**On the Order of Syntactic Operations**

**Lecture 3: A Local Optimality-Theoretic Perspective**

Compact course, University of Cambridge, June 14-19, 2013

Gereon Müller (Universität Leipzig)

*Note*:
This handout mainly reports joint work with Fabian Heck; see Heck & Müller (2013b).

## 1. Introduction

*Four claims*:

- There is a competition between Merge and Agree which can and must be resolved in one way or the other.
  (We already know that from lecture 2.)
- Resolving the competition by giving preference to (ranking) one or the other requirement accounts for (a) ergative vs. accusative patterns of argument encoding (on the vP cycle) and (b) a restriction on ergative movement (on the TP cycle). This can be modelled by invoking a (parametrized) preference requirement in minimalist syntax. (Again, we already know that from lecture 2.) Alternatively, it can be modelled in an optimality-theoretic way. (This will be new, but not very different.)
- Movement in accusative encoding systems is in some contexts accelerated, so that it applies before Agree; and movement in ergative encoding systems 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 lexically case-marked (and some ergative) arguments in some languages. Such effects can straightforwardly be derived in an optimality-theoretic approach (they signal the presence of more specific, higher-ranked constraints); they are less straightforward in a minimalist approach.
- The reasoning presupposes an extremely local, derivational approach to optimization (because of the *opacity* (counter-bleeding) effects).

## 2. Background

*Question*:
Does syntactic optimization apply once (*harmonic parallelism*: *representational syntax*) or more than once (*harmonic serialism*: *derivational syntax*)? If the latter holds: Is optimization *global* or *local*?



> Much of the analysis given in this book will be in the parallel mode, and some of the results will absolutely require it. But it is important to keep in mind that the serial/parallel distinction pertains to Gen and not to the issue of harmnonic evaluation *per se*. It is an empirical question [...] Many different theories [...] can be equally well accommodated in Gen, and the framework of Optimality Theoriy *per se* involves no commmitment to any set of such assumptions.
>
> Prince & Smolensky (2004, 95-96)

*Viability of derivational OT*:

- McCarthy (2008; 2010) and much related recent work (also cf. McCarthy (2000), and to some extent McCarthy (2007a)): A harmonic serialism (derivational OT) approach to phonology is viable (and empirically preferable).
- Legendre et al. (1998), Smolensky & Legendre (2006): A derivational OT approach to syntax is viable.

*Two kinds of derivational OT syntax*

> While some see a major divide between the derivationally-oriented MP and OT, we do not. Of course, there are likely to be differences of empirical import between the non-derivational, chain-based theory of "Shortest Move" developed here and a particular derivational MP proposal, but such differences seem comparable to those between different approaches to syntax within OT, or to those between different proposals within MP: they do not seem to follow from some major divide between the OT and MP frameworks. In fact, derivational theories can be naturally formalized within OT. *[1]* "Harmonic serialism" is a derivational version of OT developed in Prince & Smolensky (1993) in which each step of the derivation produces the optimal next representation. *[2]* Another approach, seemingly needed to formalize MP within OT has *Gen* produce *derivations*; it is these that are evaluated by the constraints, the optimal derivation being determined via standard OT evaluation. Thus, on our view, while the issue of derivations is an important one, it is largely orthogonal to OT.
>
> Legendre et al. (1998, 285-286)

*Two ways to reconcile a derivational approach to syntax with OT*:

- standard, parallel optimization of full derivations
- serial optimization

*A relevant issue in the second case*:

- In classical transformational grammar (e.g., Chomsky (1965)), syntactic transformations applying to the output of the base component effect derivational steps where the input and the output have roughly *the same size*, exactly as in phonology.
- In the minimalist program, operations of the "base" component and of the "transformational" component are systematically interspersed; syntactic structures start with two lexical items and grow throughout the derivation by iterated application of (external or internal) Merge. Here, iterated optimization cannot apply to objects of the same size – the optimal output of one optimization procedure is *smaller* than the optimal output of the next optimization procedure.

*Consequence: two dichotomies*

- Optimization may be *parallel* (i.e., apply once) or *serial* (i.e., apply more than once).
- Optimization may be *global* (applying to the full sentence) or *local* (applying also to smaller domains).

*Serial optimization and locality*

- Serial optimization in phonology must be global (phonology is not structure-building).
- *Serial optimization in minimalist syntax must be local*: The linguistic expressions created by Merge grow steadily throughout the derivation.

*What are the local domains for optimization?*

(1) *Optimization domains*:
  a. *sentence* (single or multiple optimization, derivational or representational)
  b. *minimal clause* (e.g., CP; potentially multiple optimization, derivational)
  c. *phase* (CP, vP, DP): multiple optimization, derivational)
  d. *phrase* (XP: multiple optimization, derivational)
  e. *derivational step* (multiple optimization, derivational)

*Background*:

1. Classical assumption: The whole sentence is subject to a single, parallel optimization procedure (Grimshaw (1997), Pesetsky (1998), Legendre et al. (1998), etc.). The output candidates are usually taken to be representations (also in approaches that assume them to be generated by a minimalist GEN component; see Broekhuis & Dekkers (2000), Broekhuis (2000; 2006; 2008), Salzmann (2012)); but they can also be full derivations (as, e.g., in Müller (1997)).



2. Wilson (2001), Heck (1998; 2001): multiple optimization of whole sentences (still global).
3. Multiple optimization of smaller optimization domains: closely related to developments in the minimalist program.

*Trade-offs: conceptual arguments*

- The smaller the optimization domain is, the more the complexity of the overall system is reduced (reduction of the size of candidate sets).
- The larger the optimization domain is, the less often optimization procedures have to be carried out.

*The view adopted here*:

- Iterated optimization in small domains is more economical than single optimization of extremely large domains.
- Ultimately, empirical arguments should decide the issue: If the ranked constraints have access to more/less structure, a wrong winner is predicted.

(2) *Proposals for local domains for optimization*:
  a. *Minimal clause*:
  Ackema & Neeleman (1998) on *wh*-movement in Czech; Müller (2003) on extraction from verb-second clauses in German
  b. *Phase*:
  Fanselow & Ćavar (2001) on MeN-deletion in Malay; Müller (2000a; 2002) on R-pronouns in German
  c. *Phrase*:
  Fischer (2004; 2006) on reflexivization (including long-distance reflexivization); Müller (2000b) on secondary remnant movement; Heck & Müller (2000; 2003) on *wh*-movement, superiority, quantifier raising, and sluicing
  d. *Derivational step*:
  Heck & Müller (2007; 2013a) on gender agreement with dative possessors in German DPs, expletives in German verb-second clauses, and VP topicalization and *do*-support in English; Müller (2004; 2009) on ergative and accusative argument encoding patterns; Lahne (2008; 2009) on excluding SVO in ergative languages; Georgi (2012a) on global case splits in Tauya; Georgi (2012b) on anti-agreement effects and defective intervention.

*Extremely local optimization*:
Let us pursue the consequences of the most radical position within a theory of local optimization, where the domain is the derivational step (also see McCarthy (2010)).

*See also*:

> This is tantamount to the claim that each transformational rule application constitutes a "phase," which we believe to be the null hypothesis.
>
> Epstein & Seely (2002, 77)

*Shape of the argument*:

1. Sometimes, the order of applying Agree and Merge is under-determined. If there are no simultaneous rule applications in the grammar (see Epstein & Seely (2002); contra Pullum (1979), Chomsky (2008)), then a conflict arises: Only one of them can be executed at each step.

2. 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.

3. If the optimization domain is larger than the step-level, then, ceteris paribus, the order of elementary operations that is imposed by the ranking under step-level optimization cannot be preserved. Empirically, this is the wrong result.

## 3. Merge vs. Agree on the vP level: Accusative vs. Ergative Encoding

*Refs.*: Müller (2009); Heck & Müller (2013a)

### *3.1 Assumptions*

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

(4) *Two types of features that drive operations*:

a. Structure-building features (edge features, subcategorization features) trigger Merge: [•F•].
b. Probe features trigger Agree: [∗F∗].

*Three operations* (Chomsky (2001)):

(5) *Merge*:
$\alpha$ can be merged with $\beta$, yielding $[_\alpha\ \alpha, \beta\ ]$, if $\alpha$ bears a structure-building feature [•F•] and F is the label of $\beta$.

(6) *Move*:
Move is Merge, with $\beta$ internal to $\alpha$.

(7) *Agree*:
$\alpha$ can agree with $\beta$ with respect to a feature bundle $\Gamma$ iff (a) and (b) hold:

a. $\alpha$ bears a probe feature [∗F∗] in $\Gamma$ and may thereby provide the $\alpha$-value for a matching goal feature [F] of $\beta$ in $\Gamma$.



b. $\alpha$ m-commands $\beta$.
(This permits an Agree relation between a head and its specifier.)

*Constraints*:
Designated constraints ensure that Merge (incl. Move) and Agree must take place as soon as their context of application is present. (This derives an earliness requirement for syntactic operations; see Pesetsky (1989).)

(8) Agree Condition (AC):
Probes ([∗F∗]) participate in Agree.

(9) Merge Condition (MC):
Structure-building features ([•F•]) participate in Merge.

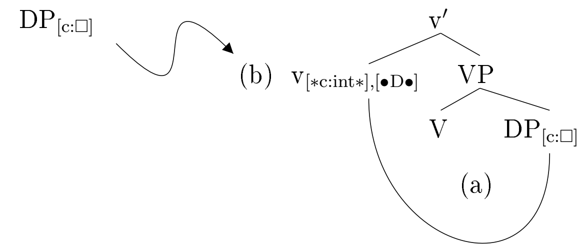
(10) *The role of T and v in argument encoding*:

a. T bears [∗case:*ext*∗] that instantiates [case:*ext*] on DP.
b. v bears [∗case:*int*∗] that instantiates [case:*int*] on DP.

