Extremely Local Optimization

Gereon Müller

Institut für Linguistik
Universität Leipzig

Beschränkungen für Bewegung
004-1006
www.uni-leipzig.de/∼muellerg

Harmonic Serialism

- Harmonic serialism in phonology:
  McCarthy (2008; 2010), McCarthy et al. (2012), Kimper (2016), Pater (2016), ...
  (also see Prince & Smolensky (1993; 2004) for the general option, and McCarthy (2000) for an early negative assessment)
- Harmonic serialism in morphology:
  Caballero & Inkelas (2013) Müller (2018)
- Harmonic serialism in syntax:
  Heck & Müller (2007; 2013; 2016), Assmann, Georgi, Heck, Müller & Weisser (2015), Müller (2004), Lahne (2008; 2009), Georgi (2009), Murphy (2017) (predecessors: Ackema & Neeleman (1998), Heck (1998; 2001), Heck & Müller (2000; 2003a))

Optimization Domains

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?

But wait:
Isn't optimality theory inherently representational/non-derivational?

Answer:
No, it isn't.

> 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 harmonic 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)

Harmonic Serialism vs. Derivations as Candidates

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. "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. 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)

Optimization domains

(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 (AgrOP), DP): multiple optimization, derivational)
d. phrase (XP: multiple optimization, derivational)
e. derivational step (multiple optimization, derivational) = true harmonic serialism

Background:

(i) Classical assumption: The whole sentence is subject to a single, parallel optimization procedure (Grimshaw (1997), Pesetsky (1998), Legendre, Smolensky & Wilson (1998) etc.). The output candidates are usually taken to be representations; but they can also be full derivations (as, e.g., in Müller (1997)).
(ii) Wilson (2001), Heck (1998; 2001): multiple optimization of whole sentences (still global).
(iii) Mutltiple optimization of smaller optimization domains: closely related to developments in the minimalist program.

Types of Arguments

Observation:
Small optimization domains presuppose a derivational approach to syntax.

Conceptual argument for small optimization domains:
The smaller the optimization domain is, the more the complexity of the overall system is reduced (reduction of the size of candidate sets).

Conceptual argument for larger optimization domains:
The larger the optimization domain is, the less often optimization procedures have to be carried out.

Empirical arguments for smaller/larger optimization domains:
If the ranked constraints have access to more/less structure, a wrong winner is predicted.

Proposals for Local Optimization

(2) 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; 2003b) on wh-movement, superiority, quantifier raising, and sluicing

d. Derivational step:
Heck & Müller (2007) on gender agreement with dative possessors in German DPs and expletives in German verb-second clauses; Müller (2004) on ergative and accusative argument encoding patterns; Lahne (2008; 2009) on excluding SVO in ergative languages; Georgi (2009) on global case splits in Tauya.

Clauses as optimization domains

Ref.: Ackema & Neeleman (1998)

(3) Long multiple wh-movement in Czech proceeds without wh-cluster formation:

[$_{VP}$ Co$_1$ [$_{VP}$ podle tebe [$_{VP}$ komu$_2$ [$_{VP}$ Petr řekl [$_{CP}$ že Jan dal t_1 t_2]]]]]
what according to you whom Petr said that Jan gave

The proposal:
Evaluation of movement constraints proceeds cyclically. That is to say, STAY is first evaluated with respect to the embedded clause, then to the combination of the embedded clause and the matrix clause. In the embedded clause, STAY favours separate movement of the two wh-expressions [...] This means that clustering can only take place when the larger cycle is taken into account, i.e., when the two whs have already been adjoined to the embedded VP. However, it is no longer possible then, because it would have to take place within the embedded clause (the initial landing site of the whs), which would go against strict cyclicity.

Ackema & Neeleman (1998, fn. 25)

Ackema & Neeleman's constraints

(4) a. Q-MARK:
Assign [+Q] to a propositional constituent.
(This can only be done by an overt functional head, which in turn needs to inherit this capacity in the matrix clause from some wh-phrase in its specifier.)

b. Q-SCOPE:
[+Q]-elements must c-command the constituent representing the proposition.

c. STAY:
Every node crossed by movement induces a violation.

