On Accelerating and Decelerating Movement: From Minimalist Preference Principles to Harmonic Serialism

Fabian Heck & Gereon Müller
Leipzig University

Abstract
Derivations in the minimalist program (MP) frequently encounter competition between elementary operations. A way in MP to resolve such competitions is to assume preference principles, the most prominent being "Merge before Move" (Chomsky 2000). Assmann et al. (2012) discuss competition between Agree and Move. There, either Agree must be procrastinated in favor of Move or the other way round. Procrastination, in turn, suggests constraint violability and thus an optimality theoretic account. While it is, in principle, possible to formulate an inviolable constraint such that it mimics the effects of two interacting violable constraints, such a move is arguably conceptually unattractive as it requires the constraint in question to be complex. Moreover, if it turns out that there are scenarios where the general preference expressed by the constraint is exceptionally reversed, then it has to be complicated further, shedding serious doubt on its plausibility. In this paper, we argue that scenarios of exceptional preference reversal do indeed exist. In optimality theory, they are captured straightforwardly by assuming a higher ranked constraint that brings about the change.

1 Introduction

Syntactic derivations in feature-based theories such as the minimalist programm (Chomsky 1995, 2000, 2001) often involve competition between elementary operations. This means that there are derivational stages where more than one operation may in principle apply. Assuming a general earliness requirement (Pesetsky 1989, Chomsky 1995: 233, Lasnik 1999: 198), operations apply as soon as their context for application is present. But if there is no simultaneous rule application in grammar (see Epstein and Seely 2002; contra Pullum 1979, Chomsky 2008), then a conflict arises: More than one operation should apply immediately, yet only one of them can be executed at each step. Consequently, competition between operations can arise, which must be resolved by giving preference to (ranking) one or the other operation.

For instance, Chomsky (1995, 2000) observes that there are derivational stages where both Merge (external Merge, EM) and Move (internal Merge, IM) can in principle apply. Chomsky (1995, 2000) argues on the basis of contrasts such as (1) that there is a general preference to apply Merge before Move.

(1) a. There$_1$ seems [$_{\text{TP}}$ t$_1$ to be [$_{\text{PP}}$ someone$_2$ in the room]]
    b. *There$_1$ seems [$_{\text{TP}}$ someone$_2$ to be [$_{\text{PP}}$ t$_2$ in the room]]

The embedded SpecT-position in (1) can be filled in two ways: Either *someone* moves or the expletive *there* is merged. If *there* is merged, then it can undergo movement to the matrix SpecT-position at some later step, deriving (1-a). If *someone* moves to the embedded SpecT first, then *there* will be merged into the higher SpecT later, deriving the ungrammatical (1-b). In order to block (1-b), Chomsky (2000) proposes the preference principle in (2).

(2) *Merge before Move*:
Suppose that the derivation has reached stage Σ_n, and Σ_{n+1} is a legitimate instance

of Merge, and Σ'_{n+1} is a legitimate instance of Move. Then, Σ_{n+1} is to be preferred over Σ'_{n+1}.

The question arises as to whether the order of Merge and Move can be derived from more general assumptions about the make-up of these operations. Chomsky (2000) suggests that Merge is simpler than Move because Move might be Merge plus Agree, plus Pied Piping, etc., which should explain the preference in terms of economy. However, in contrast to this, Chomsky (2013) states that, if anything, Move should be simpler than Merge "since it requires vastly less search" because external Merge "must access the workspace of already generated objects and the lexicon". It is also worth noting that on the basis of Chomsky's (2000) assumptions about the complexity of Merge and Move (with Move emerging as less simple, i.e., more specific), the specificity-based preference principle for ordering operations postulated by Koutsoudas (1966, 1973) and Pullum (1979) (also see van Koppen 2005, Lahne 2012, and Georgi 2013 for similar more recent concepts) would in fact also predict a reverse Move before Merge outcome. Perhaps the lack of an uncontroversial, obvious evaluation metric for ordering the two operations can be taken to indicate that both resolutions are in principle available in natural languages: Conflicts between elementary operations are resolved by ranking (giving preference to one of the two options), but there is no inherently fixed resolution strategy.[1]

Thus, sometimes the order of applying Merge and Move is under-determined. The conflict can be resolved by ranking the requirements: The highest-ranked requirement is satisfied immediately; lower-ranked ones must remain unsatisfied at the current derivational step. Such unsatisfiability does not lead to a crash of the derivation and thus suggests an analysis in terms of violable constraints.

While (2) is (mildly) transderivational in nature, one may argue that it does not require violability of the constraint demanding the application of Move (which is procrastinated due to the preference principle): If the constraint does not require movement as such but rather the filling of some specifier position, then it can be equally well fulfilled by Merge. However, it has been argued that there is also competition between Agree and Move (see Assmann et al. 2012), two operations for which it is less likely that their application is reducible to an identical trigger. Thus, in this case either Agree must be procrastinated in favor of Move or the other way round. Procrastination, in turn, presupposes constraint violability and thus suggests an optimality theoretic account. While it is, in principle, possible to formulate an inviolable constraint that mimics the effects of two interacting violable constraints, such a move is conceptually unattractive as it requires the constraint to be complex (see also the introduction to this volume). What is more, if there are scenarios where the preference expressed by the complex constraint is exceptionally reversed, then the constraint must be further complicated, increasing the conceptual burden of the approach and thus rendering an alternative approach in terms of violable constraints more plausible.

In section 2, we briefly report the analysis of Assmann et al. (2012). There, it is suggested that resolving the competition between Move and Agree in morphologically ergative languages by giving preference to Move over Agree accounts for a restriction on ergative movement (on the TP cycle) in these languages. At the same time, the analysis explains the absence of a parallel restriction on accusative movement in morphologically accusative languages, where

[1] And perhaps there are also systematic conflicts and resolutions among different types of Move, as in intermediate vs. criterial movement steps; see Georgi (2013).

the competition between Move and Agree is resolved by giving Agree preference over Move.[2]

Based on this background, we then illustrate in section 3 that movement in accusative encoding systems (where Move is usually procrastinated in favor of Agree) is in some contexts *accelerated*, so that it applies before Agree. Further we illustrate in section 4 that movement in ergative encoding systems (where Move usually applies before Agree) is in some contexts *decelerated*, so that it applies after Agree. This accounts for a priori unexpected mobility restrictions on dative arguments in German, and for a priori unexpected movement options for ergative arguments in some certain ergative languages (Chol, Basque, Avar, and Pitjantjatjara). The upshot will be that the effects illustrated in sections 3 and 4 can straightforwardly be derived in an optimality-theoretic approach (they signal the presence of more specific, higher-ranked constraints). However, as suggested above, they are less straightforwardly derivable in a more orthodox minimalist approach.

The reasoning presented here presupposes an extremely local, derivational approach to optimization. In other words, the domain for optimization (conflict resolution) is the minimal derivational step (Epstein and Seely 2002, Heck and Müller 2007, 2013, McCarthy 2010). If the optimization domain is larger than the step-level, wrong empirical predictions are made.

2 Move vs. Agree: A Constraint on Ergative Movement

2.1 The Phenomenon

The starting point of Assmann et al. (2012) is the observation that in many morphologically ergative languages ergative arguments (DP_{erg}) cannot undergo Ā-movement, i.e., they cannot undergo *wh*-movement, focussing, or relativization (see Campana 1992, Aldridge 2004, Stiebels 2006, Coon et al. 2011, Deal 2012). (3) briefly illustrates this for *wh*-movement in the Mayan language Kaqchickel.

(3) *Wh-Movement of DP_{erg} vs. DP_{abs} in Kaqchikel (Mayan)*:

a. *achike n-Ø-u-löq' jun sik'iwuj?
   Q INCOMPL-3SG.ABS-3SG.ERG-buy INDEF book
   'Who buys a book?'

b. atux n-Ø-u-löq' a Carlos?
   Q INCOMPL-3SG.ABS-3SG.ERG-buy CL Carlos
   'What does Carlos buy?'

c. achike ri n-Ø-tze'en?
   Q DET INCOMPL-3SG.ABS-laugh
   'Who laughs?'

(3-a) involves *wh*-movement of an ergative subject, which is ungrammatical.[3] Nothing is wrong with *wh*-moving a subject per se, as the grammatical case of *wh*-movement of a absolutive marked subject in (3-c) illustrates. Finally, *wh*-movement of an absolutive marked object (3-b) is also impeccable.

[2]The relevant rankings for ergative type languages and accusative type languages are independently motivated by the theory of argument encoding proposed in Müller (2009).

[3]Strictly speaking, argument DPs in Mayan languages do not bear overt case markers, but ergative and absolutive DPs trigger different kinds of agreement: DP_{ext} triggers ergative agreement whereas DP_{int} and the sole argument of an intransitive verb trigger absolutive agreement.

Assmann et al. (2012) propose an account of the restriction on Ā-movement of the ergative DP that is co-argument-based. The leading idea is that there is nothing wrong with movement of the ergative marked external argument as such. Rather, movement of the ergative invariably leads to a loss of the absolutive case provided for the internal argument, thus leaving the internal argument caseless, which leads to a violation of the case filter (Rouveret and Vergnaud 1980).

2.2 Theoretical Assumptions

To begin with, the clause structure in (4) is assumed in Assmann et al. (2012).

(4) *Clause structure*:
$[_{\text{CP}}$ C $[_{\text{TP}}$ T $[_{\text{vP}}$ DP$_{ext}$ $[_{\text{v}'}$ v $[_{\text{VP}}$ V DP$_{int}$ $]]]]]$

The internal argument DP$_{int}$ is the complement of the verb. The external argument DP$_{ext}$ is introduced as the specifier of the functional head v (Chomsky 1995, Kratzer 1996), which takes VP as its complement. There are two other functional heads above v, namely T and C.

Following Chomsky (1995), all operations are assumed to be feature-driven in Assmann et al. (2012). Agree is triggered by probe features (5-a). Merge (external and internal) is triggered by subcategorization/edge features (Svenonius 1994, Sternefeld 2006, Chomsky 2007, 2008), see (5-b).

(5) *Two types of features that drive operations*:
  a. Probe features trigger Agree: [∗F∗].
  b. Subcategorization features/edge features trigger Merge: [•F•].

The definitions of the operations Merge, Move, and Agree (cf. Chomsky 2001), which are driven by these features, are given in (6)–(8), respectively. Note that some features may lack a value, which they must acquire by entering into Agree with another feature that bears a value (Chomsky 2000, 2001). A feature [F] that lacks a value is rendered as [F:□].

(6) *Merge*:
α can undergo merge with β, yielding [α β], if α bears a structure-building feature [•F•] and F is the label of β.

(7) *Move*:
Move is Merge, with β internal to α.

(8) *Agree*:
α agrees with β with respect to a feature bundle Γ iff (a) and (b) hold:
  a. α bears a probe feature [∗F∗] in Γ and may thereby provide the α-value for a matching goal feature [F] of β in Γ.
  b. α m-commands β.

Note that (8-b) permits an Agree relation between a head and its specifier. Incidentally, Assmann et al. (2012) assume that Agree by a head H with its (innermost) specifier is not only possible but is actually preferred over Agree by H with any item bearing another structural relation towards H (see Chomsky 1986*a*, 1995, Kayne 1989, Koopman 2006 for related proposals; Béjar and Řezáč 2009 express a similar idea with the bias reversed). This principle is dubbed the Specifier-Head Bias in Assmann et al. (2012):

(9) SPECIFIER-HEAD BIAS (SHB):
Agree between (first) specifier and head is preferred to other instances of Agree.

To a certain extent, (9) replaces standard minimality conditions such as Relativized Minimality (Rizzi 1990) or the Minimal Link Condition (Fanselow 1991, Ferguson 1993, Chomsky 1995), though with somewhat different empirical coverage. Müller (2004, 2011) argues that further effects usually accounted for by standard minimality conditions can be derived from the PIC (see below). At the same time, the SHB is compatible with equi-distance effects, which pose a problem for path-based, or closest c-command-based, definitions of minimality. It is therefore assumed in Assmann et al. (2012) that minimality as such does not exist and that the effects traditionally attributed to it derive from independent principles (such as SHB and PIC).

The designated constraints in (10) and (11) ensure that Merge (incl. Move) and Agree must take place as soon as their context of application is present (Heck and Müller 2007, 2013). This derives the earliness requirement for syntactic operations that was mentioned above.

(10) AGREE CONDITION (AC):
Probes ([∗F∗]) participate in Agree.

(11) MERGE CONDITION (MC):
Structure-building features ([•F•]) participate in Merge.

Every argument must receive structural case in the syntax (so as not to violate some form of the case filter). Structural case is assigned by the heads v and T under Agree. By standard assumption, T and v have valued case probe features that assign their value α to DPs with an unvalued case feature [CASE:□]. Assmann et al. (2012) follow Murasugi (1992) (see also Jelinek 1993, Ura 2000, Müller 2009) in assuming that in ergative as well as in accusative languages T assigns the unmarked structural case (i.e., nominative = absolutive) and v assigns the marked structural case (i.e., ergative = accusative). In intransitive contexts only T is active, so the single argument receives the unmarked case.

More specifically, the assumption is that there is a single structural case feature [CASE]. This feature can have the two values ext(ernal) and int(ernal). The unmarked case (nominative/absolutive) is represented as the external case [CASE:*ext*] and the marked case (ergative/accusative) as the internal case [CASE:*int*]. Since T assigns unmarked external case and v assigns the marked internal case, these heads bear the following probe features:

(12) *The role of T and v in argument encoding*:
a. T bears [∗CASE:*ext*∗] that instantiates [CASE:*ext*] on DP.
b. v bears [∗CASE:*int*∗] that instantiates [CASE:*int*] on DP.