*A conspicuous property*:
The head v has a dual role: It participates in a Merge operation with a DP, and it also participates in an Agree relation with a DP. This dual role has far-reaching consequences for the nature of argument encoding.

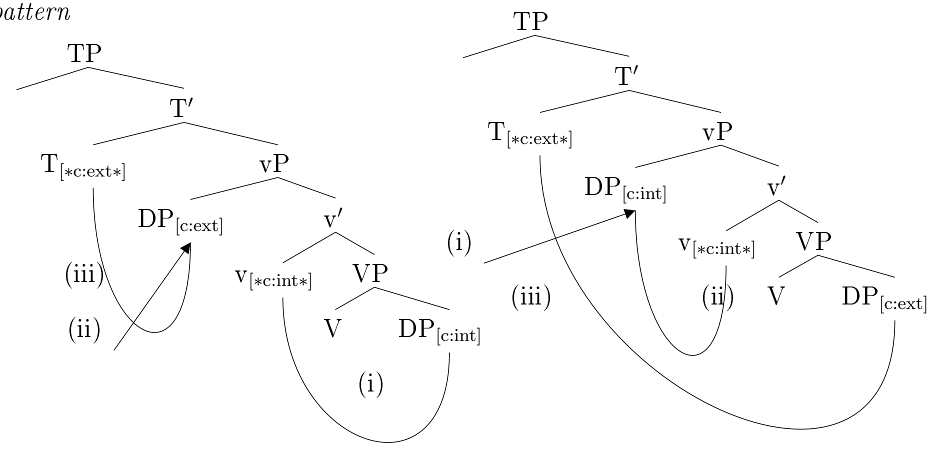
*A constraint conflict*:
Consider a simple transitive context, with two arguments DP$_{int}$, DP$_{ext}$. Suppose that the derivation has reached a stage $\Sigma$ where v has been merged with a VP containing DP$_{int}$, with DP$_{ext}$ waiting to be merged with v in the workspace of the derivation. At this point, a conflict arises: AC demands that the next operation is Agree(v,DP$_{int}$) (see (a)), MC demands that it is Merge(DP$_{ext}$,v) (see (b)).

(11) *Stage $\Sigma$:*



## 3.2 Analysis

(12) a. *Agree before Merge: accusative pattern* b. *Merge before Agree: ergative pattern*




*Note*:
The derivation of the ergative pattern presupposes that a specifier is preferred with respect to Agree with its head to an item included in the complement of that head. This can be formulated as the *Specifier-Head Bias* (Chomsky (1986; 1995), Koopman (2006); see Béjar & Řezáč (2009) for a similar idea with the bias inversed).

(13) Specifier-Head Bias (SHB):
Spec/head Agree is preferred to other instances of Agree.

This replaces standard minimality conditions (Relativized Minimality, MLC) (though with a somewhat different empirical coverage). The SHB is compatible with equidistance effects, which pose a problem for path-based definitions of minimality.

(14) *Rankings*:

a. (SHB $\gg$) AC $\gg$ MC (accusative pattern)
b. (SHB $\gg$) MC $\gg$ AC (ergative pattern)

*$T_1$: Accusative pattern, step 1 ($\Sigma$ as input): Agree*

| Input: $[_{v'}\ v_{[*CASE:int*],[\bullet D\bullet]} \ldots DP_{[CASE:\square]} \ldots]$<br>Workspace = $\{DP_{[CASE:\square]}, T_{[*CASE:ext*]}, \ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| $O_1$: $[_{v'}\ DP_{[CASE:\square]}\ [_{v'}\ v_{[*CASE:int*]} \ldots DP_{[CASE:\square]} \ldots]]$ | | *! | |
| ☞ $O_2$: $[_{v'}\ v_{[\bullet D\bullet]} \ldots DP_{[CASE:int]} \ldots]$ | | | * |

*Note*:
In intransitive contexts, only T can still assign case via Agree. This derives the basic



*$T_2$: Accusative pattern, step 2: Merge*

| Input: $DP_{[CASE:\square]} \oplus [_{v'}\ v_{[\bullet D\bullet]} \ldots DP_{[CASE:int]} \ldots]$<br>Workspace = $\{T_{[*CASE:ext*]}, \ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| ☞ $O_1$: $[_{vP}\ DP_{[CASE:\square]}\ [_{v'}\ v \ldots DP_{[CASE:int]} \ldots]]$ | | | |

*$T_3$: Accusative pattern, step 3: Agree*

| Input: $[_{T'}\ T_{[*CASE:ext*]}\ [_{vP}\ DP_{[CASE:\square]} \ldots DP_{[CASE:int]} \ldots]]$<br>Workspace = $\{\ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| ☞ $O_1$: $[_{TP}\ T\ [_{vP}\ DP_{[CASE:ext]}\ [_{v'}\ v \ldots DP_{[CASE:int]} \ldots]]]$ | | | |

*$T_4$: Ergative pattern, step 1 ($\Sigma$ as input): Merge*

| Input: $DP_{[CASE:\square]} \oplus [_{v'}\ v_{[*CASE:int*],[\bullet D\bullet]} \ldots DP_{[CASE:\square]} \ldots]$<br>Workspace = $\{T_{[*CASE:ext*]}, \ldots\}$ | SHB | MC | AC |
|---|---|---|---|
| ☞ $O_1$: $[_{v'}\ DP_{[CASE:\square]}\ [_{v'}\ v_{[*CASE:int*]} \ldots DP_{[CASE:\square]} \ldots]]$ | | | * |
| $O_2$: $[_{v'}\ v_{[\bullet D\bullet]} \ldots DP_{[CASE:int]} \ldots]$ | | *! | |

*$T_5$: Ergative pattern, step 2: Agree (with $DP_{ext}$)*

| Input: $[_{v'}\ DP_{[CASE:\square]}\ [_{v'}\ v_{[*CASE:int*]} \ldots DP_{[CASE:\square]} \ldots]]$<br>Workspace = $\{T_{[*CASE:ext*]}, \ldots\}$ | SHB | MC | AC |
|---|---|---|---|
| ☞ $O_1$: $[_{vP}\ DP_{[CASE:int]}\ [_{v'}\ v \ldots DP_{[CASE:\square]} \ldots]]$ | | | |
| $O_2$: $[_{vP}\ DP_{[CASE:\square]}\ [_{v'}\ v \ldots DP_{[CASE:int]} \ldots]]$ | *! | | |

ergative and accusative patterns of argument encoding (a version of Murasugi's (1992) approach).

## 3.3 Extremely Local vs. Less Local Optimization

*Note*:
If phrases (or phases, or clauses, or sentences) are the local domains for optimization, things do not change in ergative contexts (see $T_7$), but a wrong winner is predicted in accusative contexts (see $T_8$). Legitimate assignment of accusative case by v in the presence of $DP_{ext}$ in Specv emerges as an instance of *counter-bleeding* opacity (Chomsky (1951; 1975), Kiparsky (1973)) – one that cannot be accounted for representationally by adding postulating abstract items (like traces).

*$T_6$: Ergative pattern, step 3: Agree*

| Input: $[_{T'}\ T_{[*\text{CASE}:ext*]}\ [_{vP}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:\square]} \ldots]]]$ Workspace $= \{\ldots\}$ | SHB | MC | AC |
|---|---|---|---|
| ☞ $O_1$: $[_{T'}\ T\ [_{vP}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:ext]} \ldots]]]$ | | | |

*$T_7$: vP optimization under MC ≫ AC ("ergative") ranking: right result*

| Input: $DP_{[\text{CASE}:\square]} \oplus v_{[*\text{CASE}:int*],[\bullet V\bullet]\succ[\bullet D\bullet]} \oplus [_{VP} \ldots DP_{[\text{CASE}:\square]} \ldots]$ Workspace $= \{T_{[*\text{CASE}:ext*]}, \ldots\}$ | SHB | MC | AC |
|---|---|---|---|
| ☞ $O_1$: $[_{vP}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:\square]} \ldots]]$ | | | |
| $O_2$: $[_{vP}\ DP_{[\text{CASE}:\square]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]$ | *! | | |

*$T_8$: vP optimization under AC ≫ MC ("accusative") ranking: wrong result*

| Input: $DP_{[\text{CASE}:\square]} \oplus v_{[*\text{CASE}:int*],[\bullet V\bullet]\succ[\bullet D\bullet]} \oplus [_{VP} \ldots DP_{[\text{CASE}:\square]} \ldots]$ Workspace $= \{T_{[*\text{CASE}:ext*]}, \ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| $O_1$: $[_{vP}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:\square]} \ldots]]$ | | | |
| ☛ $O_2$: $[_{vP}\ DP_{[\text{CASE}:\square]}\ [_{v'}\ v \ldots DP_{[\text{CASE}:int]} \ldots]]$ | *! | | |

## 4. Interlude: Syntax vs. Phonology

*Opacity in Syntax vs. Opacaity in Phonology*:

- Syntax:
    1. Standard global parallel optimization cannot capture opacity effects.
    2. Extremely local serial optimization can capture opacity effects.
       (However, even slightly less local optimization, with the phrase as the optimization domain, cannot account for the instance of opacity at hand.)
- Phonology:
    1. Standard parallel optimization cannot capture opacity effects.
       (Possible ways out: (i) enrichment of representations with abstract materical encoding earlier movement steps: *turbid phonology* (Goldrick (2000)), *virtual phonology* (Bye (2001)), *coloured containment* (Oostendorp (2006; 2007), Trommer (2011)); (ii) *sympathy theory* (McCarthy (1999): non-optimal candidates encoding earlier derivational stages whose properties the optimal candidate must be faithful to); (iii) *candidate chains* (McCarthy (2007b)): faithfulness violations incurred in the derivation are recorded on output candidates, and *precedence* constraints demand a certain order of faithfulness violations in the output candidate.)
    2. Serial optimization cannot account for opacity either.