T_1: Long multiple wh-movement in Czech, optimization of embedded CP

Input: part of the numeration	Q-Scope	Stay	Q-Mark
☞ O_1: $[_{CP}$ že $[_{VP}$ co$_1$ $[_{VP}$ komu$_2$ $[_{VP}$ Jan dal t_1 t_2]]]]		***	
O_2: $[_{CP}$ komu$_2$ že $[_{VP}$ co$_1$ $[_{VP}$ Jan dal t_1 t_2]]]]		****!*	
O_3: $[_{CP}$ co$_1$ komu$_2$ že $[_{VP}$ Jan dal t_1 t_2]]		****!***	
O_4: $[_{CP}$ že $[_{VP}$ Jan dal co$_1$ komu$_2$]]	*!		

T_2: Long multiple wh-movement in Czech, optimization of matrix clause

Input: $[_{CP}$ že $[_{VP}$ co$_1$ $[_{VP}$ komu$_2$ $[_{VP}$ Jan dal t_1 t_2]]]], Petr, řekl	Q-Scope	Stay	Q-Mark
☞ O_{11}: $[_{VP}$ co$_1$ $[_{VP}$ komu$_2$ $[_{VP}$ Petr řekl $[_{CP}$ že $[_{VP}$ t_1 $[_{VP}$ t_2 $[_{VP}$ Jan dal t_1 t_2]]]]]]]		*** ******	*
O_{12}: $[_{CP}$ co$_1$ řekl $[_{VP}$ komu$_2$ $[_{VP}$ Petr $[_{CP}$ že $[_{VP}$ t_1 $[_{VP}$ t_2 $[_{VP}$ Jan dal t_1 t_2]]]]]]]		*** *******!**	
O_{13}: $[_{CP}$ co$_1$ komu$_2$ řekl $[_{VP}$ (t_1 t_2) Petr $[_{CP}$ že $[_{VP}$ t_1 $[_{VP}$ t_2 $[_{VP}$ Jan dal t_1 t_2]]]]]]		*** *******!**	

Notation:
O_{11}-O_{13} in T_2 are descendants of O_1.

Note:

Global optimization of the whole sentence would predict a wrong winner: "It seems to be predicted that when the distance to be covered by the wh-expressions in a multiple question increases, clustering [as in Bulgarian, with a high-ranked Q-Mark] will be favoured."

T_3: Global optimization: Long multiple wh-movement in Czech, wrong winner

Input: numeration	Q-Scope	Stay	Q-Mark
O_1: [$_{VP}$ co$_1$ [$_{VP}$ komu$_2$ [$_{VP}$ Petr řekl [$_{CP}$ že [$_{VP}$ Jan dal t_1 t_2]]]]]		*********!*	*
☛O_2: [$_{CP}$ co$_1$ komu$_2$ řekl [$_{VP}$ Petr [$_{CP}$ že [$_{VP}$ Jan dal t_1 t_2]]]]		********	

Underlying logic:

(i) Two short movements are better than a short movement and a longer movement:

$2+2=4$, $1+5=6$.

(ii) Two medium-sized movements can be worse than a short movement and a very long movement:

$7+7=14$, $1+10=11$.

Harmonic Serialism: Derivational Steps as Optimization Domains

Assumption:
Derivational Steps qualify as the optimization domains.

Ref.: Heck & Müller (2007; 2013; 2016), Müller (2004; 2009)

Premise:
Minimalist program and optimality theory can be combined (see Pesetsky (1998), Broekhuis & Dekkers (2000), Broekhuis (2000; 2006; 2008), Fanselow & Ćavar (2001), Heck & Müller (2000), among others).

Two constitutive properties:
(i) MP: Syntactic structure is built up derivationally.
(ii) OT: Well-formedness of syntactic objects is determined via optimization.