Turning to the issue of locality, Assmann et al. (2012) suggest that movement that starts from within the vP-domain and targets SpecC must obligatorily make an intermediate movement step to SpecT. This is achieved by assuming that, generally, movement takes place successive-cyclically, from one XP edge domain to the next one higher up (see Sportiche 1989: 36, 45-47, Boeckx 2003: 16-25, Müller 2004, Chomsky 2005: 18, among others). Given the Phase

Impenetrability Condition (PIC; Chomsky 2001) in (13) and the notion of edge in (14), this follows automatically if every XP is a phase.

(13) *Phase Impenetrability Condition (PIC; Chomsky 2001)*:
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.

(14) *Edge*:
The edge of a head X comprises all specifiers of X (and adjuncts to XP).

Additionally, it must be ensured in a theory of syntax where all operations are feature-driven that intermediate steps of movement as required under the PIC are possible in the first place. A standard assumption here is that category neutral edge features ([•X•]) can be inserted on all intervening phase heads (Chomsky 2007, 2008). These edge features then trigger intermediate movement steps.[4]

The following assumptions pertain to the activity of structural case features and the way that probe features and the goals they enter into Agree with are paired off. First, suppose that a structural case goal G can enter into Agree with a case probe P even if G has already acquired a value via Agree with another probe P′ at a previous stage of the derivation. That is, structural case features on arguments remain active throughout the derivation. This is explicitly stated in (15) (cf. Merchant 2006).

(15) *Activity of structural case features*:
Structural case features act as active goals.

Independent motivation for (15) comes from the existence of case stacking, as it exists in some of the world's languages (see Andrews 1996, Nordlinger 1998, Richards 2007, Merchant 2006).

Imagine now a situation where an argument with a structural case goal G enters into a Spec/head configuration with a functional head that bears a case probe P. Assume that P has not yet entered into Agree and that it is ultimately supposed to provide a value for another case goal G′. In this situation, P must enter into Agree with G (instead of G′) due to the Specifier-Head Bias, even though G has already acquired a case value while G′ has not. In this way, G *marauds* the functional head by taking away its case probe, which should normally be reserved for G′ (see Georgi et al. 2009 on maraudage; similar concepts are suggested in Chomsky 2001, Abels 2003, Anagnostopoulou 2005, Adger and Harbour 2007, Béjar and Řezáč 2009.)

The situation is abstractly depicted in (16). The configuration in (16-a) may involve Agree between [∗CASE:*ext*∗] on X and [CASE:*int*] on α or not. If Agree involves [CASE:*int*], this leads to a crash of the derivation because there remains an unchecked [CASE:□] on β. If, however, [∗CASE:*ext*∗] enters into Agree with [CASE:□] on β, then the derivation converges, which is sufficient to ensure grammaticality. The situation is different in (16-b). Here, α is in a Spec/head configuration with X. Thus, the SHB forces Agree between [∗CASE:*ext*∗] and [CASE:*int*], thereby leaving [CASE:□] on β without a value. It follows that (16-b) invariably leads to a crash.

(16) a. $[_{X'}\ X_{[*\text{case:ext}*]}\ [_{ZP}\ \ldots\ \alpha_{[\text{case:int}]}\ \ldots\ \beta_{[\text{case:}\square]}\ \ldots\ ]]$

[4] For reasons of perspicuity, let as assume here that edge features can only attract elements that are designated for Ā-movement.

b. $[_{\text{XP}}\ \alpha_{[\text{case:int}]}\ [_{\text{X}'}\ \text{X}_{[*\text{case:ext}*]}\ [_{\text{ZP}}\ \dots\ \text{t}_{\alpha}\ \dots\ \beta_{[\text{case:}\square]}\ \dots\ ]]]$

Recall in this context Assmann et al.'s (2012) assumption that there is no minimality condition. Given general activity of structural case goals, this is necessary to ensure convergence of at least one of the derivations based on (16-a). However, it is not sufficient. What is needed in addition is that both α and β are PIC-accessible to X in (16). This implies that the PIC is slightly less restrictive, as eventually proposed in Chomsky (2001), or that Agree operations can escape the PIC, as suggested by Bošković (2007), among others.

A question arises as to how multiple case agreement is possible. The presupposition behind (15) is that Agree of [CASE:*int*] on α with a conflicting [*CASE:*ext**] on X is harmless as such. α simply maintains its original feature value, which then accordingly surfaces in morphology; or α adds the new case feature, leading to case stacking (cf. Assmann et al. 2013). However, [*CASE:*ext**] on X is then discharged, and not available for further operations anymore.

Finally, Assmann et al. (2012) adopt an idea put forward in Müller (2009) that ergative type languages vs. accusative type languages are distinguished by the relative ordering of Merge and Agree: The ranking MC $\gg$ AC derives ergative type encoding systems (by assigning the internal case of v to the external argument in Specv, due to the SHB which preferes Specv to CompV if both external and internal argument are present in the structure when AC needs to be satisfied), while the reversed ranking generates accusative type encoding systems (by assigning the internal case of v to the internal argument in the VP, which is the only DP requiring structural case that is present at the point of the derivation where AC must be satisfied under this ranking, with Merge of the external argument delayed).

In the following section, we introduce an optimality theoretic variant of the analysis of Assmann et al. (2012), which our arguments in sections 3 and 4 will then be based upon.

2.3 Analysis

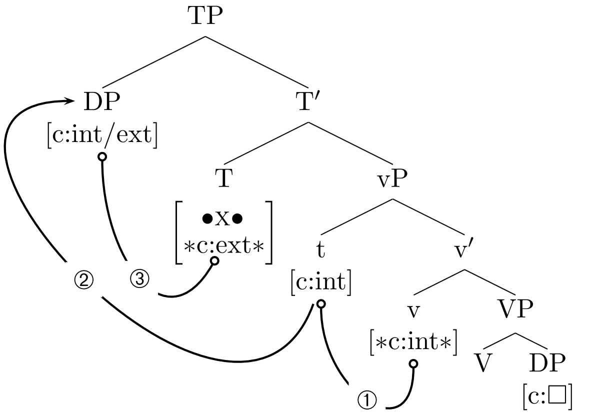
2.3.1 Displacement in Languages with Ergative Encoding Patterns

According to Müller (2009), the ordering conflict between Merge and Agree in a morphologically ergative language is resolved by the ranking MC $\gg$ AC. Imagine a situation where DP_{ext} is a *wh*-phrase that is supposed to undergo Ā-movement (*wh*-movement, relativization, focus movement) and ultimately show up in some SpecC-position. The details of deriving the ergative encoding system need not concern us here. It is sufficient to recall that v assigns internal case (= ergative). It follows that DP_{ext} must have its case feature valued as ergative while it still resides in the m-command domain of v, i.e., upon completion of the vP. At this point, DP_{int} still bears an unvalued case probe, awaiting valuation by T, see step ① in (17).[5]

Suppose now T is introduced into the structure. Given the PIC, DP_{ext} needs to move from Specv to SpecT if it is to undergo subsequent movement to SpecC. Based on the null hypothesis that the ranking MC $\gg$ AC that leads to ergative type encoding systems on the vP-cycle is also maintained on the TP cycle, movement of DP_{ext} (as an instance of internal Merge) will have to *precede* Agree of T with DP_{int}, which has not yet valued its case feature (as absolutive), see step ② in (17). The optimization of this derivational step is illustrated in tableau T_1.

[5]Here and henceforth, case probe features that have participated in Agree are signalled by underlining in trees; they are rendered here only so as to enhance perspicuity.

(17) *Illegitimate movement of DP_{erg}:*



Finally, given the SHB, DP_{ext} will next maraud T's case probe, see step ③ in (17). The relevant optimization is given in tableau T_2. The internal argument DP will consequently remain without a valued case feature. Assuming that all DPs must have their case features valued eventually (and assuming that there is no such thing as a default case in a normal transitive clause where all arguments could in principle get their cases valued), the derivation will crash. In a nutshell, ergative movement is impossible because it applies too early, thereby bleeding absolutive case assignment to DP_{int}.

T_1: Ergative movement, step 1: Move

Input: $[_{T'}\ T_{[*\text{CASE}:ext*],[\bullet X\bullet]}\ [_{vP}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots$ $\ldots DP_{[\text{CASE}:\square]} \ldots]]]]$	SHB	MC	AC
☞ O_1: $[_{TP}\ DP_{[\text{CASE}:int]}\ [_{T'}\ T_{[*\text{CASE}:ext*]}\ [_{vP}\ t\ [_{v'}\ v \ldots$ $\ldots DP_{[\text{CASE}:\square]} \ldots]]]]$			*
O_2: $[_{T'}\ T_{[\bullet X\bullet]}\ [_{vP}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:ext]} \ldots]]]$		*!	

T_2: Ergative movement, step 2: Agree (maraudage)

Input: $[_{TP}\ DP_{[\text{CASE}:int]}\ [_{T'}\ T_{[*\text{CASE}:ext*]}\ [_{vP}\ t\ [_{v'}\ v \ldots$ $\ldots DP_{[\text{CASE}:\square]} \ldots]]]]$	SHB	MC	AC
☞ O_1: $[_{TP}\ DP_{[\text{CASE}:ext/int]}\ [_{T'}\ T\ [_{vP}\ t\ [_{v'}\ v \ldots$ $\ldots DP_{[\text{CASE}:\square]} \ldots]]]]$			
O_2: $[_{TP}\ DP_{[\text{CASE}:int]}\ [_{T'}\ T\ [_{vP}\ t\ [_{v'}\ v \ldots DP_{[\text{CASE}:ext]} \ldots]]]]$	*!		

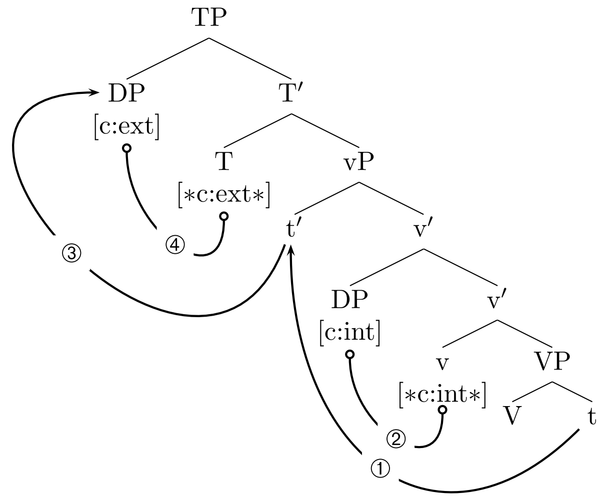
It is assumed here that a violation of the case filter eventually leads to a crash of the derivation. This means that an unvalued case feature represents an instance of Grimshaw's (1994) "no good output" approach to absolute ungrammaticality (or 'ineffability'): The optimal candidate is characterized by a property that gives rise to problems at the interfaces.[6]

Next, consider the case where DP_{int} undergoes Ā-movement. First, DP_{ext} is merged in an inner Specv and DP_{int} moves into an outer Specv (step ① in (18)), triggered by an edge feature inserted in v. These operations apply first, given the ranking MC ≫ AC. After this,

[6]For instance, it may be the case that arguments are not θ-visible at the LF-interface unless they bear case (see Chomsky 1986*b*, Baker 1988).

DP_{ext} enters into Agree with the case probe on v (② in (18)), thereby receiving ergative case (18-a).[7]

(18) *Legitimate movement of DP_{abs}:*



If DP_{int} is to remain accessible for further movement (to SpecC), it first has to raise to SpecT, see ③ in (18). MC ≫ AC forces this intermediate movement step to apply before Agree values absolutive case on DP_{int}. Tableau T_3 illustrates the optimization. Finally, the case probe on T enters into Agree with the case feature on DP_{int}, valuing the latter as absolutive (step ④). As DP_{ext} has already received its case value on the vP-level, the derivation converges. The optimization is shown in tableau T_4.

T_3: Absolutive movement, step 1: Move

Input: $[_{T'}$ $T_{[*CASE:\mathit{ext}*],[\bullet X\bullet]}$ $[_{vP}$ $DP_{[CASE:\square]}$ … …$[_{v'}$ $DP_{[CASE:\mathit{int}]}$ $[_{v'}$ v … t …]]]]	SHB	MC	AC
☞ O_1: $[_{TP}$ $DP_{[CASE:\square]}$ $[_{T'}$ $T_{[*CASE:\mathit{ext}*]}$ $[_{vP}$ t′ … …$[_{v'}$ $DP_{[CASE:\mathit{int}]}$ $[_{v'}$ v … t …]]]]]			*
O_2: $[_{T'}$ $T_{[\bullet X\bullet]}$ $[_{vP}$ $DP_{[CASE:\mathit{ext}]}$ … …$[_{v'}$ $DP_{[CASE:\mathit{int}]}$ $[_{v'}$ v … t …]]]]		*!	

T_4: Absolutive movement, step 2: Agree (with SpecT)

Input: $[_{TP}$ $DP_{[CASE:\square]}$ $[_{T'}$ $T_{[*CASE:\mathit{ext}*]}$ $[_{vP}$ t′ … …$[_{v'}$ $DP_{[CASE:\mathit{int}]}$ $[_{v'}$ v … t …]]]]]	SHB	MC	AC
☞ O_1: $[_{TP}$ $DP_{[CASE:\mathit{ext}]}$ $[_{T'}$ T $[_{vP}$ t′ … …$[_{v'}$ $DP_{[CASE:\mathit{int}]}$ $[_{v'}$ v … t …]]]]]			
O_2: $[_{TP}$ $DP_{[CASE:\square]}$ $[_{T'}$ T $[_{vP}$ t′ … …$[_{v'}$ $DP_{[CASE:\mathit{ext/int}]}$ $[_{v'}$ v … t …]]]]]	*!		