*Observation* (McCarthy (2007a, 37-38; 69-70)):
Serial optimization as such does not capture counter-bleeding and counter-feeding in phonology.

(15) *Counter-Bleeding in Bedouin Arabic*

| | |
|---|---|
| UR | /ħaːkim-iːn/ |
| Palatalization | ħaːkʲimiːn |
| Vowel deletion | ħaːkʲmiːn |
| Output | [ħaːkʲmiːn] |

See /t-ħakum-in/ → [tħakmin]

(16) *Counter-Feeding in Bedouin Arabic*

| | |
|---|---|
| UR | /dafaʕ/ |
| Vowel deletion | – |
| Vowel raising | difaʕ |
| Output | [difaʕ] |

See /ʃarib-at/ → [ʃarbat]

*Assumption* (McCarthy (2000); also see Heck & Müller (2007)):
Optimization applies cyclically: The output of one competition forms the sole input for the next competition, and so on, until the derivation *converges*, i.e., until subsequent optimizations always return the same candidate as optimal.

*Additional assumption*:
The competing outputs generated by GEN on the basis of a given input differ from the input only by at most one application of an elementary operation.

*Problem*:
Opacity cannot be derived in this way.

(17) *Counter-Bleeding does not work*

a. *Constraints*:
   (i) *$i$CV: triggers $i$-deletion.
   (ii) *$ki$: triggers palatalization of the velar before $i$.
   (iii) MAX: prohibits deletion of $i$.
   (iv) ID(back): prohibits palatalization of $k$.

b. *Cycle 1: wrong winner*

| /ħaːkim-iːn/ | *$ki$ | *$i$CV | MAX | ID(back) |
|---|---|---|---|---|
| ☛ħaːkmiːn | | | * | |
| ☞ħaːkʲimiːn | | *! | | * |
| ħaːkʲmiːn | | | * | *! |
| ħaːkimiːn | *! | * | | |

c. *Cycle 2: fatal convergence*

| /ħaːkmiːn/ | **ki* | **i*CV | Max | Id(back) |
|---|---|---|---|---|
| ☛ħaːkmiːn | | | | |
| ☞ħaːk$^j$imiːn | | *! | | * |
| ħaːkimiːn | *! | * | | |
| ☞ħaːk$^j$miːn | | | | *! |

*Underlying reason for the failure of serial optimization to capture counter-bleeding*:
(i) In a rule-based approach, early application of the palatalization rule brings about palatalization.
(ii) In a serial OT approach, a high ranking of the constraint **ki*, which triggers palatalization, does in fact *not* trigger early palatalization here: The constraint triggers palatalization only *indirectly*, and can in principle also be satisfied by other elementary operations that are also triggered in an indirect way by a markedness constraint.

(18) *Another non-solution: reranking of constraints*

a. *Cycle 1: correct intermediate winner*

| /ħaːkim-iːn/ | **ki* | Max | **i*CV | Id(back) |
|---|---|---|---|---|
| ħaːkmiːn | | *! | | |
| ☞ħaːk$^j$imiːn | | | * | * |
| ħaːkimiːn | *! | | * | |

b. *Cycle 2: wrong final winner, fatal convergence*

| /ħaːk$^j$im-iːn/ | **ki* | Max | **i*CV | Id(back) |
|---|---|---|---|---|
| ☞ħaːk$^j$miːn | | *! | | |
| ☛ħaːk$^j$imiːn | | | * | |
| ħaːkimiːn | *! | | * | |

*Reason*:
The high ranking of Max vis-à-vis the counteracting markedness constraint **i*CV is needed in cycle 1 to generate the intended intermediate output, but afterwards it proves fatal since *i*-deletion is then not permitted anymore. (More generally, *i*-deletion should then be blocked.)

*Conjecture*:
Counter-bleeding can be derived by serial optimization in syntax, but not in phonology, because in syntax, it typically does not happen that the application of *one* elementary operation leads to a satisfaction of *two* markedness constraints (e.g., AC and MC). Strategy for phonology: Ensure (by adding an appropriate highly ranked, or inviolable, constraint) that *only one new markedness constraint* can be satisfied by an optimal candidate.



## 5. Merge vs. Agree on the TP level: A Constraint on Ergative Movement

*Ref.*: Assmann, Georgi, Heck, Müller & Weisser (2012)

### *5.1 Data*

*Observation* (Campana (1992), Aldridge (2004), Stiebels (2006), Coon et al. (2011), Deal (2012)):
In many morphologically ergative languages, ergative arguments (DP$_{erg}$) cannot undergo Ā-movement (wh-movement, focussing, relativization).

(19) *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?'

### *5.2 Assumptions*

*Locality of movement*:
Movement takes place successive-cyclically, from one XP edge domain to the next one higher up. Given the Phase Impenetrability Condition (PIC; Chomsky (2001)), this follows automatically if every XP is a phase.

*Edge features*:
It must be ensured that intermediate steps of movement as required under the PIC are possible in the first place in a model of syntax where all operations are feature-driven. A standard assumption here is that *edge features* ([•X•]) that trigger intermediate movement steps can be inserted on all intervening phase heads.

*Maraudage*:
Certain goal features require checking in Spec/head configurations; this way, they may "maraud" a functional head and take away features that should normally be reserved for some other item. (See Georgi, Heck & Müller (2009); similar concepts are suggested in Chomsky (2001), Abels (2003), Anagnostopoulou (2005), Adger & Harbour (2007), Béjar & Řezáč (2009).)

*Case features and maraudage*:
Structural case features trigger maraudage in Spec/head configurations even if they have already been checked (or valued). Independent motivation: the existence of *case*

*stacking* in the world's languages (see Andrews (1996), Nordlinger (1998), Richards (2007)).

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

*Note*:
Given the SHB, the configuration in (21-a) may involve checking of [case:int] by X or not (leading to a crash of the derivation or not because of an unchecked [case:□]), whereas the configuration in (21-b) must involve checking of [case:int] by X (which invariably leads to a crash).

(21) a. $[_{X'}\ X_{[*case:ext*]}\ [_{ZP}\ ...\ \alpha_{[case:int]}\ ...\ \beta_{[case:\square]}\ ...\ ]]$
b. $[_{XP}\ \alpha_{[case:int]}\ [_{X'}\ X_{[*case:ext*]}\ [_{ZP}\ ...\ t_{\alpha}\ ...\ \beta_{[case:\square]}\ ...\ ]]]$

*Recall*:
There is no minimality condition on Agree or Merge; minimality effects are derivable from the PIC; see Müller (2011). Suppose that both $\alpha$ and $\beta$ are PIC-accessible to X in (21); this will imply 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.

*Assumption*:
Checking of [case:int] on $\alpha$ with a conflicting [*case:ext*] on X is harmless as such; $\alpha$ will simply maintain its original feature value. However, [*case:ext*] is then discharged, and not available for further operations anymore.

### *5.3 Analysis*

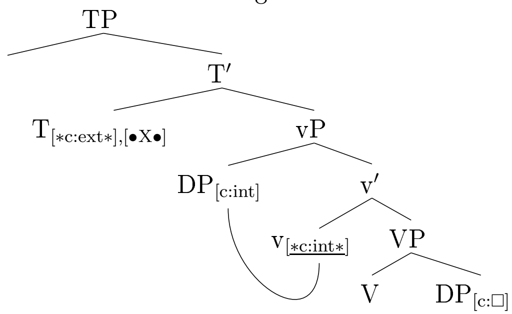
#### *5.3.1 Displacement in Languages with Ergative Encoding Patterns*

*(i) *DP$_{erg}$ Movement*
Given the PIC, DP$_{erg}$ needs to move from Specv to SpecT if it is to undergo subsequent movement to SpecC (*wh*-movement, relativization, focus movement). Given that the "ergative" ranking MC ≫ AC is also maintained on the TP cycle, movement of DP$_{erg}$ (as an instance of internal Merge) will have to *precede* Agree of T with the VP-internal DP that has not yet valued its case feature (as absolutive). Given the SHB, DP$_{erg}$ will next maraud T's case probe; the internal argument DP will consequently remain without a checked case feature. Assuming that all DPs must have their case features checked eventually (and assuming that there is no such thing as a default case), the derivation will crash. In a nutshell, ergative movement is impossible because the remaining argument cannot get absolutive case in this context.