Combining the properties:

(i) Syntactic structure is built up derivationally and is subject to repeated *local* optimization: structure building operations and optimization apply in a cyclic interleaving fashion.
(ii) Based on a given input, the operations Merge, Move, and Agree create various output candidates α_1,...,α_n: the candidate set M. M is subject to optimization.
(iii) The optimal output α_i serves as the input for the next cycle, and so on, until the numeration is empty.

Aim:
Pursue the consequences of the most radical position within a theory of local optimization: *extremely local optimization*.

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

(i) 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.

(ii) 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.

(iii) 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.

Constraints, Features, and Operations

(5) Two types of features that drive operations (see Sternefeld (2003), Adger (2003)):

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

b. Probe features trigger Agree: [$*$F$*$].

(6) Merge:
α can be merged with 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. (But cf. the following section for a qualification.)

(8) Agree:
α can agree 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. (This permits an Agree relation between a head and its specifier.)

Strict Cycle Condition

(11) Strict Cycle Condition (SCC, Chomsky (1973; 1993)):
Merge of α and β is possible only if β has no active features.
(A feature is active if it is a [•F•] or [$*$F$*$] feature that has not yet participated in Merge or Agree).

(12) Last Resort (LR):
Move of α and β is accompanied by Agree of α and β.

Comment:
The (perhaps less ordinary) treatment of Move in (12) as a binary operation rests on the assumption that Move is Merge (with β internal to α), which is binary.

Empirical Evidence for Extremely Local Optimization 1: Argument Encoding

(13) Basic patterns of argument encoding:

a. Accusative marking

DP_{ext}-V_i	DP_{int}-V_i
DP_{ext}-V_t	DP_{int}-V_t
nom	acc

b. Ergative marking

DP_{ext}-V_i	DP_{int}-V_i
DP_{ext}-V_t	DP_{int}-V_t
erg	abs

Note:
DP_{ext} = external argument DP; DP_{int} = internal argument DP.
V_i = intransitive verb; V_t = transitive verb.

Assumptions about argument encoding

(i) There is one structural argument encoding feature: [case].
(ii) [case] can have two values: ext(ernal) and int(ernal) (determined with respect to vP).
(iii) [case:ext] = nominative/absolutive, [case:int] = accusative/ergative (see Murasugi (1992)).
(iv) [case] features figure in Agree relations involving T/v and DP, as in (14).

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

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

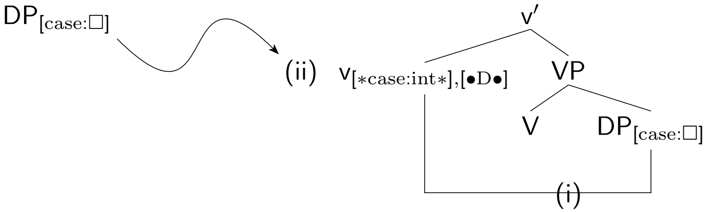
A dual role and a conflict

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 Σ 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 (i)), MC demands that it is Merge(DP_{ext},v) (see (ii)). (Application of these constraints at each derivational step derives the effects of Pesetsky's (1989) Earliness Principle, see Chomsky (2001, 15).)

Convention:
A feature [F] whose value is not yet determined is written as "[F:□]."

(15) Stage Σ:



Resolving the indeterminacy

Problem:
There is an indeterminacy in rule application that poses a problem for canonical minimalist approaches. What can be done?

- Rezáč (2004): Agree always applies before Merge.
- Epstein, Kitahara & Seely (2009): Merge always applies before Agree ('No Search before Search'), but Agree is confined ot c-command contexts. Hence, it must apply counter-cyclically.

(16) MINIMAL LINK CONDITION (Chomsky (1995; 2001)):
An Agree operation involving α and β can only take place if there is no δ such that (i) and (ii) hold:

a. δ is closer to α than β.

b. δ bears a feature that has not yet participated in Agree.

Specifiers vs. daughters of complements

(17) Closeness:
δ is closer to α than β if the path from δ to α is shorter than the path from β to α.

