in English comes in two versions distinguished only by their strength. Strong C blocks wh-movement of subjects (but not of objects, which are themselves stronger than subjects); weak C does not. Transferring this analysis to the present approach in terms of MC and AL, this follows if weak C has a weight of [0.5] in English, strong C has a weight of [1.0], [wh] has a weight of [0.8], and subject and object DPs have weights of [0.4] and [0.8], respectively. Crucially, Lee (2018) shows that these different weight assignments to declarative finite C in English can also be assumed to govern post-syntactic morphological realization. A strong C:[1.0] gives rise to a severe (and fatal) violation of a constraint demanding vocabulary insertion if it is not post-syntactically realized by *that*; in contrast, with a weak C:[0.5], the violation of this constraint is not so severe anymore, and the violation of a DÉP constraint prohibiting vocabulary insertion that is incurred by the presence of *that* becomes fatal. Thus, the complementizer-trace effect in (3b) (vs. (3a)) is derived without giving up the assumption that the morphological shape of C is determined only post-syntactically.

In the same way, the fact that finite declarative C can be morphologically realized by *dass* in German whereas the non-finite Cs of control infinitives are not realized by morphological exponents does not emerge as fully accidental under present assumptions: The former kind of C is stronger than the latter ones ([0.9] vs. [0.6], [0.7], [0.8]). Thus, whereas one might abstractly conceive of a variety of German where, e.g., Cs of non-restructuring infinitives are also overtly realized in some way whereas Cs of the third construction and restructuring infinitives are not, the prediction clearly is that it would ceteris paribus be impossible to have a variety of German where the Cs that are more transparent to movement are overtly realized, and Cs that are less transparent remain without morphological exponence. I take this to be a non-trivial and welcome result.
References


Torres-Tamarit, Francesc (2016). Compensatory and Opaque Vowel Lengthening in