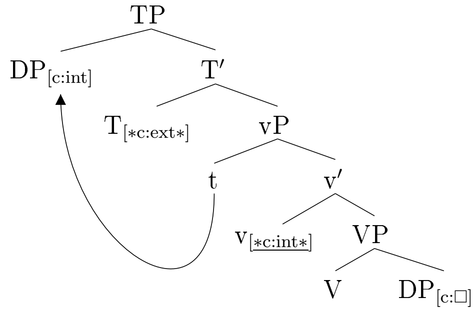


(22) *Illegitimate movement of DP$_{erg}$*

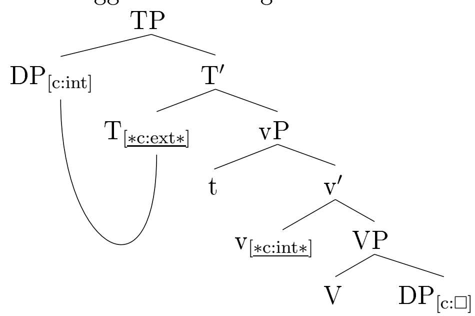
a. Structure after T is merged




b. MC ≫ AC triggers movement of DP$_{erg}$ first




c. SHB triggers maraudage of T




*T$_9$: Ergative movement, step 1: Move*

| Input: $[_{T'}\ T_{[*CASE:ext*],[\bullet X\bullet]}\ [_{vP}\ DP_{[CASE:int]}\ [_{v'}\ v\ ...\ DP_{[CASE:\square]}\ ...]]]]$ Workspace = {...} | SHB | MC | AC |
|---|---|---|---|
| ☞ O$_1$: $[_{TP}\ DP_{[CASE:int]}\ [_{T'}\ T_{[*CASE:ext*]}\ [_{vP}\ t\ [_{v'}\ v\ ...\ DP_{[CASE:\square]}\ ...]]]]$ | | | * |
| O$_2$: $[_{T'}\ T_{[\bullet X\bullet]}\ [_{vP}\ DP_{[CASE:int]}\ [_{v'}\ v\ ...\ DP_{[CASE:int]}\ ...]]]$ | | *! | |

*$T_{10}$: Ergative movement, step 2: Agree (maraudage)*

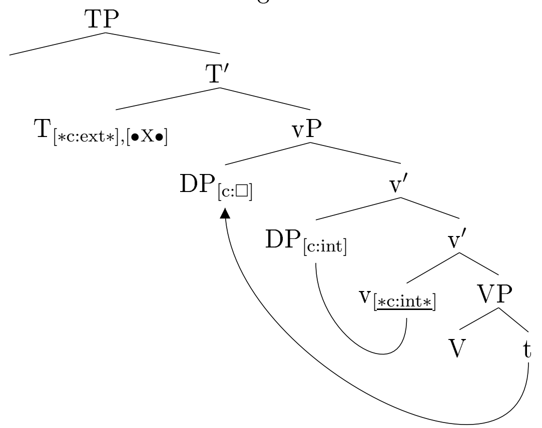
| Input: $[_{\text{TP}}$ DP$_{[\text{CASE}:int]}$ $[_{\text{T}'}$ T$_{[*\text{CASE}:ext*]}$ $[_{\text{vP}}$ t $[_{\text{v}'}$ v ... DP$_{[\text{CASE}:\square]}$ ... ]]]] Workspace = {...} | SHB | MC | AC |
|---|---|---|---|
| ☞ O$_1$: $[_{\text{TP}}$ DP$_{[\text{CASE}:ext,int]}$ $[_{\text{T}'}$ T $[_{\text{vP}}$ t $[_{\text{v}'}$ v ... DP$_{[\text{CASE}:\square]}$ ... ]]]] | | | |
| O$_2$: $[_{\text{TP}}$ DP$_{[\text{CASE}:int]}$ $[_{\text{T}'}$ T $[_{\text{vP}}$ t $[_{\text{v}'}$ v ... DP$_{[\text{CASE}:ext]}$ ... ]]]] | *! | | |

*Assumption re ineffability*:
A DP with an unchecked case feature will eventually lead to a crash of the derivation; this is an instance of Grimshaw's (1994) "no good output" approach to absolute ungrammaticality.

*(ii) DP$_{abs}$ Movement*
No such problem arises for movement of DP$_{abs}$ because DP$_{erg}$ has already been assigned case when DP$_{abs}$ moves to SpecT.

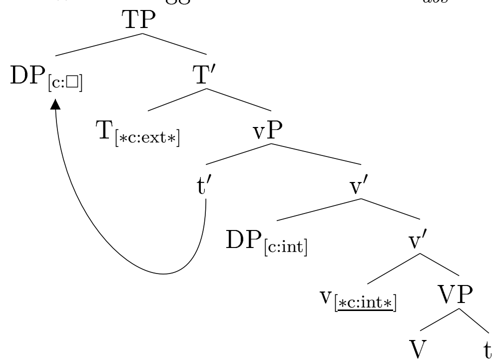
(23) *Legitimate movement of DP$_{abs}$*

a. Structure after T is merged

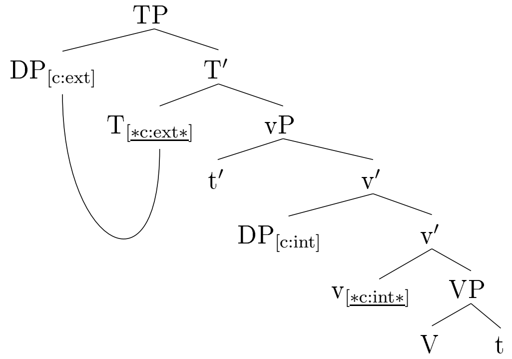





b. MC ≫ AC triggers movement of DP$_{abs}$ first




c. Finally, Agree with T ensures external case of DP$_{abs}$; no maraudage




*Note*:
On the vP cycle in (23-a), MC ≫ AC ensures that external Merge of DP$_{ext}$ and (subsequent; Chomsky (2001; 2008)) internal Merge of DP$_{int}$ (both triggered by ([•X•]) features on v) both precede Agree. Since there is no MLC-like constraint and both items occupy a Specv position (so the SHB does not discriminate the options), the derivation can now proceed in two ways: Agree(v,DP$_{ext}$) ultimately leads to a well-formed output, as indicated; in contrast, Agree(v,DP$_{int}$) in (23-a) would lead to a crash because DP$_{ext}$ would then never be assigned case.

### *5.3.2 Displacement in Languages with Accusative Encoding Patterns*

*(iii) DP$_{acc}$ Movement*
The ranking Agree ≫ Merge that gives rise to an accusative pattern in the first place (on the vP cycle) is also active on the TP cycle. Here it ensures that Agree with the DP$_{nom}$ in Specv can be carried out *before* the DP$_{acc}$ undergoes successive-cyclic movement to SpecT (and then to a higher position).

*$T_{11}$: Absolutive movement, step 1: Move*

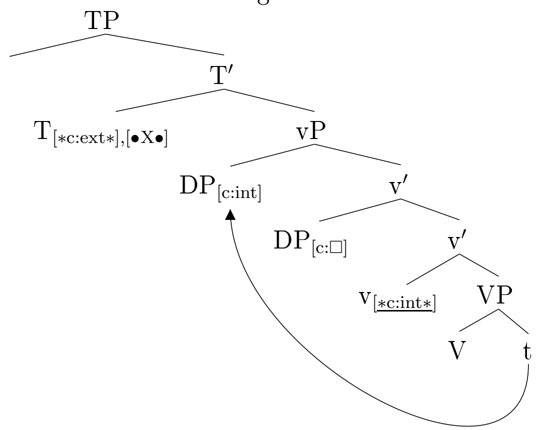
| Input: $[_{T'}\ T_{[*\text{CASE}:ext*],[\bullet X\bullet]}\ [_{vP}\ DP_{[\text{CASE}:\square]}\ [_{v'}\ DP_{[\text{CASE}:int]}\ [_{v'}\ v \ldots t \ldots]]]]$ Workspace = $\{\ldots\}$ | SHB | MC | AC |
|---|---|---|---|
| ☞ $O_1$: $[_{TP}\ DP_{[\text{CASE}:\square]}\ [_{T'}\ T_{[*\text{CASE}:ext*]}\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

*$T_{12}$: Absolutive movement, step 2: Agree (with SpecT)*

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

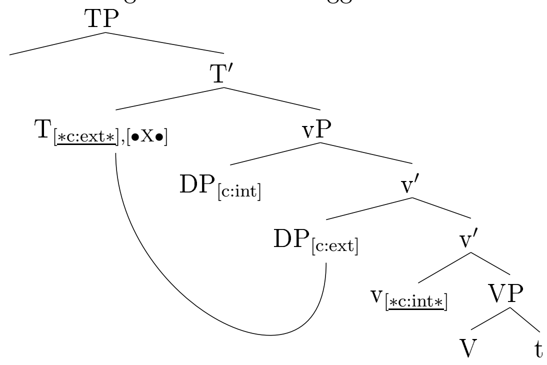
(24) *Legitimate movement of $DP_{acc}$*

a. Structure after T is merged

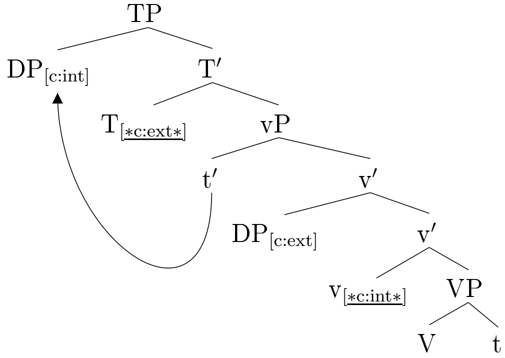





b. No maraudage: AC ≫ MC triggers case valuation of $DP_{nom}$ next



c. Finally, movement of $DP_{acc}$ takes place to SpecT




*$T_{13}$: Accusative Movement, step 1: Agree*

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

*$T_{14}$: Accusative Movement, step 2: Move*

| Input: $[_{T'}\ T_{[\bullet X\bullet]}\ [_{vP}\ DP_{[CASE:int]}\ [_{v'}\ DP_{[CASE:ext]}\ [_{v'}\ v \ldots t \ldots]]]]$ <br> Workspace = $\{\ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| ☞ $O_1$: $[_{TP}\ DP_{[CASE:int]}\ [_{T'}\ T\ [_{vP}\ t'\ [_{v'}\ DP_{[CASE:ext]}\ [_{v'}\ v \ldots t \ldots]]]]]$ | | | |
| ☞ $O_2$: $[_{TP}\ DP_{[CASE:ext]}\ [_{T'}\ T\ [_{vP}\ DP_{[CASE:int]}\ [_{v'}\ t'\ [_{v'}\ v \ldots t \ldots]]]]]$ | | | |

*Note*:
Nothing so far rules out $O_2$ in $T_{14}$. However, because of the PIC, only $DP_{ext}$ can move on. Eventually, this leads to unchecked operator features on the attracting head and $DP_{int}$, and thus to a crash of the derivation.