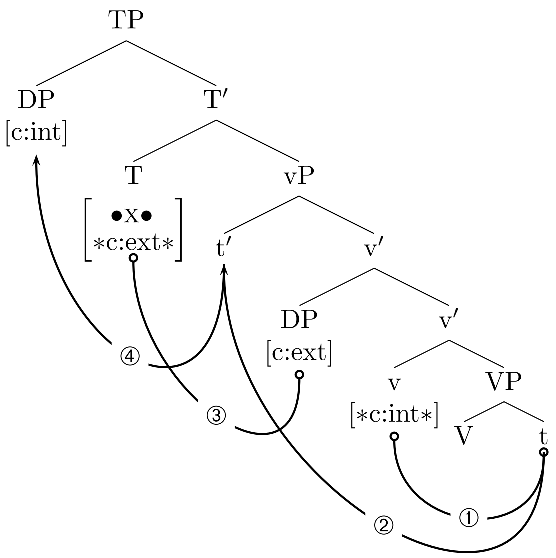
[7] The SHB blocks Agree between DP_{int} and v. This raises the question as to what prevents DP_{int} from occupying the innermost specifier (leaving the outer specifier for DP_{ext}) and thus receiving internal case. Such a derivation would wrongly lead to an accusative encoding pattern. To block it, Assmann et al. (2012) assume a preference for Merge over Move.

2.3.2 Displacement in Languages with Accusative Encoding Patterns

We now show how Assmann et al. (2012) account for the absence of a parallel restriction on movement of the accusative argument in morphologically accusative type languages.

According to Müller (2009), the ranking in accusative type languages is AC $\gg$ MC. This ranking, giving rise to an accusative pattern in the first place (on the vP cycle), is also active on the TP cycle. Thus, in a derivation where DP_{int} is supposed to undergo extraction, it will target an outer specifier of vP – see step ② in (19) – after its case feature has been valued accusative by the probe on v (step ①). Once T is merged, AC $\gg$ MC ensures that case on DP_{ext} gets valued nominative in step ③ *before* DP_{int} moves on to SpecT, see tableau T_5.

(19) *Legitimate movement of* $\textit{DP}_{acc}$:



Finally, in step ④ DP_{int} moves on to SpecT to satisfy an edge feature on the T-head. The optimization is shown in tableau T_6.

T_5: Accusative Movement, step 1: Agree

Input: $[_{\text{T}'}\ \text{T}_{[*\text{CASE}:ext*],[\bullet\text{X}\bullet]}\ [_{\text{vP}}\ \text{DP}_{[\text{CASE}:int]} \ldots \ldots [_{\text{v}'}\ \text{DP}_{[\text{CASE}:\square]}\ [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]$	SHB	AC	MC
O_1: $[_{\text{TP}}\ \text{DP}_{[\text{CASE}:int]}\ [_{\text{T}'}\ \text{T}_{[*\text{CASE}:ext*]}\ [_{\text{vP}}\ \text{t}' \ldots \ldots [_{\text{v}'}\ \text{DP}_{[\text{CASE}:\square]}\ [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]]$		*!	
☞ O_2: $[_{\text{T}'}\ \text{T}_{[\bullet\text{X}\bullet]}\ [_{\text{vP}}\ \text{DP}_{[\text{CASE}:int]} \ldots \ldots [_{\text{v}'}\ \text{DP}_{[\text{CASE}:ext]}\ [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]$			*

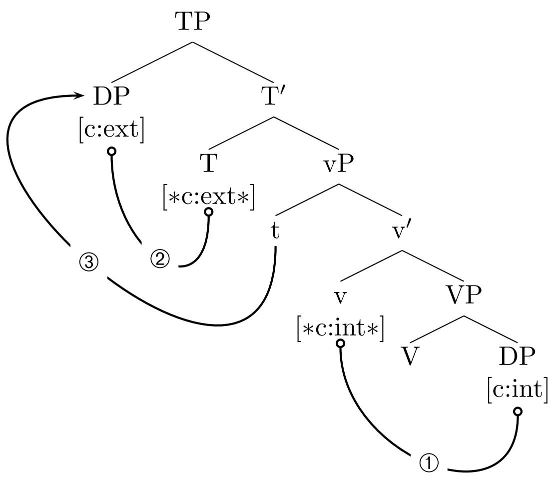
T_6: Accusative Movement, step 2: Move

Input: $[_{\text{T}'}\ \text{T}_{[\bullet\text{X}\bullet]}\ [_{\text{vP}}\ \text{DP}_{[\text{CASE}:int]} \ldots \ldots [_{\text{v}'}\ \text{DP}_{[\text{CASE}:ext]}\ [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]$	SHB	AC	MC
☞ O_1: $[_{\text{TP}}\ \text{DP}_{[\text{CASE}:int]}\ [_{\text{T}'}\ \text{T}\ [_{\text{vP}}\ \text{t}' \ldots \ldots [_{\text{v}'}\ \text{DP}_{[\text{CASE}:ext]}\ [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]]$			
☞ O_2: $[_{\text{TP}}\ \text{DP}_{[\text{CASE}:ext]}\ [_{\text{T}'}\ \text{T}\ [_{\text{vP}}\ \text{DP}_{[\text{CASE}:int]} \ldots \ldots [_{\text{v}'}\ \text{t}'\ [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]$			

Nothing so far rules out O_2 in T_6. However, because of the PIC, only DP_{ext} can move on in the subsequent steps of the derivation. Eventually, this leads to unchecked operator features on the attracting head and DP_{int}, and thus to a crash of the derivation.[8]

Finally consider a derivation involving Ā-movement of DP_{ext}. Similarly to movement of DP_{int} in ergative type systems, there is no problem for movement of DP_{ext} (=DP_{nom}) in accusative type systems because DP_{int} has already been assigned case when DP_{ext} moves. As step ① in (20) shows, the case feature of DP_{int} is valued as accusative already within vP. When T has been merged, AC ≫ MC dictates valuation of the case feature of DP_{ext} to apply before movement of DP_{ext}, see step ② and tableau T_7. Finally, when all cases have been valued, DP_{ext} moves to SpecT, see step ③ and tableau T_8. From there, it can move on to SpecC.

(20) *Legitimate movement of DP_{nom}:*



T_7: Nominative movement, step 1: Agree

Input: $[_{T'}\ T_{[*\text{CASE}:ext*],[\bullet X\bullet]}\ [_{vP}\ DP_{[\text{CASE}:\square]} \ldots \ldots [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]]$	SHB	AC	MC
O_1: $[_{TP}\ DP_{[\text{CASE}:\square]}\ [_{T'}\ T_{[*\text{CASE}:ext*]}\ [_{vP}\ t \ldots \ldots [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]]]$		*!	
☞ O_2: $[_{T'}\ T_{[\bullet X\bullet]}\ [_{vP}\ DP_{[\text{CASE}:ext]} \ldots \ldots [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]]$			*

T_8: Nominative movement, step 2: Move

Input: $[_{T'}\ T_{[\bullet X\bullet]}\ [_{vP}\ DP_{[\text{CASE}:ext]} \ldots \ldots [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]]$	SHB	AC	MC
☞ O_1: $[_{TP}\ DP_{[\text{CASE}:ext]}\ [_{T'}\ T_{[\bullet X\bullet]}\ [_{vP}\ t \ldots \ldots [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]]]$			

[8]Again, we assume, as is often done, that unchecked probe features cause legibility problems at LF.

2.4 Extremely Local vs. Less Local Optimization

It is crucial for the analysis that optimization applies to the single derivational step. If the optimization domain is not the derivational step but rather comprises phrases (phases, clauses, sentences), then a wrong prediction is made for accusative contexts: Maraudage would be expected to arise, and thus one would expect (at least some) morphologically accusative languages to exhibit a restriction on Ā-movement of the accusative argument, parallel to the restriction that shows up in many morphologically ergative languages. To wit, if optimization applies at the phrase level, then the order of operations induced by the accusative type ranking AC ≫ MC is lost: The optimal TP will always have its specifier filled by DP_{int} and thus SHB will force Agree between [∗CASE:*ext*∗] on T and DP_{int}, and make case assignment to DP_{ext} impossible. This is shown in tableau T_9.

T_9: TP optimization under AC ≫ MC ("accusative") ranking: wrong result

Input: $\mathrm{T}_{[*\text{CASE}:ext*],[\bullet X\bullet]} \oplus [_{\text{vP}}\ \mathrm{DP}_{[\text{CASE}:int]}\ [_{\text{v}'}\ \mathrm{DP}_{[\text{CASE}:\square]} \ldots \ldots [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]$	SHB	AC	MC
O_1: $[_{\text{TP}}\ \mathrm{DP}_{[\text{CASE}:int]}\ [_{\text{T}'}\ \text{T}\ [_{\text{vP}}\ \text{t}'\ [_{\text{v}'}\ \mathrm{DP}_{[\text{CASE}:ext]} \ldots \ldots [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]]$	*!		
☛ O_2: $[_{\text{TP}}\ \mathrm{DP}_{[\text{CASE}:ext/int]}\ [_{\text{T}'}\ \text{T}\ [_{\text{vP}}\ \text{t}'\ [_{\text{v}'}\ \mathrm{DP}_{[\text{CASE}:\square]} \ldots \ldots [_{\text{v}'}\ \text{v} \ldots \text{t} \ldots]]]]]$			

Finally, note that the derivation with extraction of the accusative includes a stage that represents an interesting case of opacity, namely counter-bleeding (Chomsky 1951, 1975, Kiparsky 1973): When the moved accusative DP_{int} occupies SpecT, one would expect it to maraud T's case probe, thereby bleeding nominative case valuation of DP_{ext}. However, no such bleeding takes place. The reason is, of course, that nominative case valuation already took place at a previous step in the derivation. The interesting aspect of this instance of counter-bleeding is that it cannot be accounted for representationally by postulating abstract items (like traces). It therefore provides a good argument in favor of a derivational grammar.

Having presented an optimality theoretic version of the analysis proposed in Assmann et al. (2012), we are now in a position to move on to the central argument of the present paper. So far, the empirical evidence and theoretical analyses are compatible both with postulating (parametrized) preference principles like *Merge before Agree* and *Agree before Merge* (as in Assmann et al. 2012), and with postulating local optimization involving parametrized rankings of violable AC and MC constraints (as in our reconstruction in this section). In what follows, we are going to propose that the respective ranking established for ergative type languages and accusative type languages can be overwritten in particular contexts. We suggest that in both cases this happens in order to satisfy a higher ranked, more specific requirement. This is exactly what one would expect under an optimality-theoretic account, but it comes as a surprise under a preference principle-based analysis.

3 Accelerating Move: A Constraint on Dative Movement in Accusative Systems

The first argument concerns the reversal of the general preference for Agree over Move. It is based on a restriction against movement of dative arguments out of ECM-complements in German. The idea is that movement of the dative applies too early, namely before accusative case agreement can apply, thereby creating problems for the co-argument of the dative.

3.1 Data

It is a long-standing observation in the literature on German syntax that extraction of dative arguments out of ECM-complements leads to ungrammaticality (see Höhle 1978: 56-57, Thiersch 1978: 168-169, Fanselow 1986: 4, Grewendorf 1989: 150, Fanselow 1990: 121). This is illustrated for different contexts and movement types in (21)–(23).

(21) *Scrambling and pronoun movement of a DP_{dat} object from ECM complements:*

a. *dass keiner [$_{DP}$ dieser Frau]$_1$ [$_{XP}$ den Jungen t_1 helfen sah/ließ]
   that no-one$_{nom}$ this woman$_{dat}$ the boy$_{acc}$ help saw/let

b. *dass er [$_{DP}$ ihm]$_1$ [$_{XP}$ den Jungen t_1 helfen sah/ließ]
   that he$_{nom}$ him$_{dat}$ the boy$_{acc}$ help saw/let

c. *weil mir$_1$ niemand [$_{XP}$ Karl t_1 helfen ließ]
   because me$_{dat}$ no-one$_{nom}$ Karl$_{acc}$ help let

(22) *Wh-movement and topicalization of a DP_{dat} object from ECM complements:*

a. *Wem$_1$ sah/ließ Karl [$_{XP}$ den Jungen t_1 helfen]?
   whom$_{dat}$ saw/let Karl$_{nom}$ the boy$_{acc}$ help

b. *Dem Lehrer$_1$ sah/ließ Karl [$_{XP}$ den Jungen t_1 helfen]
   the teacher$_{dat}$ saw/let Karl$_{nom}$ the boy$_{acc}$ help

(23) *Movement of DP_{dat} from ECM complements with double object constructions:*

a. *Wem$_1$ ließ/sah Karl [$_{XP}$ den Jungen t_1 das Buch geben]?
   whom$_{dat}$ let/saw Karl the boy$_{acc}$ the book$_{acc}$ give

b. *dass keiner dieser Frau$_1$ [$_{XP}$ den Jungen t_1 das Buch geben]
   that no-one$_{nom}$ this woman$_{dat}$ the boy$_{acc}$ the book$_{acc}$ give
   ließ/sah
   let/saw

c. *dass er ihm$_1$ [$_{XP}$ den Jungen t_1 das Buch geben] ließ/sah
   that he$_{nom}$ him$_{dat}$ the boy$_{acc}$ the book$_{acc}$ give let/saw

It is clear that the ungrammaticality of the previous examples cannot be attributed to the extraction of DP_{dat} as such: In other contexts, movement of a dative argument produces impeccable results, see (24). (24-a) involves wh-movement from a finite clause (embedded by a bridge verb); (24-b) is a case of topicalization from a non-restructuring (i.e., fully clausal) infinitive; and (24-c) instantiates scrambling from a restructuring infinitive (that we here assume to be a vP).