(18) Path (Müller (1998, 130); also cf. Pesetsky (1982, 289), Collins (1994, 56)):
The path from X to Y is the set of categories Z such that (a) and (b) hold:

a. Z is reflexively dominated by the minimal XP that dominates both X and Y.
b. Z dominates X or Y.

The length of a path is determined by its cardinality.

Consequences:
(i) The specifier and the complement of a head qualify as equally close to the head.
(ii) The specifier of a head is closer to the head than a category that is further embedded in the complement of the head.
(iii) DP_{ext} is now closer to v than DP_{int}.

AC vs. MC: Accusative patterns

Proposal:
This conflict of AC and MC is resolved by language-specific constraint ranking; the two possibilities yield accusative and ergative patterns of argument encoding.

(19) Rankings:

a. Accusative patterns:
(MLC $\gg$) $\mathrm{AC} \gg \mathrm{MC}$

b. Ergative patterns:
(MLC $\gg$) $\mathrm{MC} \gg \mathrm{AC}$

T$_4$: Accusative pattern, step 1 (Σ as input): Agree

Input: $[_{\mathrm{v'}}$ v$_{[*\mathrm{case:int}*],[\bullet\mathrm{D}\bullet]}$ ... DP$_{[\mathrm{case:}\square]}$ ... $]$ Workspace = $\{$DP$_{[\mathrm{case:}\square]}$, T$_{[*\mathrm{case:ext}*]}$, ...$\}$	MLC	AC	MC
O$_1$: $[_{\mathrm{v'}}$ DP$_{[\mathrm{case:}\square]}$ $[_{\mathrm{v'}}$ v$_{[*\mathrm{case:int}*]}$ ... DP$_{[\mathrm{case:}\square]}$ ... $]]$		*!	
☞ O$_2$: $[_{\mathrm{v'}}$ v$_{[\bullet\mathrm{D}\bullet]}$ ... DP$_{[\mathrm{case:int}]}$ ... $]$			*

T_5: Accusative pattern, step 2: Merge

Input: $[_{v'}$ v$_{[\bullet D \bullet]}$ ... DP$_{[case:int]}$ ... $]$ Workspace = $\{$DP$_{[case:\Box]}$, T$_{[*case:ext*]}$, ...$\}$	MLC	AC	MC
☞ O_{21}: $[_{vP}$ DP$_{[case:\Box]}$ $[_{v'}$ v ... DP$_{[case:int]}$ ... $]]$			

T_6: Accusative pattern, step 3: Merge

Input: $T_{[*case:ext*],[\bullet v\bullet]}$ + $[_{vP}$ $DP_{[case:\square]}$ $[_{v'}$ v … $DP_{[case:int]}$ …]] Workspace = { }	MLC	AC	MC
☞ O_{211}: $[_{T'}$ $T_{[*case:ext*]}$ $[_{vP}$ $DP_{[case:\square]}$ $[_{v'}$ v … $DP_{[case:int]}$ …]]]		*	

T_7: Accusative pattern, step 4: Agree

Input: $[_{T'}$ T$_{[*case:ext*]}$ $[_{vP}$ DP$_{[case:\square]}$ $[_{v'}$ v … DP$_{[case:int]}$ …]]] Workspace = { }	MLC	AC	MC
☞ O_{2111}: $[_{TP}$ T $[_{vP}$ DP$_{[case:ext]}$ $[_{v'}$ v … DP$_{[case:int]}$ …]]]			

T_8: Ergative pattern, step 1 (Σ as input): Merge

Input: $[_{v'}\ v_{[*case:int*],[\bullet D\bullet]}\ ...\ DP_{[case:\square]}\ ...\ ]$ Workspace = $\{DP_{[case:\square]},\ T_{[*case:ext*]},\ ...\}$	MLC	MC	AC
☞ O_1: $[_{v'}\ DP_{[case:\square]}\ [_{v'}\ v_{[*case:int*]}\ ...\ DP_{[case:\square]}\ ...\ ]]$			*
O_2: $[_{v'}\ v_{[\bullet D\bullet]}\ ...\ DP_{[case:int]}\ ...\ ]$		*!	