*(iv) $DP_{nom}$ Movement*
Similarly to the $DP_{abs}$ case, there is no problem for movement of $DP_{nom}$ because $DP_{acc}$ has already been assigned case when $DP_{nom}$ moves.

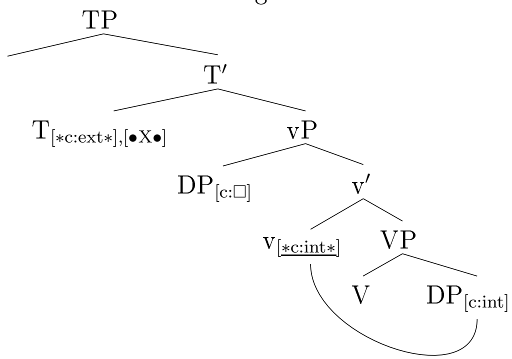


*$T_{15}$: Nominative movement, step 1: Agree*

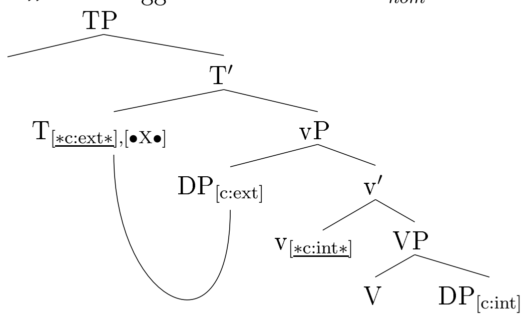
| Input: $[_{T'}\ T_{[*CASE:ext*],[\bullet X\bullet]}\ [_{vP}\ DP_{[CASE:\square]}\ [_{v'}\ v \ldots DP_{[CASE:int]} \ldots]]]$ <br> Workspace = $\{\ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| $O_1$: $[_{TP}\ DP_{[CASE:\square]}\ [_{T'}\ T_{[*CASE:ext*]}\ [_{vP}\ t\ [_{v'}\ v \ldots DP_{[CASE:int]}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

(25) *Legitimate movement of $DP_{nom}$*

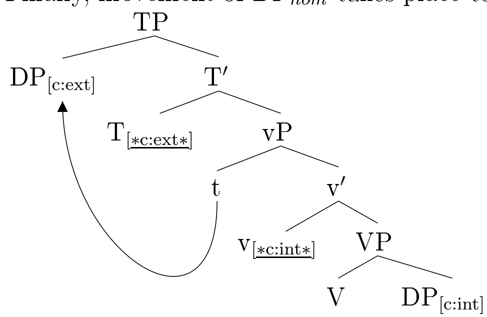
a. Structure after T is merged




b. AC ≫ MC triggers valuation of $DP_{nom}$ next




c. Finally, movement of $DP_{nom}$ takes place to SpecT



*$T_{16}$: Nominative movement, step 2: Move*

| Input: $[_{T'}\ T_{[\bullet X \bullet]}\ [_{vP}\ DP_{[CASE:\mathit{ext}]}\ [_{v'}\ v \ldots DP_{[CASE:\mathit{int}]} \ldots]]]$ <br> Workspace = $\{\ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| ☞ $O_1$: $[_{TP}\ DP_{[CASE:\mathit{ext}]}\ [_{T'}\ T_{[\bullet X \bullet]}\ [_{vP}\ t\ [_{v'}\ v \ldots DP_{[CASE:\mathit{int}]} \ldots]]]]$ | | | |

*5.4 Extremely Local vs. Less Local Optimization*

*Note*:
Again, if phrases (phases, clauses, sentences) are optimization domains, a wrong prediction is made for accusative contexts: Maraudage would be expected to arise. As before, non-occurrence of Agree with the moved accusative DP, and thus maraudage of T's case features, in the presence of $DP_{int}$ in SpecT illustrates *counter-bleeding* opacity.

*$T_{17}$: TP optimization under AC ≫ MC ("accusative") ranking: wrong result*

| Input: $T_{[*CASE:\mathit{ext}*],[\bullet X \bullet]} \oplus [_{vP}\ DP_{[CASE:\mathit{int}]}\ [_{v'}\ DP_{[CASE:\square]}\ [_{v'}\ v \ldots t \ldots]]]$ <br> Workspace = $\{\ldots\}$ | SHB | AC | MC |
|---|---|---|---|
| $O_1$: $[_{TP}\ DP_{[CASE:\mathit{int}]}\ [_{T'}\ T\ [_{vP}\ t'\ [_{v'}\ DP_{[CASE:\mathit{ext}]}\ [_{v'}\ v \ldots t \ldots]]]]]$ | *! | | |
| ☛ $O_2$: $[_{TP}\ DP_{[CASE:\mathit{ext/int}]}\ [_{T'}\ T\ [_{vP}\ t'\ [_{v'}\ DP_{[CASE:\square]}\ [_{v'}\ v \ldots t \ldots]]]]]$ | | | * |

## 6. Accelerating Merge: A Constraint on Dative Movement in Accusative Systems

*6.1 Data*

*Observation* (Höhle (1978, 56-57), Thiersch (1978, 168-169), Fanselow (1986, 4), Grewendorf (1989, 150), Fanselow (1990, 121)):
A $DP_{dat}$ object cannot be extracted from an ECM complement in German.

(26) *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

(27) *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



(28) *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

(29) *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

*Hypothesis*:
This restriction has the same source as the ban on ergative movement in ergative systems: The dative arguments move too early, and thus maraud the matrix v's [*CASE:*int**] feature, thereby precluding accusative case assignment to the ECM subject.

*Complication*:
German is an accusative language that has a ranking AC ≫ 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.

*Proposal*:
Movement of the dative DP is accelerated by a higher-ranked constraint in this particular context, one that regulates proper and improper movement.

(30) *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

(31) *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 ]

(32) *A dilemma for improper movement, given that every XP is a phase*:

a. Welches $\text{Buch}_1$ hat [$_{\text{vP}}$ $t_1'''$ Karl gemeint [$_{\text{CP}}$ $t_1''$ dass [$_{\text{vP}}$ $t_1'$ jeder $t_1$
which $\text{book}_{acc}$ has Karl meant that everyone
lesen möge ]]] ?
read should

b. *dass Karl [$_{\text{vP}}$ das $\text{Buch}_1$ glaubt [$_{\text{CP}}$ $t_1''$ dass [$_{\text{vP}}$ $t_1'$ keiner $t_1$
that $\text{Karl}_{nom}$ the $\text{book}_{acc}$ thinks that $\text{no-one}_{nom}$
liest ]]]
reads

### 6.2 *Assumptions*

#### 6.2.1 *Improper Movement*

*Ref.*: Müller (2012)

*Basic assumptions about improper movement*:
(i) Edge features are defective copies of categorial features of phase heads.
(ii) Edge features successively value movement-related features of moved items, creating lists that record aspects of the derivational history of movement.
(iii) Such information is deleted when information of the same type is encountered.
(iv) When a criterial landing site is reached, the functional sequence of categories (*f-seq*: C-T-v-V) must be respected on such lists; see Williams (1974; 2003). This constraint will be called the WILLIAMS CYCLE.

(33) *Legitimate long-distance wh-movement*:
$\text{What}_2$ do you think [$_{\text{CP}}$ C [$_{\text{TP}}$ $\text{she}_1$ T [$_{\text{vP}}$ $t_1$ v [$_{\text{VP}}$ said $t_2$ ]]]] ?
[$_{\text{CP}}$ what$_{[\text{wh:}\boxed{\text{CTvV}\cancel{\text{CTvV}}}]}$ [$_{\text{C}'}$ C [$_{\text{TP}}$ you think she said ]]] ($\surd$f-seq)

(34) *Illegitimate long-distance scrambling*:
*dass Karl das Buch glaubt [$_{\text{CP}}$ dass keiner $t_1$ liest ]
that $\text{Karl}_{nom}$ the $\text{book}_{acc}$ thinks that $\text{no-one}_{nom}$ 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)

#### 6.2.2 *Exceptional Case Marking in German*

*Observation*:
German ECM complements lack typical properties associated with TP (or CP) (Stechow & Sternefeld (1988), Fanselow (1991), Wurmbrand (2001)). There is no separate temporal specification, there is obligatory wide scope for negation, and a systematic absence of *zu* ('to'): XP in the above examples is vP.