(24) *Legitimate movement of DP_{dat} in other contexts*:

a. Wem$_1$ meint sie [$_{CP}$ dass wir t$_1$ das Buch geben sollten]?
whom$_{dat}$ thinks that we the book give should

b. Diesem Plan$_1$ habe ich abgelehnt [$_{CP}$ PRO t$_1$ meine Unterstützung zu geben]
this plan$_{dat}$ have I rejected my support to give

c. dass ihm$_1$ der Fritz [$_{vP}$ t$_1$ das Buch zu geben] versuchte
that him$_{dat}$ the Fritz the book to give tried

Here, we would like to put forward the hypothesis that this restriction has the same source as the ban on ergative movement in morphologically ergative languages: In all of the cases (21)–(23), the dative argument moves too early, and thus marauds the matrix v's [∗CASE:*int*∗] feature. This ultimately precludes accusative case assignment to the ECM subject, which consequently leads to a crash of the derivation.

But there is a complication: German is an accusative language and therefore exhibits the ranking AC $\gg$ MC, which would normally order case assignment of v to the embedded DP$_{ext}$ before an intermediate movement step of the dative DP to matrix Specv. In order to overcome this problem, we propose that movement of the dative DP is exceptionally *accelerated* by a higher-ranked constraint in this particular context. The constraint in question is an independently motivated one that regulates proper and improper movement.

Traditionally, the notion of improper movement is meant to cover instances of a composite movement that decomposes into smaller movements that apply in a particular order, each targeting positions of different types. A classical case is movement that first targets a SpecC-position and then a SpecT-position (called super-raising), as illustrated for English in (25-b).

(25) *Raising vs. Super-Raising in English*:

a. Mary$_1$ seems [$_{TP}$ t$_1$ to like John]

b. *Mary$_1$ seems [$_{CP}$ t$'_1$ that t$_1$ likes John]

In contrast, movement from one SpecT-position to another, as in (25-a), is unproblematic (hence an instance of proper movement). A case of improper movement from German involves movement to SpecC followed by movement to a scrambling position, presumably a specifier of vP, resulting in long-distance scrambling (26-b). In contrast, movement to an outer Specv from within the VP of the same clause is unproblematic in German (26-a).

(26) *Long-Distance Scrambling in German*:

a. dass das Buch$_1$ keiner t$_1$ liest
that the book$_{acc}$ no-one-nom reads

b. *dass Karl das Buch$_1$ glaubt [$_{CP}$ dass keiner t$_1$ liest]
that Karl$_{nom}$ the book$_{acc}$ thinks that no-one$_{nom}$ reads

Interestingly, the assumption that the PIC forces movement to SpecC to proceed via Specv creates a representation in the context of non-clause bound movement that superficially bears the hallmark of improper movement as it arises in long-distance scrambling: Compare (27-a,b), which both involve a local movement step from SpecC to Specv.

(27) *A dilemma for improper movement, given the PIC*:

a. Welches Buch_1 hat [$_{\text{vP}}$ t_1''' Karl gemeint [$_{\text{CP}}$ t_1'' dass [$_{\text{vP}}$ t_1' jeder t_1 lesen
which book_{acc} has Karl meant that everyone read
möge]]]?
should

b. *dass Karl [$_{\text{vP}}$ das Buch_1 glaubt [$_{\text{CP}}$ t_1'' dass [$_{\text{vP}}$ t_1' keiner t_1 liest]]]
that Karl_{nom} the book_{acc} thinks that no-one_{nom} reads

Thus, any theory of improper movement which exclusively concentrates on this local step faces the problem of how to account for the difference in grammaticality between (27-a) and (27-b); see Neeleman and van de Koot (2010, 346-347), Bader (2011, ch. 5), and Müller (2012, sect. 2.6.3).

In response to this problem, a new version of a standard theory of improper movement is proposed in Müller (2012). In what follows, we would like to suggest that this approach provides the constraint that is responsible for the early movement of the dative in German ECM-contexts.

3.2 Assumptions

3.2.1 Improper Movement

The assumptions about improper movement made in Müller (2012) are the following. First, when an edge feature attracts some category, it values a movement-related feature on this category. In this way, successive cyclic movement triggered by edge features creates a list on the moved item that records aspects of the derivational history of its movement. The information on the list is deleted when information of the same type is encountered in the course of the movement. Finally, there is a constraint to the effect that if the moved item reaches a criterial landing site, then the functional sequence of categories (*f-seq*: C-T-v-V) must be respected on the list containing the history of the movement steps performed by the item so far (cf. Williams 1974, 2003). This constraint will be called the WILLIAMS CYCLE (an explicit formulation of the WILLIAMS CYCLE will be given in section 3.2.2).
To illustrate the mechanics of this, consider the contrast between legitimate long *wh*-movement in (28) and illegitimate long-distance scrambling in (29).

(28) *Legitimate long-distance wh-movement*:
What_2 do you think [$_{\text{CP}}$ C [$_{\text{TP}}$ she_1 T [$_{\text{vP}}$ t_1 v [$_{\text{VP}}$ said t_2]]]]?
[$_{\text{CP}}$ $\text{what}_{[\text{wh:}\boxed{\text{CTvV}\cancel{\text{CTvV}}}]}$ [$_{\text{C}'}$ C [$_{\text{TP}}$ you think she said]]] ($\surd$f-seq)

By assumption, movement of the *wh*-phrase in (28) proceeds through the specifiers of all phrases on the path to the matrix SpecC. In the embedded clause, the *wh*-phrase collects categorial information of all intervening phrase boundaries encountered there, resulting in the partial list C-T-v-V. Movement within the matrix clause creates the same sequence on the list again, leading to successive deletion of each of the elements on the list collected in the embedded clause. When the *wh*-phrase reaches its criterial position, the matrix SpecC, the list exclusively contains the categorial information collected within the matrix clause: C-T-v-V. Since this sequence conforms with f-seq, the WILLIAMS CYCLE is satisfied.

In principle, the derivation of long-distance scrambling proceeds along the same lines. The difference, however, is that the categorial information collected in the embedded clause is not fully mirrored by the information collected in the matrix clause. As a consequence, when the criterial Specv-position is reached not all of the items on the list stemming from the embedded clause have been deleted (29), and the remaining list thus does not conform to f-seq.[9] As a consequence, the WILLIAMS CYCLE is violated and ungrammaticality results.

(29) *Illegitimate long-distance scrambling:*

*dass	Karl	das Buch	glaubt	[$_{\text{CP}}$ dass	keiner	t_1 liest]
that	Karl$_{nom}$	the book$_{acc}$	thinks	that	no-one$_{nom}$	reads

[$_{\text{CP}}$ [$_{\text{C}}$ dass] [$_{\text{TP}}$ Karl [$_{\text{T}'}$ [$_{\text{vP}}$ [$_{\text{v}'}$ [$_{\text{DP}}$ das Buch]$_{[\Sigma:\boxed{\text{vVCT}\cancel{\text{vV}}}]}$ [$_{\text{v}'}$ [$_{\text{v}'}$ glaubt dass keiner liest]]]] T]]] (*f-seq → *WILLIAMS CYCLE → crash)

3.2.2 Exceptional Case Marking in German

German ECM complements lack typical properties associated with TP (or CP) (Stechow and Sternefeld 1988, Fanselow 1991, Wurmbrand 2001). For instance, they do not host any separate temporal specification, there is obligatory wide scope for negation, and there is a systematic absence of *zu* ('to'), see (30-a-c). The conclusion from this is that XP in the examples (21)–(23) is vP.

(30) *German ECM complements are vPs*:

a. *Wir sehen [$_{\text{vP}}$ ihn den Diener erschossen haben]
we see him$_{acc}$ the servant$_{acc}$ shot have

b. Wir lassen [$_{\text{vP}}$ den Diener den Mann nicht schlagen]
we let the servant$_{acc}$ the man$_{acc}$ not hit
'We do not make/allow the servant (to) hit the man.'

c. *Wir hören [$_{\text{vP}}$ ihn zu schnarchen]
we hear him$_{acc}$ to snore

3.3 Analysis

Against this background, consider now the analysis of extraction of a dative-marked argument out of ECM complements (e.g., (21)–(23)) in German. In this configuration, matrix v has a dual role (see (31-a)): First, it assigns accusative case to the ECM subject (i.e., it bears [∗CASE:*int*∗]), and secondly, it has an (edge) feature to effect the (intermediate) movement step. Thus, the derivation faces the familiar conflict between Agree and Move on the vP cycle.

Under the accusative type ranking AC ≫ MC, this conflict is expected to be resolved by giving preference to Agree over Move. But note now that a moved dative DP originating in the embedded ECM complement (and having been assigned lexical case there by V) has a chance on the matrix vP-cycle to *immediately* remedy temporary f-seq violations on the feature list of its movement-related feature, and to thereby satisfy the WILLIAMS CYCLE (WC) quickly.[10] The specific version of WC that is required for this to happen is given in (31). It is formulated such that a temporary violation can be initiated without violating WC,

[9]It is assumed here that scrambling is ultimately triggered by a feature Σ on v, see Grewendorf and Sabel (1999), Sauerland (1999).

[10]For this to be the case, it is crucial that ECM complements in German are vPs, not TPs (or V clusters, for that matter).

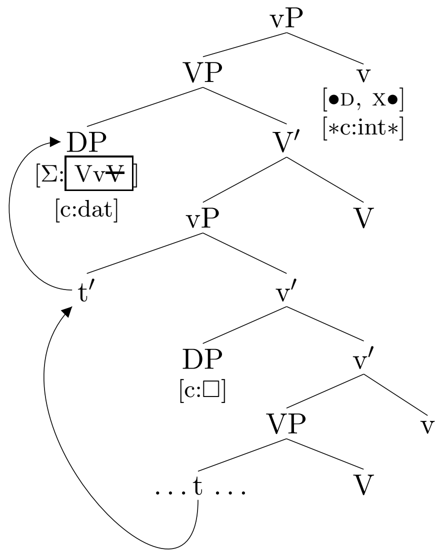
which must be possible given that WC is ranked high (above AC).[11,12]

(31) Williams Cycle (WC):
If categorial information on a list of a movement-related feature does not conform to f-seq (C-T-v-V) in the input, it must conform to f-seq in the output.

Assuming that WC in (31) outranks AC in German, movement of the dative DP to Specv will have to *precede* case assignment by v to the embedded DP_{ext} in German (32-b). The competition is shown in tableau T_{10}. Together with the SHB, this gives rise to maraudage of v's case feature [∗case:*int*∗], see (32-c) and tableau T_{11}, and the derivation will ultimately crash because the embedded DP_{ext}'s case feature remains permanently unvalued.

(32) *Illegitimate movement of DP_{dat} from ECM complements:*

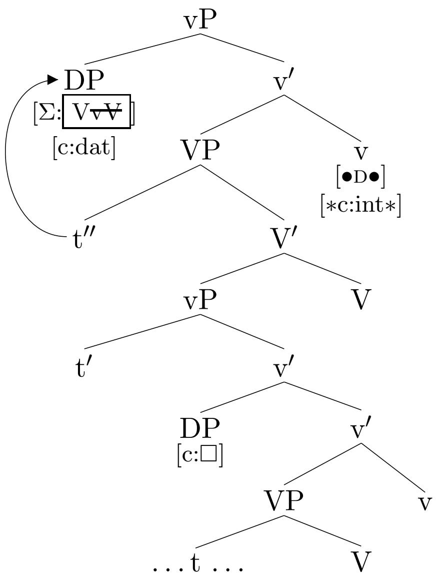
a. Structure after matrix v is merged; DP_{dat} almost satisfies f-seq



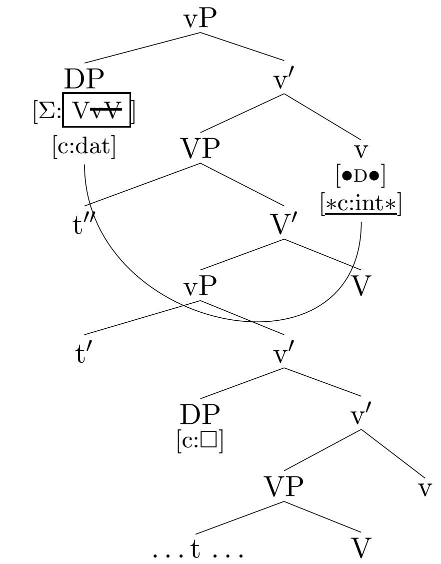
b. WC ≫ AC ≫ MC triggers movement of DP_{dat} to Specv

[11] Note that there is a certain similarity with anti-faithfulness constraints in phonology here; see Alderete (2001). Also see Baković and Wilson (2000) on targeted constraints.