T_9: Ergative pattern, step 2: Agree (with DP_{ext})

Input: $[_{v'}$ DP$_{[case:\Box]}$ $[_{v'}$ v$_{[*case:int*]}$ ... DP$_{[case:\Box]}$ ...]] Workspace = $\{$T$_{[*case:ext*]}$, ...$\}$	MLC	MC	AC
☞ O$_{11}$: $[_{vP}$ DP$_{[case:int]}$ $[_{v'}$ v ... DP$_{[case:\Box]}$ ...]]			
O$_{12}$: $[_{vP}$ DP$_{[case:\Box]}$ $[_{v'}$ v ... DP$_{[case:int]}$ ...]]	*!		

T_{10}: Ergative pattern, step 3: Merge

Input: $T_{[*case:ext*],[\bullet v \bullet]}$ + $[_{vP}$ $DP_{[case:int]}$ $[_{v'}$ v ... $DP_{[case:\Box]}$ ...]] Workspace = { }	MLC	MC	AC
☞ O_{111}: $[_{T'}$ $T_{[*case:ext*]}$ $[_{vP}$ $DP_{[case:int]}$ $[_{v'}$ v ... $DP_{[case:\Box]}$ ...]]]			*

T_{11}: Ergative pattern, step 4: Agree

Input: $[_{T'}$ $T_{[*case:ext*]}$ $[_{vP}$ $DP_{[case:int]}$ $[_{v'}$ v … $DP_{[case:\square]}$ …]]] Workspace = { }	MLC	MC	AC
☞ O_{1111}: $[_{T'}$ T $[_{vP}$ $DP_{[case:int]}$ $[_{v'}$ v … $DP_{[case:ext]}$ …]]]			

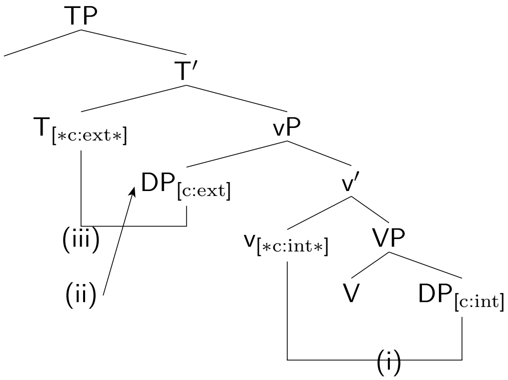
Note:
Agree in T_{11} is just local enough to be in accordance with the PIC in Chomsky (2001, 14)); cf. (20). (Also, DP_{ext} does not intervene, given the definition of the MLC).

(20) VP-internal nominative arguments in Icelandic:

Honum lík-a sjálf-s-elsk-ir leikar-ar
PRON.3.SG.MASC.DAT like-3.PL self-GEN-love-3.PL.NOM actor-PL.NOM
'He likes selfish actors.' (Sigurðsson (2002, 702))

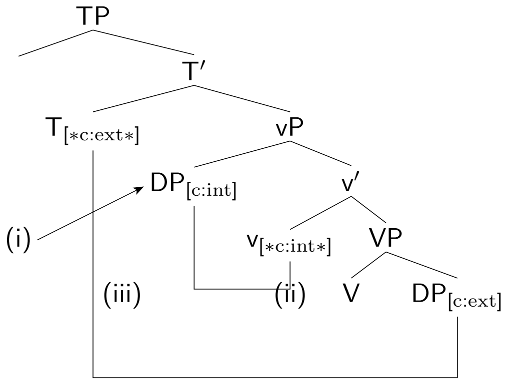
Agree before Merge

(21) a. Agree before Merge: accusative



Merge before Agree

b. Merge before Agree: ergative



Less local optimization

Suppose that the optimization domain is the phrase, the phase, the clause, or the whole sentence. Other things being equal, this makes wrong empirical predictions (a candidate chosen wrongly as optimal is marked ☛, as before): Given the MLC, [∗CASE:int∗] can never be instantiated on DP_{int}, but must be instantiated on DP_{ext}: Once DP_{ext} is part of the structure, Agree(v,DP_{int}) w.r.t. [case] is impossible.