(35) *German ECM complements are vPs*:

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



b. Wir lassen [$_{\text{vP}}$ den Diener den Mann nicht schlagen ]
we let the $\text{servant}_{acc}$ the $\text{man}_{acc}$ not hit

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

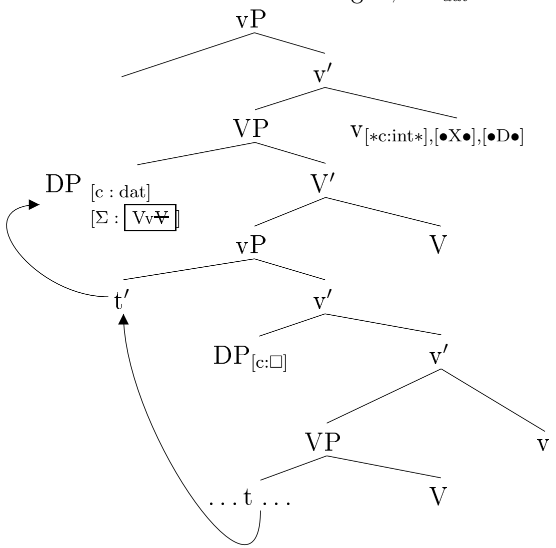
*Note*:

- With extraction from ECM complements, matrix v has a dual role: It assigns case to the ECM subject ([∗CASE:*int*∗]), and it has an edge feature ([•X•]) to effect the intermediate movement step.
- On the vP cycle, a moved dative DP originating in the embedded ECM complement (and having been assigned lexical case there by V) has a chance to immediately remedy temporary f-seq violations on the feature list of its movement-related feature, and thereby satisfy the WILLIAMS CYCLE quickly. Assuming that this constraint outranks AC in German, movement of the dative DP to Specv will have to precede case assignment by v to the embedded $\text{DP}_{ext}$ in German.
- This gives rise to maraudage, and the derivation will ultimately crash because the embedded $\text{DP}_{ext}$'s case feature remains unvalued.
- The specific version of the WILLIAMS CYCLE (WC) that is required is given in (36). It is formulated such that a temporary violation can be initiated without violating WC, which must be possible given that WC is ranked high (above AC). (Also see Alderete (2001) on anti-faithfulness constraints in phonology; and Baković & Wilson (2000) on targeted constraints.)

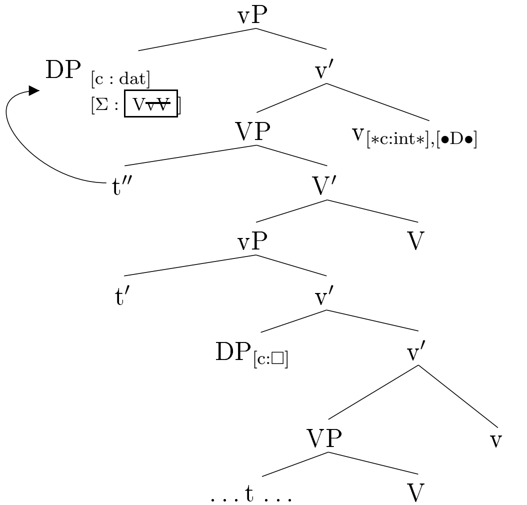
(36) WILLIAMS CYCLE (WC):
If categorial information on a list of a movement-related feature does not conform to f-seq (C-T-v-V) in the input, it must conform to f-seq in the output.

(37) *Illegitimate movement of $DP_{dat}$ from ECM complements*

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

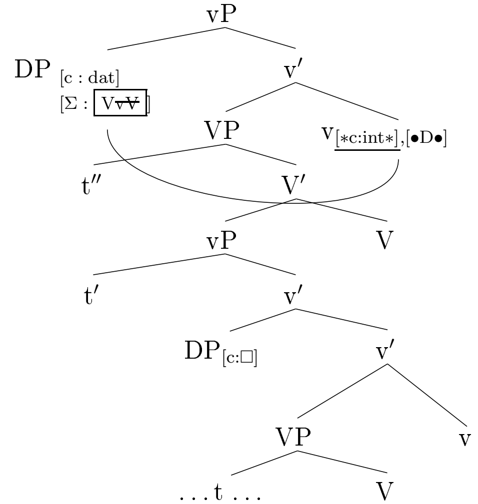



b. WC ≫ AC ≫ MC triggers movement of $DP_{dat}$ to Specv






c. SHB triggers maraudage of v




*Note*:
To trigger early dative movement in (37-b), it is thus crucial that ECM complements in German are vPs, not TPs (or V clusters, for that matter).

*Consequence*:
Since dative movement in ECM contextsa applies too early (because of the WC), it will maraud matrix v's case feature, and the embedded $DP_{ext}$ will permanently remain without a valued case feature; the derivation crashes.

*$T_{18}$: Dative movement in ECM contexts, step 1: Move*

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

*Note*:
The derivation continues by merging the matrix $DP_{ext}$; however, it will ultimately crash because the case feature of the embedded $DP_{ext}$ can never be valued.

## *6.4 Extremely Local vs. Less Local Optimization*

*Note*:
This time, assuming larger optimization domains like the phrase, does not make a

$T_{19}$*: Dative movement in ECM contexts, step 2: Agree (maraudage)*

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

wrong prediction: At the vP phrase level, WC, MC and AC are all satisfied, and SHB will continue to pick a maraudage output.

*6.5 Consequences*

*(i) Predictions*:

- If there is no embedded DP$_{ext}$ in what is otherwise the same construction, movement of the dative DP should be fine (there can be no maraudage). This is borne out; see the so-called *lassen*-passive construction in (38).
- If the dative DP and the embedded DP$_{ext}$ both undergo movement in ECM construction, the result should also be fine. As a tendency, this may also seem to be right; see (39).

(38) *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

(39) *DP$_{dat}$ movement where the embedded DP$_{ext}$ also moves*

Diesen Jungen ließ der Frau keiner [$_{XP}$ $t_1$ $t_2$ helfen ]
this boy$_{acc}$ let the woman$_{dat}$ no-one$_{nom}$ help

*(ii) Why can dative DPs undergo extraction from ECM constructions with an unaccusative embedded predicate?*
The problem here is that the embedded DP$_{int}$ gets case from the matrix v and does not block dative movement; see (40) (from Fanselow (1990)). 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.



(40) *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

*(iii) What about extraction of other kinds of DPs from ECM contexts?*
DP$_{acc}$ can undergo such movement easily (see (41), (42)); DP$_{gen}$ movement is arguably much more restricted. This would follow without further ado if maraudage is blocked if exactly the same case is involved; and genitive and accusative are sufficiently different.

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

a. dass der Kollege [$_{DP}$ den Antrag ]$_1$ [$_{XP}$ seine Mitarbeiter $t_1$
that the colleague$_{nom}$ the proposal$_{acc}$ his co-workers$_{acc}$
gerade schreiben lässt ]
currently write lets

b. dass [$_{DP}$ den Antrag ]$_1$ der Kollege [$_{XP}$ seine Mitarbeiter $t_1$
that the proposal$_{acc}$ the colleague$_{nom}$ his co-workers$_{acc}$
gerade schreiben lässt ]
currently write lets

(42) *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

(43) *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
that they$_{gen}$/the dead$_{gen}$ no-one$_{nom}$ the boy$_{acc}$ commemorate
sieht/lässt
sees/lets

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

## 7. Decelerating Merge: Mobility of Lexical/Oblique Arguments in Ergative Systems

*Starting point*:
Not all morphologically ergative languages ban extraction of the ergative. In some, the ergative extracts freely and without any special morphology (such as the agent focus morphology encountered in many Mayan languages).

(44) *Wh-movement of DP$_{erg}$ in Chol* (Mayan, 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?'

(45) *A-bar movement of DP$_{erg}$ in Basque* (Hualde & Ortiz de Urbina (2003))

a. Jonek$_1$ omen daki t$_1$ hori.
JON.ERG apparently knows that
'Apparently, JON knows that.'

b. Nork$_1$ lagunduko die t$_1$ sure lagunei?
who.ERG help.FUT AUX your friends.DAT
'Who will help your friends?'

c. [ Ø.ERG$_1$ t$_1$ Bonaparteren lanari ekin zioten ] langile
Bonaparte.GEN work.DAT tackle AUX worker
apalak
humble.DET.PL
'the humble workers who took on the work of Bonaparte'

(46) *Relativization of DP$_{erg}$ in Avar* (Nakh-Dagestanian, Polinsky et al. (2011))

[ Ø.ERG$_1$ t$_1$ ʕoloqana-y yas repetici-yal-de y-ač:-un
unmarried-II girl.ABS rehearsal-OBL-LOC II-bring-GER
y-ač'-ara-y ] artistka$_1$ bercina-y y-igo.
II-come-PRTC-II actress.ABS beautiful-II II-AUX

'The actress that brought the young girl to the rehearsal is pretty.'

(47) *Relativization of DP$_{erg}$ in* Pitjantjatjara (Pama-Nyungan, Bowe (1990, 101)):

Wati panya [ Ø.ERG$_1$ t$_1$ wa<u>r</u>u atu-ntja-lu ] ngayu-nya u-ngu.
man ANAPH wood chop-INF-ERG 1.SG-ACC give-PAST

'The man who chops wood gave me some.'

*Possible explanation:*
Extraction of the ergative subject is an option in these languages because Agree applies before Merge on the TP-cycle despite the 'ergative ranking' MC ≫ AC.

*Basic assumption*:
The ergative in Chol, Avar, Basque, and Pitjantjatjara is not structural but lexical (see Nash (1996), Alexiadou (2001), Woolford (1997; 2001), Legate (2008), Sheehan (2013) for related claims).