[12] The reader may observe that WC is interpreted differently depending on the stage the derivation is currently in: If the moved item has reached a position that is criterial for the movement in question, a violation of WC is fatal and leads to an immediate crash. In non-criterial position, however, a violation of WC must not be fatal. Technically, the difference in interpretation could be expressed by assuming that besides WC there is also a (higher-ranked) constraint coming about via local conjunction (Smolensky 1995) of WC with a constraint making reference to the criterial/non-criterial distinction. We leave the details open here.



c. SHB triggers maraudage of v



T_{10}: Dative movement in ECM contexts, step 1: Move

Input: $\mathrm{DP}_{ext} \oplus [_{\mathrm{v}'}\ [_{\mathrm{VP}}\ \mathrm{DP}_{[\textsc{case}:\textit{dat}]}\ [_{\mathrm{vP}} \ldots \mathrm{DP}_{[\textsc{case}:\square]} \ldots \ldots \mathrm{v}\ ]]\ \mathrm{v}_{[*\textsc{case}:\textit{int}*],[\bullet \mathrm{X} \bullet],[\bullet \mathrm{D} \bullet]}]$	SHB	WC	AC	MC
☞ O_1: $[_{\mathrm{v}'}\ \mathrm{DP}_{[\textsc{case}:\textit{dat}]}\ [_{\mathrm{v}'}\ [_{\mathrm{VP}}\ \mathrm{t}''\ [_{\mathrm{vP}} \ldots \mathrm{DP}_{[\textsc{case}:\square]} \ldots \ldots \mathrm{v}\ ]]\ \mathrm{v}_{[*\textsc{case}:\textit{int}*],[\bullet \mathrm{D} \bullet]}]]$			*	*
O_2: $[_{\mathrm{v}'}\ [_{\mathrm{VP}}\ \mathrm{DP}_{[\textsc{case}:\textit{dat}]}\ [_{\mathrm{vP}} \ldots \mathrm{DP}_{[\textsc{case}:\textit{int}]} \ldots \ldots \mathrm{v}\ ]]\ \mathrm{v}_{[\bullet \mathrm{X} \bullet],[\bullet \mathrm{D} \bullet]}]$		*!		**
O_3: $[_{\mathrm{vP}}\ \mathrm{DP}_{ext}\ [_{\mathrm{v}'}\ [_{\mathrm{VP}}\ \mathrm{DP}_{[\textsc{case}:\textit{dat}]}\ [_{\mathrm{vP}} \ldots \mathrm{DP}_{[\textsc{case}:\square]} \ldots \ldots \mathrm{v}\ ]]\ \mathrm{v}_{[*\textsc{case}:\textit{int}*],[\bullet \mathrm{X} \bullet]}]]$		*!	*	*

T_{11}: Dative movement in ECM contexts, step 2: Agree (maraudage)

Input: $DP_{ext} \oplus [_{v'}\ DP_{[CASE:\mathit{dat}]}\ [_{v'}\ [_{VP}\ t''\ [_{vP} \ldots \ldots DP_{[CASE:\square]} \ldots v\ ]]\ v_{[*CASE:\mathit{int}*],[\bullet D\bullet]}]]$	SHB	WC	AC	MC
☞ O_1: $[_{v'}\ DP_{[CASE:\mathit{dat/int}]}\ [_{v'}\ [_{VP}\ t''\ [_{vP} \ldots \ldots DP_{[CASE:\square]} \ldots v\ ]]\ v_{[\bullet D\bullet]}]]$				*
O_2: $[_{v'}\ DP_{[CASE:\mathit{dat}]}\ [_{v'}\ [_{VP}\ t''\ [_{vP} \ldots \ldots DP_{[CASE:\mathit{int}]} \ldots v\ ]]\ v_{[\bullet D\bullet]}]]$	*!			*
O_3: $[_{vP}\ DP_{ext}\ [_{v'}\ DP_{[CASE:\mathit{dat}]}\ [_{v'}\ [_{VP}\ t''\ [_{vP} \ldots \ldots DP_{[CASE:\square]} \ldots v\ ]]\ v_{[*CASE:\mathit{int}*]}]]]$			*!	

3.4 Extremely Local vs. Less Local Optimization

This time, assuming larger optimization domains like the phrase does not make a wrong prediction: At the vP phrase level, WC, MC and AC are all satisfied, and SHB will continue to pick a maraudage output. Of course, the analysis is nevertheless also compatible with an approach where optimization applies at the step-level, as illustrated.

3.5 Consequences

The approach makes at least one interesting additional prediction, but it also raises various questions. In what follows, these issuess are briefly addressed.

First, if there is no embedded DP_{ext} in what is otherwise the same construction, then the prediction is that movement of the dative DP should be fine because there is no external co-argument DP that could violate the case filter after early (WC-driven) movement of the dative DP from the ECM infinitive to the matrix Specv position. This prediction is borne out. Consider the so-called *lassen*-passive construction in (33), where the external argument of the embedded infinitive is demoted exactly as in standard passive constructions (including the option of realizing it as a PP, not indicated here) even though no morphological reflex of passive is present; see Höhle (1978), among many others.

(33) *DP_{dat} movement where an embedded DP_{ext} is not present:*

a. dass keiner [$_{DP}$ dieser Frau]$_1$ gestern/gerne [$_{XP}$ t_1 helfen ließ]
   that no-one$_{nom}$ this woman$_{dat}$ yesterday/gladly help let

b. dass er [$_{DP}$ ihm]$_1$ gestern/ungern [$_{XP}$ t_1 helfen ließ]
   that he$_{nom}$ him$_{dat}$ yesterday/reluctantly help let

c. Wem$_1$ ließ Karl gestern/ungern [$_{XP}$ t_1 helfen] ?
   whom$_{dat}$ let Karl$_{nom}$ yesterday/reluctantly help

d. Dem Lehrer$_1$ ließ Karl gestern/ungern [$_{XP}$ t_1 helfen]
   the teacher$_{dat}$ let Karl$_{nom}$ yesterday/reluctantly help

This effect is fully parallel to the one identified in Assmann et al. (2012) with respect to legitimate ergative movement in the absence of an internal argument DP that requires structural (i.e., absolutive) case.

Second, there is the question as to why extraction of dative DPs becomes possible again if the predicate of the ECM complement is an unaccusative verb. The problem here is that the embedded DP$_{int}$ gets case from the matrix v and does not block dative movement; see (34) (from Fanselow 1990).

(34) *DP$_{dat}$ movement is fine in unaccusative contexts:*

a. *dass mir$_1$ niemand [$_{XP}$ Karl t$_1$ helfen] ließ/saw
that me$_{dat}$ no-one$_{nom}$ Karl$_{dat}$ help let/saw

b. dass mir$_1$ niemand [$_{XP}$ t$_1$ ein Unglück zustoßen] ließ/sah
that me$_{dat}$ no-one$_{nom}$ an accident happen to let/saw

This problem is solved if there is no vP with unaccusative predicates, pace Legate (2003). Under this assumption, WC does not force early DP$_{dat}$ movement because there is no improper f-seq when DP$_{dat}$ enters the matrix VP domain; hence, there is no maraudage.

Third, one may wonder what happens if DPs with other cases undergo extraction from ECM complements. It turns out that DP$_{acc}$ can undergo such movement easily (see (35), (36)). In contrast, DP$_{gen}$ movement is arguably much more restricted, see (37).

(35) *Scrambling of a DP$_{acc}$ object from ECM complements:*

a. dass der Kollege [$_{DP}$ den Antrag]$_1$ [$_{XP}$ seine Mitarbeiter t$_1$ gerade schreiben lässt]
that the colleague$_{nom}$ the proposal$_{acc}$ his co-workers$_{acc}$ currently write lets

b. dass [$_{DP}$ den Antrag]$_1$ der Kollege [$_{XP}$ seine Mitarbeiter t$_1$ gerade schreiben lässt]
that the proposal$_{acc}$ the colleague$_{nom}$ his co-workers$_{acc}$ currently write lets

(36) *Pronoun movement of a DP$_{acc}$ object from ECM complements:*

a. dass er es$_1$ [$_{XP}$ den Jungen t$_1$ lesen sah]
that he$_{nom}$ it$_{acc}$ the boy$_{acc}$ read saw

b. dass er es$_1$ [$_{XP}$ den Jungen t$_1$ machen ließ]
that he$_{nom}$ it$_{acc}$ the boy$_{acc}$ make let

(37) *Movement of a DP$_{gen}$ object from ECM complements.*

a. Karl sieht/lässt den Jungen der Toten gedenken
Karl$_{nom}$ sees/lets the boy$_{acc}$ the dead$_{gen}$ commemorate

b. ?*dass derer/der Toten keiner den Jungen gedenken sieht/lässt
that they$_{gen}$/the dead$_{gen}$ no-one$_{nom}$ the boy$_{acc}$ commemorate sees/lets

c. ?*Der Toten$_1$ sieht/lässt Karl den Jungen gedenken
the dead$_{gen}$ sees/lets Karl$_{nom}$ the boy$_{acc}$ commemorate

This would follow without further ado if maraudage is blocked if exactly the same case is involved; and genitive and accusative are sufficiently different.

4 Decelerating Move: Mobility of Lexical/Oblique Arguments in Ergative Systems

The second argument deals with a reversal of the general preference for Move over Agree. Starting point is the observation that not all morphologically ergative languages exhibit the ban against extraction of the ergative subject.

4.1 Data

In some morphologically ergative languages, the ergative extracts freely and without any special morphology (such as the agent focus morphology encountered in many Mayan languages). Below, this is illustrated for Chol (Mayan). Other ergative languages that exhibit this pattern are Basque (isolate; Hualde and Ortiz de Urbina 2003), Avar (Nakh-Dagestanian; Polinsky et al. 2011), and Pitjantjatjara (Pama-Nyungan; Bowe 1990: 101).

(38) *Wh-movement of DP_{erg} in Chol* (Coon 2010: 226, Coon et al. 2011):

a. Maxki$_{1/2}$ tyi y-il-ä (t$_1$) aj-Maria (t$_2$)?
   who ASP A3-see-DTV DET-Maria
   'Who saw Maria?' / 'Who did Maria see?'

b. Maxki$_1$ tyi y-il-ä t$_1$ a-wakax?
   who PRFV A 3-see-DTV A 2-cow
   'Who saw your cow?'

The idea of the analysis will be that extraction of the ergative subject is an option in these languages because Agree exceptionally applies before Merge on the TP-cycle despite the ergative ranking MC $\gg$ AC. This will be analyzed as there being a higher ranked constraint which, in this particular context, procrastinates Merge in favor of Agree. (In ergative languages that exhibit the ban on moving the ergative argument this higher ranked constraint is vacuously fulfilled, giving rise to emergence of the unmarked.)

4.2 Assumptions

We propose that the possibility of moving the ergative argument in these languages is due the nature of the ergative case involved. More precisely, we would like to suggest that ergative case in Chol, Avar, Basque, and Pitjantjatjara is not structural but lexical (see Nash 1996, Alexiadou 2001, Woolford 1997, 2001, Legate 2008 for related claims). However, our overall argument here presupposes the theory of argument encoding put forward in Müller (2009), which is based on the idea that ergative type encoding systems involve a marked case [∗CASE:*int*∗] on v, which is spelled out as ergative case.

In order to reconcile these two views, we make the following assumptions. The case probe [∗CASE:*int*∗] representing internal structural case is in fact composed of the two subfeatures [–OBL] and [+GOV]. (Similarly, [∗CASE:*ext*∗] is actually composed of [–OBL] and [–GOV].) Here, [±GOV] maintains the external/internal distinction, and [–OBL] indicates that the cases associated with T and v are structural (non-oblique). Such a decomposition of case features is first and foremost motivated by morphological considerations relating to syncretism: This way, natural classes of cases can be defined by referring to underspecified case information on morphological case exponents (e.g., [–OBL] captures the natural class of structural cases

– nominative and accusative in accusative systems, and absolutive and ergative in ergative systems; [+GOV] captures the natural class of accusative/ergative, dative, and other governed cases; and so on); cf. Bierwisch (1967), Wiese (1999), and much recent work in Distributed Morphology.

In what follows, we will make use of case decomposition in the syntax. Accordingly, DP arguments bear unvalued variants of these subfeatures: [OBL:□] and [GOV:□]. A DP is valued with ergative (or, for that matter, accusative) case if its case subfeatures are valued [CASE:–OBL,+GOV]. In morphologically ergative languages with a *structural* ergative, these two case subfeatures are located on v. From there, they compositionally value case on DP_{ext}. Thus, here everything still works exactly as laid out above – the fine structure of the case feature may be relevant in morphology, but is in fact invisible in syntax. For languages with a *lexical* ergative, we assume that v only bears [–OBL] while [+GOV] is located on V; i.e., internal case is split between a [∗CASE:–OBL∗] probe on v and a [∗CASE:+GOV∗] probe on V. This reflects the hypothesis that lexical ergative is assigned by V in interaction with v. The compositional assignment of the two a priori separated probes comes about via V-to-v movement. Note that we take it that any case probe (including a composite v-V probe) cannot distribute its decomposed feature values over different goals.

In addition, we assume that the two subfeatures [–OBL] and [+GOV] involved in lexical case assignment differ with respect to the structural conditions they require for entering into Agree: For [–OBL], m-command is sufficient (see (8)); but the [+GOV] feature on V that makes the composite case lexical is discharged under a stricter locality condition: It must *c-command* the goal that it is supposed to establish Agree with. This corresponds to the observation that "pure" lexical case assignment (e.g., a lexically assigned genitive in German) typically ends up on the lowest argument DP of a predicate (see Fanselow 2001).[13]

Next, we propose that arguments with partially valued case are inactive in the sense that they are invisible for structure-building features triggering Merge (cf. Richards 2008; also cf. Chomsky's 2001 *Activity Condition*). As a consequence, inactive elements cannot undergo movement. This is explicitly stated by the constraint in (39).

(39) ACTIVITY CONDITION (ACTC):
Inactive elements cannot undergo movement.

Thus, DPs that have at least one but not yet all of their case features valued must first finish valuation before they can undergo movement. In contrast, DPs that have their case features completey valued – or completely unvalued – are active.[14] The guiding hypothesis here is that there is a general *contiguity* requirement for syntactic operations: An operation consisting of several subparts must be fully completed once it has begun before the affected item can be accessed by other operations (i.e., qualify as active).

Finally, given that lexical (i.e., V-based) [+GOV] on V can only be assigned under c-command, and given that v-V does not c-command Specv, something needs to be said about

[13]No such strict c-command requirement holds for [+GOV] on v, where it is not a lexical case feature. This presupposes that [+GOV] on v (structural case) and [+GOV] on V (lexical case) can be distinguished accordingly. Given the minimal contextual difference (part of v vs. part of V), this would seem to be unproblematic.

[14]There is evidence suggesting that a DP that has not received any case value so far must not qualify as inactive under these assumptions, and can accordingly undergo movement. First, this is required by classical approaches to case-driven raising; second, it is in fact required for the derivation of constructions involving absolutive movement in ergative systems under present assumptions; recall the derivation in (18).

how [+GOV] can eventually be assigned to an external argument DP in Specv. The obvious conclusion would seem to be that this is effected by v-V-to-T movement: T c-commands Specv. We will assume that this is indeed correct. If so, there are basically two options how the required head movement operation can be brought about. First, it might be that v-V-to-T movement must independently exist in a language to make lexical ergative assignment possible; this would imply that a lexical ergative results from the conspiracy of two independent parameter settings (viz., (a) [+GOV] on V (not v) and (b) obligatory v-V-to-T movement). Second, it might be that the need for V to get rid of its lexical case subfeature [+GOV] may directly trigger v-V-to-T movement, as an instance of repair-driven movement (see Heck and Müller 2000), and irrespective of any general parameter setting for the head movement operation. Since nothing hinges on this, we will not choose between the two options in what follows; we tentatively adopt the first option for reasons of exposition alone.[15]

4.3 Analysis

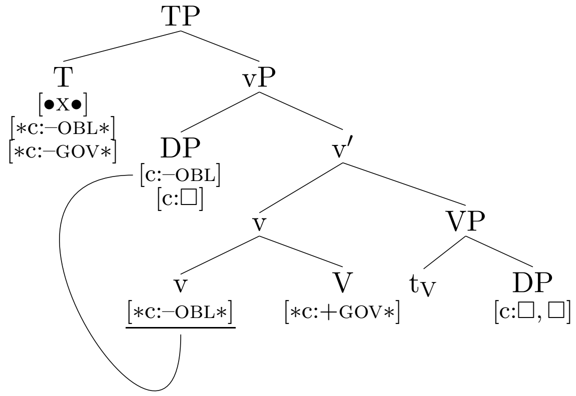
Imagine a scenario where DP_{ext} is supposed to undergo Ā-movement in an ergative system where the ergative is lexical. Given MC ≫ AC, DP_{ext} is merged before v can trigger Agree; the same ranking may also be assumed to trigger V-to-v movement early. After being merged in Specv, DP_{ext}'s case feature is partially valued by [*CASE:–OBL*] on v (due to the SHB), yielding [CASE:–OBL,□]. However, DP_{ext} is not in the c-command domain of v-V, so the remaining (lexical case) probe [*CASE:+GOV*] cannot participate in Agree at this point (recall that distributing decomposed feature values over different goals is barred; thus [*CASE:+GOV*] cannot enter into Agree with DP_{int} either). As a consequence, DP_{ext} is inactive when T is merged; see (40-a). Due to the ranking ActC ≫ MC ≫ AC, the inactive DP_{ext} now cannot immediately move to SpecT once T has been introduced into the structure, despite the presence of the "ergative" ranking MC ≫ AC (DP_{int} cannot move because it is not designated for Ā-movement). However, this ranking successfully triggers head movement of v-V to T; see (40-b). At this point, there are two case probes in v-V-T: On the one hand, there is [*CASE:–OBL,–GOV*] (i.e., [*CASE:*ext**]) on T, and on the other hand, there is the partial probe [*CASE:+GOV*] on V. There are four possibilities as to what can happen next. T's case probe may undergo Agree with DP_{int}, T's case probe may undergo Agree with DP_{ext}, V's partial case probe may undergo Agree with DP_{int}, or V's partial case probe may undergo Agree with DP_{ext}. The present system does not distinguish between these four options; they all satisfy ActC, violate MC once, and violate AC once. (In contrast, movement of DP_{ext} would fatally violate ActC.) Suppose that the first option is chosen: [*CASE:–OBL,–GOV*] on T undergoes Agree with DP_{int}, valuing the latter's case feature, as in (40-c). Then, in the next step, [*CASE:+GOV*] on V values DP_{ext}, which must still be in situ, given ActC; see (40-d). However, after this final Agree operation, DP_{ext} is active again, and it can and must finally undergo the intermediate movement step to SpecT; cf. (40-e). Of course, this means that movement of DP_{ext} comes too late to maraude T's absolutive case feature for DP_{int}.

[15] Another issue that must be clarified in this context but is orthogonal to our main concerns is how head movement of V to v, and subsequently of v to T, can result in proper c-command by V (of Specv, as required for lexical case valuation, but also of its own trace). A standard solution to this problem is to minimally relax the locality condition on c-command, such that if a head α is adjoined to another head β, α c-commands whatever β c-commands (see Baker 1988). Alternatively, following Roberts (2010), we may assume that complete copying of a feature set derives the effects of head movement (without actual movement taking place). The copying operation would then also comprise the case feature [+GOV] on V.

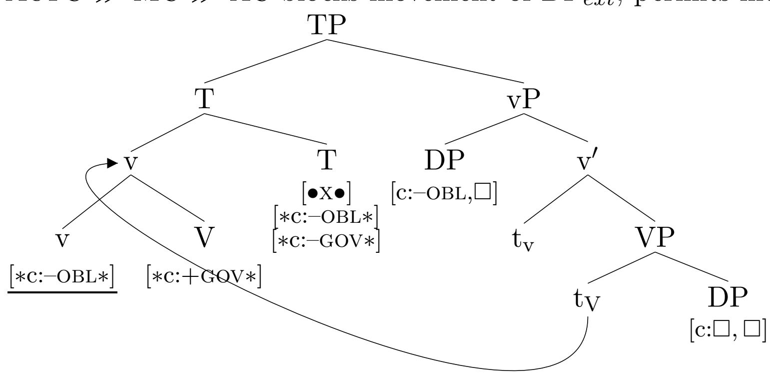
Hence, extraction of the lexically ergative-marked DP_{ext} does not lead to ungrammaticality.

(40) *Legitimate movement of DP_{erg} if the ergative is lexical*

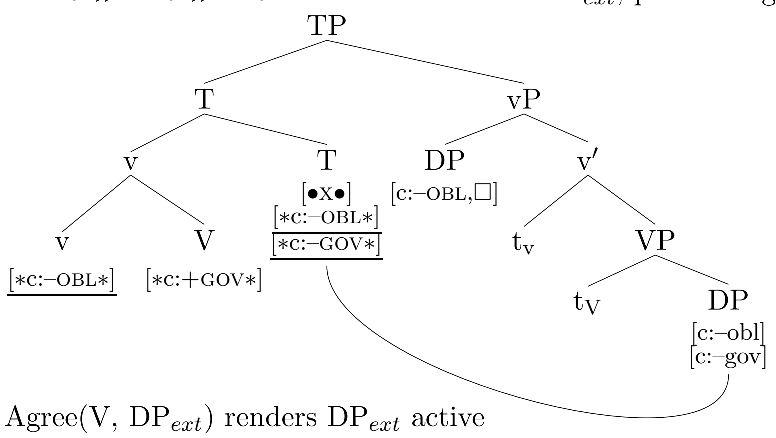
a. Structure after T is merged; partial case assignment to DP_{ext}



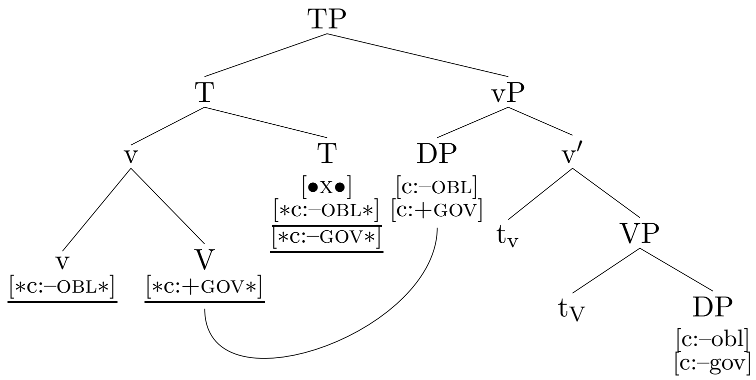
b. ActC ≫ MC ≫ AC blocks movement of DP_{ext}; permits movement of v-V



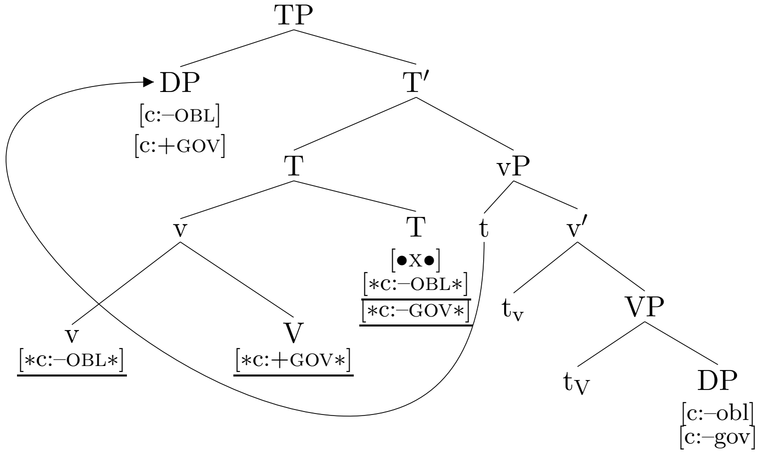
c. ActC ≫ MC ≫ AC blocks movement of DP_{ext}; permits Agree(T, DP_{int})



d. Agree(V, DP_{ext}) renders DP_{ext} active



e. DP_{ext} finally moves but comes too late to effect maraudage



The relevant competitions are given in tableaux T_{12}, T_{13}, and T_{14}. Note that there are four locally optimal continuations O_2–O_5 (that all carry out an Agree operation) in T_{12} (which illustrates the crucial step from (40-b) to (40-c)), in addition to O_1, which executes movement of DP_{ext} and thereby fatally violates ActC. However, of these four optimal outputs only O_2 (where T undergoes Agree with DP_{int}) will eventually lead to a well-formed output: In O_3 and O_5, DP_{ext} gets its case valued (by T and V, respectively), which means that it becomes active and will have to move in the next step, thereby marauding case features required for DP_{int}. Similarly, O_4 will invariably lead to a crash because DP_{int} undergoes Agree with V here, and will therefore never acquire a fully specified case feature. (Alternatively, DP_{ext} will fail to do so if DP_{int} receives T's features as well; again, recall that a case probe cannot distribute its decomposed feature values over different goals.)

Assuming a continuation with O_2, tableau T_{13} shows that the situation is still such that DP_{ext} cannot move without fatally violating ActC.

Finally, tableau T_{14} illustrates the trivial final competition on the TP cycle: DP_{ext} is now active, and movement can finally be carried out.