(48) *Decomposition of case features* (Bierwisch (1967), Wiese (1999), work in Distributed Morphology):

a. [∗CASE:*int*∗] has two subfeatures [–OBL], [+GOV]: [∗CASE:–OBL,+GOV∗]
b. [∗CASE:*ext*∗] has two subfeatures [–OBL], [–GOV]: [∗CASE:–OBL,–GOV∗]
c. [±GOV] maintains the external/internal distinction; [–OBL] indicates that the cases associated with T and v are structural (non-oblique).
d. DP arguments bear [OBL:□], [GOV:□] before valuation.
e. Structural ergative case: Both subfeatures are on v.
f. Lexical ergative case: [∗CASE:–OBL∗] is on v, [∗CASE:+GOV∗] is on V.

(49) *Structural vs. lexical case assignment*:

a. [∗CASE:–OBL∗] is discharged by Agree under m-command (as before).
b. [∗CASE:+GOV∗] is discharged under c-command.

*Activitiy*:
Arguments with partially valued case are inactive (cf. Richards (2008); also Chomsky's (2001) 'Activity Condition').

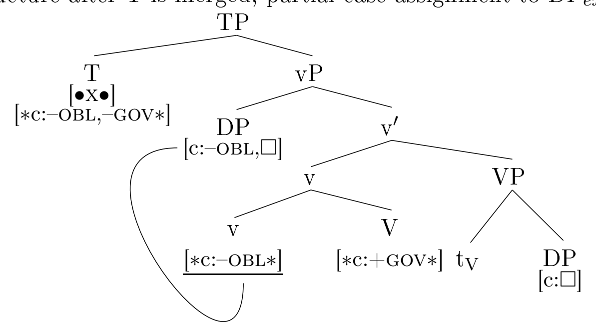
(50) ACTIVITY CONDITION (ACTC):
Inactive elements cannot undergo internal Merge.

(51) *Ranking:*
ACTC ≫ MC ≫ AC

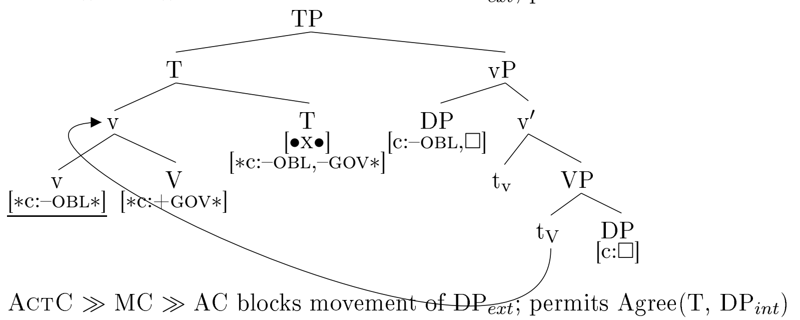
*Conclusion*:
Assignment of lexical ergative is only possible if v+V moves to T, either because this generally happens in the language, or as a repair strategy. Further proviso: Successive head movement maintains c-command (cf., e.g., Roberts (2010)).

(52) *Legitimate movement of DP$_{erg}$ if the ergative is lexical*

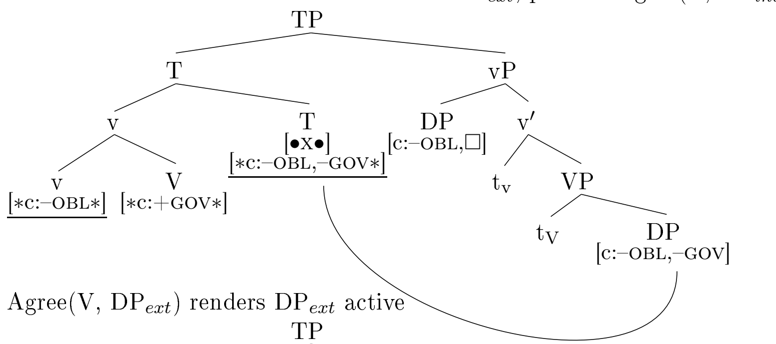
a. Structure after T is merged; partial case assignment to DP$_{ext}$



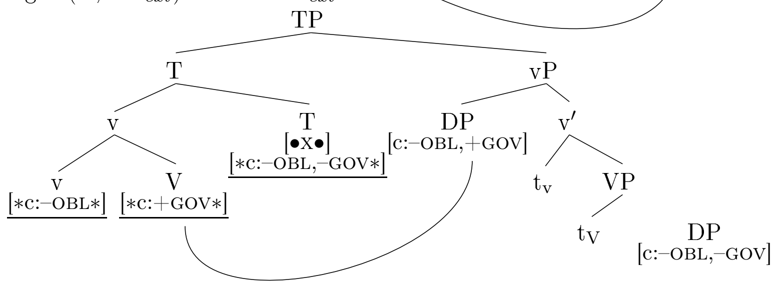
b. ACTC ≫ MC ≫ AC blocks movement of DP$_{ext}$; permits movement of v-V



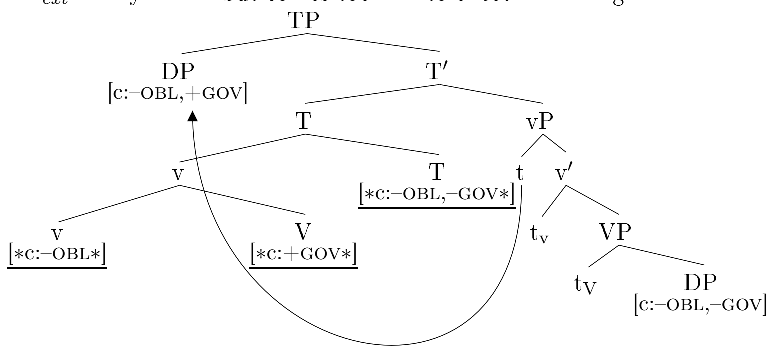

c. ACTC ≫ MC ≫ AC blocks movement of DP$_{ext}$; permits Agree(T, DP$_{int}$)




d. Agree(V, DP$_{ext}$) renders DP$_{ext}$ active




e. DP$_{ext}$ finally moves but comes too late to effect maraudage






*Note*:
The relevant competitions are given in tableaux T$_{20}$, T$_{21}$, and T$_{22}$. Note that there are four locally optimal continuations O$_2$–O$_5$ (that all carry out an Agree operation) in T$_{20}$ (which illustrates the crucial step from (52-b) to (52-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$_{int}$ marauds T's features as well).

*T$_{20}$: Lexical ergative movement, step 1: Agree(T,DP$_{int}$)*

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

Assuming a continuation with O$_2$, tableau T$_{21}$ shows that the situation is still such that DP$_{ext}$ cannot move without fatally violating ACTC.

*T$_{21}$: Lexical ergative movement, step 2: Agree(V,DP$_{ext}$)*

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

Finally, tableau T$_{22}$ illustrates the trivial final competition on the TP cycle: DP$_{ext}$ is now active, and movement can finally be carried out.

$T_{22}$: *Lexical ergative movement, step 3: Move(T,DP$_{ext}$)*

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

*7.1 Extremely Local vs. Less Local Optimization*

$T_{23}$: *TP optimization under* ActC $\gg$ *MC* $\gg$ *AC ranking: wrong result*

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

## 8. Ranked Constraints vs. Parametrized Preference Principles

*Question*:
How could acceleration and deceleration effects be addressed in a non-optimality-theoretic approach, such as the one laid out in lecture 2?

*Implausible preference principles*:

- Agree before Move, unless satisfaction of the Williams Cycle demands otherwise.
- Move before Agree unless satisfaction of the Activity Condition demands otherwise.

*Problem here*:
Constraint interaction is ignored as a phenomenon, and is simply built into the formulation of a single constraint. (See the remarks on Barss's (1984) chain-binding approach to movement and reflexivization in lecture 1.)

*Possible solution*:

1. Only *convergent* steps are considered (Chomsky (1995)).
2. Violations of WC and ActC lead to non-convergence.
3. However: It is far from obvious that a simple notion of convergence can be devised that covers all relevant contexts in a natural way; see Collins (1994), Sternefeld (1996).



## 9. Conclusion

*Four claims*:

- *(i) Competition of Merge and Agree*:
  1. vP cycle: v[*CASE:*int**],[•D•]
  2. TP cycle: T[*CASE:*int**],[•X•]
- *(ii) Ranking vs. preference principle* (see Chomsky (1995; 2001) on Merge before Move):
  1. MC $\gg$ AC on the vP cycle: ergative encoding system
  2. AC $\gg$ MC on the vP cycle: accusative encoding system
  3. MC $\gg$ AC on the TP cycle: *ergative movement (too early), $\surd$absolutive movement
  4. AC $\gg$ MC on the TP cycle: $\surd$accusative movement, $\surd$nominative movement
- *(iii) Acceleration and deceleration of movement*: expected in OT, less so in Minimalism
  1. WC $\gg$ AC $\gg$ MC on the vP cycle: *dative movement in German ECM (too early)
  2. ActC $\gg$ MC $\gg$ AC on the TP cycle: $\surd$ergative movement (late)
- *(iv) Extremely local vs. less local optimization* (opacity): If the domain is larger than the derivational step, then
  1. AC $\gg$ MC on the vP cycle does not derive accusative encoding systems;
  2. AC $\gg$ MC on the TP cycle wrongly blocks accusative movement;
  3. ActC $\gg$ MC $\gg$ AC on the TP cycle cannot circumvent the ban on ergative movement.

*Take-home message*:

Good things come to those who wait.