4.4 Extremely Local vs. Less Local Optimization

As with the very option of accusative movement under present assumptions, an argument for extremely local optimization emerges in the case of decelerating ergative movement if the ergative is lexical. If the whole TP (or an even larger domain) is considered, the ban

T_{12}: Lexical ergative movement, step 1: Agree(T,DP$_{int}$)

Input: $[_{\text{TP}}$ v-V$_{[*\text{C:+GOV}*]}$-T$_{[*\text{C:–OBL,–GOV}*],[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL},\square]}$ ... DP$_{[\text{C:}\square,\square]}$]]	ActC	MC	AC
O$_1$: $[_{\text{TP}}$ DP$_{[\text{C:–OBL},\square]}$ v-V$_{[*\text{C:+GOV}*]}$-T$_{[*\text{C:–OBL,–GOV}*]}$ $[_{\text{vP}}$ t ... DP$_{[\text{C:}\square]}$]]	*!		**
☞ O$_2$: $[_{\text{TP}}$ v-V$_{[*\text{C:+GOV}*]}$-T$_{[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL},\square]}$ ... DP$_{[\text{C:–OBL,–GOV}]}$]]		*	*
☞ O$_3$: $[_{\text{TP}}$ v-V$_{[*\text{C:+GOV}*]}$-T$_{[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL,–OBL,–GOV}]}$ ... DP$_{[\text{C:}\square,\square]}$]]		*	*
☞ O$_4$: $[_{\text{TP}}$ v-V-T$_{[*\text{C:–OBL,–GOV}*],[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL},\square]}$ ... DP$_{[\text{C:}\square,\text{+GOV}]}$]]		*	*
☞ O$_5$: $[_{\text{TP}}$ v-V-T$_{[*\text{C:–OBL,–GOV}*],[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL,+GOV}]}$ ... DP$_{[\text{C:}\square,\square]}$]]		*	*

T_{13}: Lexical ergative movement, step 2: Agree(V,DP$_{ext}$)

Input: $[_{\text{TP}}$ v-V$_{[*\text{C:+GOV}*]}$-T$_{[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL},\square]}$ ... DP$_{[\text{C:–OBL,–GOV}]}$]]	ActC	MC	AC
O$_1$: $[_{\text{TP}}$ DP$_{[\text{C:–OBL},\square]}$ v-V$_{[*\text{C:+GOV}*]}$-T $[_{\text{vP}}$ t ... DP$_{[\text{C:–OBL,–GOV}]}$]]	*!		*
☞O$_2$: $[_{\text{TP}}$ v-V-T$_{[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL,+GOV}]}$ ... DP$_{[\text{C:–OBL,–GOV}]}$]]		*	

T_{14}: Lexical ergative movement, step 3: Move(T,DP$_{ext}$)

Input: $[_{\text{TP}}$ v-V-T$_{[\bullet\text{X}\bullet]}$ $[_{\text{vP}}$ DP$_{[\text{C:–OBL,+GOV}]}$ ... DP$_{[\text{C:–OBL,–GOV}]}$]]	ActC	MC	AC
☞O$_1$: $[_{\text{TP}}$ DP$_{[\text{C:–OBL,+GOV}]}$ v-V-T $[_{\text{vP}}$ t ... DP$_{[\text{C:–OBL,–GOV}]}$]]			

on ergative movement that follows from the ranking MC $\gg$ AC cannot be circumvented anymore, although this time for a different reason. Because of ActC, the optimal TP cannot host DP$_{ext}$ in its Spec. Rather, the optimal TP candidate is one where DP$_{ext}$ receives its second case value in Specv. At this point, DP$_{ext}$ becomes active; but since the TP-cycle is over now, movement to SpecT cannot apply. On the CP-cycle, movement to SpecT is banned by the Strict Cycle Condition (Chomsky 1973) and movement to SpecC in one fell swoop is impossible due to the PIC. The derivation therefore crashes due to an unchecked feature on C. This is shown in tableau T_{15}.

5 Conclusion and Outlook

In this paper, we have tried to defend four related claims.

First, it seems to be a fact that given standard minimalist assumptions about structure-

T_{15}: TP optimization under ACTC $\gg$ *MC* $\gg$ *AC ranking: wrong result*

Input: $T_{[*C:-OBL,-GOV*],[\bullet X\bullet]} \oplus$ $[_{vP}\ DP_{[C:-OBL,\Box]}\ v\text{-}V_{[*C:+GOV*]} \ldots DP_{[C:\Box,\Box]}\ ]]$	ACTC	SHB	MC	AC
O_1: $[_{TP}\ DP_{[C:-OBL,+GOV]}\ v\text{-}V\text{-}T$ $[_{vP}\ t \ldots DP_{[C:-OBL,-GOV]}\ ]]$	*!	*		
O_2: $[_{TP}\ DP_{[C:-OBL,+GOV,-OBL,-GOV]}\ v\text{-}V\text{-}T$ $[_{vP}\ t \ldots DP_{[C:\Box,\Box]}\ ]]$	*!			
☛ O_3: $[_{TP} - v\text{-}V\text{-}T\ [_{vP}\ DP_{[C:-OBL,+GOV]}$ $\ldots DP_{[C:-OBL,-GOV]}\ ]]$			*	

building, competition between Merge (Move) and Agree can arise in the derivation. In particular, such a situation occurs when the head of a phrase has to carry out more than one operation. On the vP cycle, this is the case with a v that introduces an external argument DP and assigns structural case (v[*CASE:*int**],[•D•]); on the TP cycle this is the case with a T that triggers an intermediate movement step via an edge feature and also assigns structural case (T[*CASE:*ext**],[•X•]).

Second, the conflicts that arise between Merge (Move) and Agree can and must be resolved in one way or the other. We have argued that there may be no intrinsic, fixed way of resolution; rather, the empirical evidence suggests that how conflicts are resolved is a matter of parametrization. This can be implemented either by invoking parametrized preference principles (of the type of the Merge before Move principle in Chomsky 1995, 2001); or by postulating constraint ranking. Assuming (for concreteness) the latter, we have seen that MC $\gg$ AC on the vP cycle gives rise to an ergative encoding system whereas the reverse resolution strategy following from AC $\gg$ MC on the vP cycle predicts an accusative encoding system; and that a ranking MC $\gg$ AC on the TP cycle accounts for the immobility of DPs bearing structural ergative case (because these items move too early, bringing about maraudage of T's case feature) whereas a reverse ranking AC $\gg$ MC on the TP cycle correctly predicts movement of DPs bearing structural accusative case to be possible (because these items move late, thereby avoiding maraudage of T's case features).

Third, and most importantly, we have argued that the ban on dative movement from ECM contexts in German, and the option of lexical ergative movement in Chol, Avar, Basque, and Pitjantjatjara, can be accounted for straightforwardly if it is assumed that there can be an acceleration of movement in accusative systems that normally give preference to Agree over Merge (Move) in the case of conflict if this is forced by an independent factor; and that there can also be a deceleration of movement in ergative systems that normally give preference to Merge (Move) over Agree in the case of conflict if this is forced by an independent factor. These situations are fully expected under an optimality-theoretic approach, but less so under a more orthodox minimalist approach employing (parametrized) preference principles. Thus, a ranking WC $\gg$ AC $\gg$ MC on the vP cycle correctly predicts dative movement from German ECM constructions to be impossible (because it comes too early), and a ranking ACTC $\gg$ MC $\gg$ AC on the TP cycle correctly predicts lexically marked ergative subjects to be mobile in Chol, Avar, Basque, and Pitjantjatjara. In contrast, a modified preference principle like "Agree before Move unless satisfaction of the Williams Cycle demands otherwise" does not per se look like a plausible candidate for a constraint of grammar; and the same goes for a modified preference principle like "Move before Agree unless satisfaction of the Activity Condition demands otherwise".[16]

Fourth and finally, the analyses presented in this paper provide evidence for extremely local serial optimization in syntax, and against less local optimization procedures (including ones where the whole sentence is subject to a single, parallel optimization): If the domain is larger than the derivational step, then (i) AC $\gg$ MC on the vP cycle does not derive accusative encoding systems; (ii) AC $\gg$ MC on the TP cycle wrongly blocks accusative movement; and (iii) ActC $\gg$ MC $\gg$ AC on the TP cycle cannot circumvent the ban on ergative movement. The reasonings for (i) and (ii) rely on standard arguments based on opacity of rule interaction in generative grammar.

From a more general point of view, the present study can be seen as an attempt to sketch the outlines of a new approach to an empirical domain that received a lot of attention in earlier work in the Principles and Parameters tradition but has arguably been given much less attention in more recent minimalist approaches, viz., asymmetries between types of categories with respect to their extractability. It has often been observed that some kinds of linguistic expressions are less mobile than others in the sense that they may not cross domains that are transparent for other items. Such asymmetries have been noted for objects vs. subjects, for arguments vs. adjuncts, for referential vs. non-referential phrases, for items that have an "address" vs. others that don't (see Manzini 1992), and so on. Standardly, these kinds of asymmetries were captured by imposing appropriate constraints on empty categories that are assumed to be left behind by displacement operations (cf., e.g., Chomsky's 1981 Empty Category Principle (ECP) for traces, or the different constraints for trace vs. *pro* in Cinque 1990). However, such options do not exist anymore under minimalist assumptions according to which all constraints are either principles of efficient computation or imposed by the interfaces (see Chomsky 2001, 2008). Furthermore, traces – as special items enriching the syntactic ontology for which designated constraints can be formulated – have come to be widely regarded as suspect from a minimalist viewpoint.[17]

Taken together, this means that there is a gap in current minimalist approaches to syntax: It is a priori unclear how asymmetries between moved items can be accounted for. The present approach can be viewed as a program for filling this gap. The basic premise is that if some items are less mobile than others, this must be so because their movement may lead to problems elsewhere (i.e., in domains not directly related to the movement operation), either for themselves or for other items in the clause. We have argued that movement of certain items (α) may create problems for other, sufficiently similar items (β). Thus, by pursuing this program, we end up with a *relational*, co-argument-based approach to displacement (α cannot move in the presence of β because α-movement creates problems for β-licensing) of the type that has sometimes been suggested for case assignment (α is assigned x-case in the presence of β; see Marantz 1991, Bittner and Hale 1996, Wunderlich 1997, Stiebels 2000, McFadden 2004). More specifically, a common pattern emerges that captures the legitimate

[16] That said, it might eventually not be impossible to save the preference principle-based approach, by postulating that only *convergent* steps are considered, and further assuming that violations of WC and ActC lead to non-convergence. It is far from obvious, however, that a simple notion of convergence can be devised that covers all relevant contexts in a natural way; see Sternefeld (1996) for related discussion.

[17] This may be so because displacement does not leave a reflex in the original position to begin with; see Epstein and Seely (2002), Unger (2010), Müller (2011) for some options; or because a multidominance approach is adopted; see Gärtner (2002), Starke (2001), Abels (2004), Frampton (2004), among others.

and illegitimate instances of movement discussed in the present paper: On the one hand, movement of some category α that takes place early on a given cycle brings with it the danger of maraudage of features that would be needed for the licensing of some other category β, and may thereby lead to ungrammaticality; this holds for DPs that bear structural ergative case and for dative DPs that have a chance to immediately remedy a temporary improper movement configuration.[18] On the other hand, movement of some category α that takes place late on a given cycle will more likely be able to circumvent maraudage effects for some other category β, and will therefore more often lead to grammaticality; this holds for DPs that bear structural accusative case and for ergative DPs where the ergative is lexical and the DP in question is therefore not yet active (hence, not yet accessible by movement).[19] Overall, then, a simple generalization emerges:

Good things come to those who wait.

References

Abels, Klaus (2003): Successive Cyclicity, Anti-Locality, and Adposition Stranding. PhD thesis, University of Connecticut, Storrs, Connecticut.

Abels, Klaus (2004): Right Node Raising: Ellipsis or Across the Board Movement?. *In:* K. Moulton and M. Wolf, eds, *Proceedings of NELS 34*. GLSA, Amherst, Mass., pp. 45–60.

Adger, David and Daniel Harbour (2007): 'Syntax and Syncretisms of the Person Case Constraint', *Syntax* **10**, 2–37.

Alderete, John (2001): *Morphologically Governed Accent in Optimality Theory*. Routledge, New York.

Aldridge, Edith (2004): Ergativity and Word Order in Austronesian Languages. PhD thesis, Cornell University.

Alexiadou, Artemis (2001): *Functional Structure in Nominals*. Benjamins, Amsterdam.

Anagnostopoulou, Elena (2005): Strong and Weak Person Restrictions. A Feature Checking Analysis. *In:* L. Heggie and F. Ordoñez, eds, *Clitics and Affix Combinations*. Benjamins, Amsterdam, pp. 199–235.

Andrews, Avery (1996): 'Semantic Case-Stacking and Inside-Out Unification', *Australian Journal of Linguistics* **16**, 1–54.

[18]This latter case shows inherent myopia (see Collins 1997), i.e., the strict lack of look-ahead of the current approach: By trying to correct a problem that can in principle be tolerated (viz., a temporary improper movement configuration), much greater damage is done, and the derivation eventually crashes.

[19]Needless to say, the new approach to asymmetries between moved items raises many new questions that will have to be addressed before it can be regarded as viable. From an empirical point of view, obvious challenges that need to be faced include *that*-trace effects, Left Branch Condition effects (i.e., the immobility of pre-nominal genitive DPs), and the stronger restrictions on adjunct movement. From a conceptual point of view, the availability of re-ranking among the relevant constrains remains an open issue; for instance, the fact that ActC and WC outrank the more basic constraints AC and MC in all cases addressed above does not seem to be accidental – but if it isn't, the task is to make this follow in some principled way that is compatible with basic optimality-theoretic tenets.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller and Philipp Weisser (2012): Ergatives Move Too Early. To appear in *Syntax*.

Assmann, Anke, Svetlana Edygarova, Doreen georgi, Timo Klein and Philipp Weisser (2013): Possessor Case in Udmurt: A Local Reanalysis as Fusional Case Stacking. To appear in *The Linguistic Review*.

Bader, Monika (2011): Constraining Operations. A Phase-Based Account of Improper Movement and Anaphoric Binding. PhD thesis, CASTL, Tromsø University., Tromsø.

Baker, Mark (1988): *Incorporation. A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.

Baković, Eric and Colin Wilson (2000): Transparency, Strict Locality, and Targeted Constraints. *In:* R. Billerey and B. D. Lillehaugen, eds, *WCCFL 19 Proceedings*. Cascadilla Press, Somerville, MA, pp. 43–56.

Béjar, Susana and Milan Řezáč (2009): 'Cyclic Agree', *Linguistic Inquiry* **40**, 35–73.

Bierwisch, Manfred (1967): Syntactic Features in Morphology: General Problems of So-Called Pronominal Inflection in German. In: *To Honor Roman Jakobson*. Mouton, The Hague/Paris, pp. 239–270.

Bittner, Maria and Ken Hale (1996): 'The Structural Determination of Case and Agreement', *Linguistic Inquiry* pp. 1–68.

Boeckx, Cedric (2003): *Islands and Chains. Resumption as Stranding*. Benjamins, Amsterdam.

Bošković, Željko (2007): 'Agree, Phases, and Intervention Effects', *Linguistic Analysis* **33**, 54–96.

Bowe, Heather (1990): *Categories, Constituents and Constituent Order in Pitjantjatjara: An Aboriginal Language of Australia*. Routledge, New York.