## Bibliography


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

Ackema, Peter & Ad Neeleman (1998): Optimal Questions, *Natural Language and Linguistic Theory* 16, 443–490.

Adger, David & 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 & F. Ordoñez, eds., *Clitics and Affix Combinations*. Benjamins, Amsterdam, pp. 199–235.
Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller & Philipp Weisser (2012): Ergatives Move Too Early. Ms., Universität Leipzig (*Linguistische Arbeits Berichte* 90, 2013). To appear in *Syntax*.
Baković, Eric & Colin Wilson (2000): Transparency, Strict Locality, and Targeted Constraints. In: R. Billerey & B. D. Lillehaugen, eds., *WCCFL 19 Proceedings*. Cascadilla Press, Somerville, MA, pp. 43–56.
Barss, Andrew (1984): Chain Binding. Ms., MIT, Cambridge, Mass. (Talk at GLOW).
Béjar, Susana & 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.
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.
Broekhuis, Hans (2000): Against Feature Strength: The Case of Scandinavian Object Shift, *Natural Language and Linguistic Theory* 18, 673–721.
Broekhuis, Hans (2006): Derivations (MP) and Evaluations (OT). In: R. Vogel & H. Broekhuis, eds., *Optimality Theory and Minimalism: A Possible Convergence?*. Vol. 25, Linguistics in Potsdam, Potsdam, pp. 137–193.
Broekhuis, Hans (2008): *Derivations and Evaluations. Object Shift in the Germanic Languages*. Mouton de Gruyter, Berlin.
Broekhuis, Hans & Joost Dekkers (2000): The Minimalist Program and Optimality Theory: Derivations and Evaluation. In: J. Dekkers, F. van der Leeuw & J. van de Weijer, eds., *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford University Press, Oxford, pp. 386–422.
Bye, Patrick (2001): Virtual Phonology: Rule Sandwiching and Multiple Opacity in North Saami. PhD thesis, University of Tromsø.
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 (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.



Chomsky, Noam (1975): *The Logical Structure of Linguistic Theory*. Plenum Press, New York.
Chomsky, Noam (1986): *Barriers*. MIT Press, Cambridge, Mass.
Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.
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 (2008): On Phases. In: R. Freidin, C. Otero & M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.
Collins, Chris (1994): Economy of Derivation and the Generalized Proper Binding Condition, *Linguistic Inquiry* 25, 45–61.
Coon, Jessica (2010): Complementation in Chol (Mayan). A Theory of Split Ergativity. PhD thesis, MIT, Cambridge, Mass.
Coon, Jessica, Pedro Mateo Pedro & 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 & T. Daniel Seely (2002): Rule Applications as Cycles in a Level-Free Syntax. In: S. D. Epstein & 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 & W. Sternefeld, eds., *Scrambling and Barriers*. Benjamins, Amsterdam, pp. 113–140.
Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.
Fanselow, Gisbert & Damir Ćavar (2001): Remarks on the Economy of Pronunciation. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton de Gruyter, Berlin, pp. 107–150.
Fischer, Silke (2004): Towards an Optimal Theory of Reflexivization. PhD thesis, Universität Tübingen.
Fischer, Silke (2006): Matrix Unloaded: Binding in a Local Derivational Approach, *Linguistics* 44, 913–935.
Georgi, Doreen (2012a): A Local Derivation of Global Case Splits. In: A. Alexiadou, T. Kiss & G. Müller, eds., *Local Modelling of Non-Local Dependencies in Syntax*. Linguistische Arbeiten, De Gruyter, Berlin, pp. 306–336.
Georgi, Doreen (2012b): Opaque interaction of Merge and Agree: On Two Types of Internal Merge. Ms., Universität Leipzig.
Georgi, Doreen, Fabian Heck & 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).
Goldrick, Matt (2000): Turbid Output Representations and the Unity of Opacity. In:

M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim, eds., *Proceedings of NELS 30*. GLSA, Amherst, Mass, pp. 231–345.

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

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

Grimshaw, Jane (1997): Projection, Heads, and Optimality, *Linguistic Inquiry* 28, 373–422.

Heck, Fabian (1998): Relativer Quantorenskopus im Deutschen – Optimalitätstheorie und die Syntax der Logischen Form. Master's thesis, Universität Tübingen.

Heck, Fabian (2001): Quantifier Scope in German and Cyclic Optimization. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 175–209.

Heck, Fabian & Gereon Müller (2000): Successive Cyclicity, Long-Distance Superiority, and Local Optimization. In: R. Billerey & B. D. Lillehaugen, eds., *Proceedings of WCCFL*. Vol. 19, Cascadilla Press, Somerville, MA, pp. 218–231.

Heck, Fabian & Gereon Müller (2003): Derivational Optimization of Wh-Movement, *Linguistic Analysis* 33, 97–148. (Volume appeared 2007).

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

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

Heck, Fabian & Gereon Müller (2013b): On Accelerating and Decelarating Movement: From Minimalist Preference Principles to Harmonic Serialism. In: *Rule Interaction in Grammar*. Vol. 90 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig, pp. 511–558.

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

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

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.

Lahne, Antje (2008): Excluding SVO in Ergative Languages. In: F. Heck, G. Müller & J. Trommer, eds., *Varieties of Competition*. Vol. 87 of *Linguistische Arbeitsberichte*, Universität Leipzig, pp. 65–80.

Lahne, Antje (2009): Where There is Fire There is Smoke. Local Modelling of Successive-Cyclic Movement. PhD thesis, Universität Leipzig.

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.

Legendre, Géraldine, Paul Smolensky & Colin Wilson (1998): When is Less More? Faithfulness and Minimal Links in Wh-Chains. In: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds., *Is the Best Good Enough?*. MIT Press and MITWPL, Cambridge, Mass., pp. 249–289.

McCarthy, John (1999): Sympathy and Phonological Opacity, *Phonology* 16:3, 331–399.

McCarthy, John (2000): Harmonic Serialism and Parallelism. In: M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim, eds., *Proceedings of NELS 30*. GLSA, Amherst, Mass., pp. 501–524.

McCarthy, John (2007a): *Hidden Generalizations. Phonological Opacity in Optimality Theory*. Equinox, London.

McCarthy, John (2007b): *Hidden Generalizations. Phonological Opacity in Optimality Theory*. Equinox, London.

McCarthy, John (2008): The Serial Interaction of Stress and Syncope, *Natural Language and Linguistic Theory* 26, 499–546.

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

Müller, Gereon (1997): Partial Wh-Movement and Optimality Theory, *The Linguistic Review* 14, 249–306.

Müller, Gereon (2000a): Das Pronominaladverb als Reparaturphänomen, *Linguistische Berichte* 182, 139–178.

Müller, Gereon (2000b): Shape Conservation and Remnant Movement. In: M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim, eds., *Proceedings of NELS 30*. GLSA, Amherst, Mass., pp. 525–539. Revised and extended version appeared 2002 in Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, & Hans-Martin Gärtner (eds.), *Dimensions of Movement*, 209-241. Amsterdam: Benjamins.

Müller, Gereon (2002): Harmonic Alignment and the Hierarchy of Pronouns in German. In: H. Simon & H. Wiese, eds., *Pronouns: Grammar and Representation*. Benjamins, Amsterdam, pp. 205–232.

Müller, Gereon (2003): Local vs. Global Optimization in Syntax: A Case Study. In: J. Spenader, A. Eriksson & Ö. Dahl, eds., *Variation within Optimality Theory. Proceedings of the Stockholm Workshop*. Stockholm University, Department of Linguistics, pp. 82–91.

Müller, Gereon (2004): Argument Encoding and the Order of Elementary Operations. Ms., Universität Leipzig.

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.

Oostendorp, Marc van (2006): A Theory of Morphosyntactic Colours. Ms., Meertens Institute, Amsterdam.

Oostendorp, Marc van (2007): Derived Environment Effects and Consistency of Exponence. In: S. Blaho, P. Bye & M. Krämer, eds., *Freedom of Analysis?*. Mouton de Gruyter, Berlin, pp. 123–148.

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

Pesetsky, David (1998): Some Optimality Principles of Sentence Pronunciation. In: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds., *Is the Best Good Enough?*. MIT Press and MITWPL, Cambridge, Mass., pp. 337–383.

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

Prince, Alan & Paul Smolensky (1993): Optimality Theory. Constraint Interaction in Generative Grammar. Book ms., Rutgers University.

Prince, Alan & Paul Smolensky (2004): *Optimality Theory. Constraint Interaction in Generative Grammar*. Blackwell, Oxford.

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 & S. Fischer, eds., *Agreement Restrictions*. Mouton de Gruyter, pp. 181–213. [P5].

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

Salzmann, Martin (2012): On Three Types of Variation in Resumption: Evidence in Favor of Violable and Ranked Constraints. In: H. Broekhuis & R. Vogel, eds., *Linguistic Derivations and Filtering*. Equinox, Sheffield, pp. 76–108.

Sheehan, Michelle (2013): Towards a Parameter Hierarchy for Alignment. Ms., University of Cambridge. To appear in *Proceedings of WCCFL* 31, Arizona State University.

Smolensky, Paul & Geraldine Legendre (2006): *The Harmonic Mind*. MIT Press, Cambridge, Mass.

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

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

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

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

Trommer, Jochen (2011): Phonological Aspects of Western Nilotic Mutation Morphology. Habilitationsschrift, Universität Leipzig.



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.

Wilson, Colin (2001): Bidirectional Optimization and the Theory of Anaphora. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 465–507.

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 & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 509–543.

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