Campana, Mark (1992): A Movement Theory of Ergativity. PhD thesis, McGill University.

Chomsky, Noam (1951): Morphophonemics of Modern Hebrew. Master's thesis, University of Pennsylvania.

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

Chomsky, Noam (1975): *The Logical Structure of Linguistic Theory*. Plenum Press, New York.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam (1986*a*): *Barriers*. MIT Press, Cambridge, Mass.

Chomsky, Noam (1986*b*): *Knowledge of Language*. Praeger, New York.

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

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. *In:* R. Martin, D. Michaels and J. Uriagereka, eds, *Step by Step*. MIT Press, Cambridge, Mass., pp. 89–155.

Chomsky, Noam (2001): Derivation by Phase. *In:* M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2005): 'Three Factors in Language Design', *Linguistic Inquiry* **36**, 1–22.

Chomsky, Noam (2007): Approaching UG from Below. *In:* U. Sauerland and H.-M. Gärtner, eds, *Interfaces + Recursion = Language?*. Mouton de Gruyter, Berlin, pp. 1–31.

Chomsky, Noam (2008): On Phases. *In:* R. Freidin, C. Otero and M. L. Zubizarreta, eds, *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.

Chomsky, Noam (2013): 'Problems of Projection', *Lingua* **130**, 33–49.

Cinque, Guglielmo (1990): *Types of A-bar Dependencies*. MIT Press, Cambridge, Mass.

Collins, Chris (1997): *Local Economy*. MIT Press, Cambridge, Mass.

Coon, Jessica (2010): Complementation in Chol (Mayan). A Theory of Split Ergativity. PhD thesis, MIT, Cambridge, Mass.

Coon, Jessica, Pedro Mateo Pedro and Omer Preminger (2011): The Role of Case in A-Bar Extraction Asymmetries: Evidence from Mayan. Ms., McGill University.

Deal, Amy Rose (2012): Ergativity. Ms., University of California, Santa Cruz. To appear in *Syntax. An International Handbook of Contemporary Research*, ed. Tibor Kiss & Artemis Alexiadou. Berlin: De Gruyter.

Epstein, Samuel David and T. Daniel Seely (2002): Rule Applications as Cycles in a Level-Free Syntax. *In:* S. D. Epstein and T. D. Seely, eds, *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 65–89.

Fanselow, Gisbert (1986): Scrambling, Anaphern und Barrieren. GGS-Vortrag, Universität Regensburg.

Fanselow, Gisbert (1990): Scrambling as NP-Movement. *In:* G. Grewendorf and W. Sternefeld, eds, *Scrambling and Barriers*. Benjamins, Amsterdam, pp. 113–140.

Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.

Fanselow, Gisbert (2001): Optimal Exceptions. *In:* B. Stiebels and D. Wunderlich, eds, *The Lexicon in Focus*. Akademie Verlag, Berlin, pp. 173–209.

Ferguson, K. Scott (1993): 'Notes on the Shortest Move Metric and Object Checking', *Harvard Working Papers in Linguistics* **3**, 65–80.

Frampton, John (2004): Copies, Traces, Occurrences, and All That. Ms., Northeastern University.

Gärtner, Hans-Martin (2002): *Generalized Transformations and Beyond*. Akademie-Verlag, Berlin.

Georgi, Doreen (2013): Opaque Interaction of Internal Merge and Agree. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar*. Linguistische Arbeitsberichte 90, Leipzig University, pp. 413–461.

Georgi, Doreen, Fabian Heck and Gereon Müller (2009): Maraudage. Ms., Universität Leipzig. (Talk at the Repairs Workshop of the DGfS conference 2009, Osnabrück, and at the Potsdam/Leipzig Workshop on Movement and Morphology, Leucorea, April 2009).

Grewendorf, Günther (1989): *Ergativity in German*. Foris, Dordrecht.

Grewendorf, Günther and Joachim Sabel (1999): 'Scrambling in German and Japanese', *Natural Language and Linguistic Theory* **17**, 1–65.

Grimshaw, Jane (1994): Heads and Optimality. Ms., Rutgers University. (Handout of talk, Universität Stuttgart).

Heck, Fabian and Gereon Müller (2000): Repair-Driven Movement and the Local Optimization of Derivations. Ms., Universität Stuttgart and IDS Mannheim.

Heck, Fabian and Gereon Müller (2007): Extremely Local Optimization. *In:* E. Brainbridge and B. Agbayani, eds, *Proceedings of the 26th WECOL*. California State University, Fresno, pp. 170–183. [P5].

Heck, Fabian and Gereon Müller (2013): Extremely Local Optimization. *In:* H. Broekhuis and R. Vogel, eds, *Linguistic Derivations and Filtering*. Equinox, Sheffield, pp. 136–165.

Höhle, Tilman (1978): *Lexikalistische Syntax: Die Aktiv-Passiv-Relation und andere Infinitkonstruktionen im Deutschen*. Niemeyer, Tübingen.

Hualde, José Ignacio and Jon Ortiz de Urbina, eds (2003): *A Grammar of Basque*. Mouton de Gruyter, Berlin.

Jelinek, Eloise (1993): Ergative 'Splits' and Argument Type. *In:* J. Bobaljik and C. Phillips, eds, *Papers on Case and Agreement I*. Vol. 18 of *MIT Working Papers in Linguistics*, MITWPL, MIT: Cambridge, Mass., pp. 15–42.

Kayne, Richard (1989): Facets of Participle Agreement in Romance. *In:* P. Benincà, ed., *Dialect Variation and the Theory of Grammar*. Foris, Dordrecht, pp. 85–104.

Kiparsky, Paul (1973): Abstractness, Opacity and Global Rules. *In:* O. Fujimura, ed., *Three Dimensions in Linguistic Theory*. TEC, Tokyo, pp. 57–86.

Koopman, Hilda (2006): Agreement Konfigurations. *In:* C. Boeckx, ed., *Agreement Systems*. Benjamins, Amsterdam, pp. 159–201.

Koutsoudas, Andreas (1966): *Writing Transformational Grammars: An Introduction*. McGrah-Hill Book Company, New York.

Koutsoudas, Andreas (1973): Unordered Rule Hypotheses. Ms., Indiana University Linguistics Club.

Kratzer, Angelika (1996): Severing the External Argument from its Verb. *In:* J. Rooryck and L. Zaring, eds, *Phrase Structure and the Lexicon.* Kluwer, Dordrecht, pp. 109–137.

Lahne, Antje (2012): Specificity-driven Syntactic Derivation. *In:* M. Uribe-Etxebarria and V. Valmala, eds, *Ways of Structure Building.* Oxford University Press, Oxford, pp. 271–296.

Lasnik, Howard (1999): 'On Feature Strength: Three Minimalist Approaches to Overt Movement', *Linguistic Inquiry* **30**, 197–217.

Legate, Julie Anne (2003): 'Some Interface Properties of the Phase', *Linguistic Inquiry* **34**, 506–516.

Legate, Julie Anne (2008): 'Morphological and Abstract Case', *Linguistic Inquiry* **39**(1), 55–101.

Manzini, Rita (1992): *Locality. A Theory and Some of Its Empirical Consequences.* MIT Press, Cambridge, Mass.

Marantz, Alec (1991): Case and Licensing. *In:* German Westphal, B. Ao and H.-R. Chae, eds, *Proceedings of the Eigth Eastern States Conference on Linguistics.* University of Maryland, pp. 234–253.

McCarthy, John (2010): 'An Introduction to Harmonic Serialism', *Language and Linguistics Compass* **4**, 1001–1018.

McFadden, Thomas (2004): The Position of Morphological Case in the Derivation: A Study on the Syntax-Morphology Interface. PhD thesis, University of Pennsylvania.

Merchant, Jason (2006): Polyvalent case, geometric hierarchies, and split ergativity. *In:* J. Bunting, S. Desai, R. Peachey, C. Straughn and Z. Tomkova, eds, *Proceedings of the 42nd annual meeting of the Chicago Linguistics Society.* Chicago Linguistics Society, Chicago, pp. 47–67.

Müller, Gereon (2004): Phrase Impenetrability and Wh-Intervention. *In:* A. Stepanov, G. Fanselow and R. Vogel, eds, *Minimality Effects in Syntax.* Mouton de Gruyter, Berlin, pp. 289–325.

Müller, Gereon (2009): Ergativity, Accusativity, and the Order of Merge and Agree. *In:* K. K. Grohmann, ed., *Explorations of Phase Theory. Features and Arguments.* Mouton de Gruyter, Berlin, pp. 269–308.

Müller, Gereon (2011): *Constraints on Displacement. A Phase-Based Approach.* Vol. 7 of *Language Faculty and Beyond*, Benjamins, Amsterdam.

Müller, Gereon (2012): A Local Reformulation of the Williams Cycle. Ms., Universität Leipzig.

Murasugi, Kumiko (1992): Crossing and Nested Paths. PhD thesis, MIT, Cambridge, Mass.

Nash, Léa (1996): The Internal Ergative Subject Hypothesis. *In:* K. Kusumoto, ed., *Proceedings of NELS 26.* GLSA, Amherst, Mass, pp. 195–209.

Neeleman, Ad and Hans van de Koot (2010): 'A Local Encoding of Syntactic Dependencies and its Consequences for the Theory of Movement', *Syntax* **13**, 331–372.

Nordlinger, Rachel (1998): *Constructive Case: Evidence from Australian Languages*. CSLI, Stanford.

Pesetsky, David (1989): Language-Particular Processes and the Earliness Principle. Ms., MIT, Cambridge, Mass.

Polinsky, Maria, Carlos Gallo, Peter Graff and Ekaterina Kravtchenko (2011): 'Subject Preference and Ergativity', *Lingua* pp. 267–277.

Pullum, Geoffrey (1979): *Rule Interaction and the Organization of a Grammar*. Garland, New York.

Richards, Marc (2008): Quirky Expletives. *In:* G. H. H. Roberta d'Alessandro and S. Fischer, eds, *Agreement Restrictions*. Mouton de Gruyter, pp. 181–213. [P5].

Richards, Norvin (2007): Lardil "Case Stacking" and the Structural/Inherent Case Distinction. Ms., MIT.

Rizzi, Luigi (1990): *Relativized Minimality*. MIT Press, Cambridge, Mass.

Roberts, Ian (2010): *Agreement and Head Movement: Clitics, Incorporation and Defective Goals*. MIT Press, Cambridge, Mass.

Rouveret, Alain and Jean-Roger Vergnaud (1980): 'Specifying Reference to the Subject', *Linguistic Inquiry* **11**, 97–202.

Sauerland, Uli (1999): 'Erasability and Interpretation', *Syntax* **3**, 161–188.

Smolensky, Paul (1995): On the Internal Structure of Con, the Constraint Component of UG. Ms., Johns Hopkins University.

Sportiche, Dominique (1989): 'Le Mouvement Syntaxique: Contraintes et Paramètres', *Langages* **95**, 35–80.

Starke, Michal (2001): Move Dissolves Into Merge: A Theory of Locality. PhD thesis, University of Geneva.

Stechow, Arnim von and Wolfgang Sternefeld (1988): *Bausteine syntaktischen Wissens*. Westdeutscher Verlag, Opladen.

Sternefeld, Wolfgang (1996): Comparing Reference Sets. *In:* C. Wilder, H.-M. Gärtner and M. Bierwisch, eds, *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 81–114.

Sternefeld, Wolfgang (2006): *Syntax*. Stauffenburg, Tübingen. Two volumes.

Stiebels, Barbara (2000): Linker Inventories, Linking Splits and Lexical Economy. *In:* B. Stiebels and D. Wunderlich, eds, *Lexicon in Focus*. Akademie-Verlag, Berlin, pp. 211–245.

Stiebels, Barbara (2006): ‘Agent Focus in Mayan Languages’, *Natural Language and Linguistic Theory* **24**, 501–570.

Svenonius, Peter (1994): ‘C-Selection as Feature-Checking’, *Studia Linguistica* **48**, 133–155.

Thiersch, Craig (1978): Topics in German Syntax. PhD thesis, MIT, Cambridge, Mass.

Unger, Christina (2010): A Computational Approach to the Syntax of Displacement and the Semantics of Scope. PhD thesis, Universiteit Utrecht, LOT.

Ura, Hiroyuki (2000): *Checking Theory and Grammatical Functions in Universal Grammar*. Oxford University Press, New York.

van Koppen, Marjo (2005): One Probe – Two Goals. Aspects of Agreement in Dutch Dialects. PhD thesis, Leiden University.

Wiese, Bernd (1999): ‘Unterspezifizierte Paradigmen. Form und Funktion in der pronominalen Deklination’, *Linguistik Online* **4**. (www.linguistik-online.de/3_99).

Williams, Edwin (1974): Rule Ordering in Syntax. PhD thesis, MIT, Cambridge, Mass.

Williams, Edwin (2003): *Representation Theory*. MIT Press, Cambridge, Mass.

Woolford, Ellen (1997): ‘Four-Way Case Systems: Ergative, Nominative, Objective and Accusative’, *Natural Language and Linguistic Theory* **15**, 181–227.

Woolford, Ellen (2001): Case Patterns. *In:* G. Legendre, J. Grimshaw and S. Vikner, eds, *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 509–543.

Wunderlich, Dieter (1997): ‘Cause and the Structure of Verbs’, *Linguistic Inquiry* **27**, 27–68.

Wurmbrand, Susanne (2001): *Infinitives. Restructuring and Clause Structure*. Mouton de Gruyter, Berlin.