T_{12}: vP optimization under MC ≫ AC ('ergative') ranking: right result

Input: $DP_{[case:\square]}$, $v_{[*case:int*],[\bullet D\bullet],[\bullet V\bullet]}$, $[_{VP}$ ... $DP_{[case:\square]}$ ... $]$ Workspace = $\{T_{[*case:ext*]}, ...\}$	MLC	MC	AC
☞ O_1: $[_{vP}$ $DP_{[case:int]}$ $[_{v'}$ v ... $DP_{[case:\square]}$ ... $]]$			
O_2: $[_{vP}$ $DP_{[case:\square]}$ $[_{v'}$ v ... $DP_{[case:int]}$ ... $]]$	*!		

T_{13}: vP optimization under AC ≫ MC ('accusative') ranking: wrong result

Input: $DP_{[case:\square]}$, $v_{[*case:int*],[\bullet D\bullet],...}$, $[_{VP}$ ... $DP_{[case:\square]}$ ... $]$ Workspace = $\{T_{[*case:ext*]}, ...\}$	MLC	AC	MC
☛ O_1: $[_{vP}$ $DP_{[case:int]}$ $[_{v'}$ v ... $DP_{[case:\square]}$ ... $]]$			
O_2: $[_{vP}$ $DP_{[case:\square]}$ $[_{v'}$ v ... $DP_{[case:int]}$ ... $]]$	*!		

Empirical Evidence for Extremely Local Optimization 2: Prenominal Dative Possessors in German

Observation:
(i) German exhibits a construction with a dative-marked possessor DP_2 in SpecD of a matrix DP_1 (see, e.g., Haider (1988), Zifonun (2004)).
(ii) D_1 is realized by a possessive pronoun.
(iii) The root of the pronoun agrees with DP_{dat} with respect to [num] and [gen].
(iv) The inflection of the pronoun agrees with its complement NP with respect to [num], [gen], and [case]. We focus here on agreement with respect to [gen] (see (22)), but everything can be transferred to the other features as well.

(22) Gender agreement with dative possessors in German:

a.	$[_{DP}$	dem	Fritz]	sein	-e	Schwester	"Fritz's sister"
		the.masc	Fritz	his.masc	-fem	sister.fem	
b.	$*[_{DP}$	dem	Fritz]	ihr	-Ø	Schwester	"Fritz's sister"
		the.masc	Fritz	her.fem	-masc	sister.fem	

Assumptions

Analysis:
(i) DP_{dat} is merged as a complement of the possessee (de Vries (2005)) and undergoes [$\bullet\text{EPP}\bullet$]-driven movement to SpecD.
(ii) Functional elements like pronouns are realized by post-syntactic morphology (see, e.g., Halle & Marantz 1993).
(iii) The pronoun's inflectional features occupy a structurally higher position than its root ($\sqrt{}$) features.

A Dual Role, Again

Consequence:

It follows that the pronoun has a dual role: It bears [$*$gen:$\square*$] probes that trigger Agree and an [$\bullet$EPP$\bullet$]-feature that triggers (internal) Merge. This causes a conflict. Suppose the derivation has reached stage Σ, where the pronoun has been merged. Then AC demands Agree(D,DP_{dat}) or Agree(D, NP); and MC demands DP_{dat} raising to SpecD. The conflict can be resolved by ranking AC over MC, yielding the correct agreement pattern.

(23) MLC $\gg$ AC $\gg$ MC $\gg$ LR (Ranking for German)

Analysis

Suppose we want to derive (22-a). We enter the derivation at stage Σ. Due to AC $\gg$ MC, Agree must apply first. Since the pronoun's inflectional probes are structurally higher than its root probes, the former count as closer to both NP and DP_{dat}. Thus the MLC constrains Agree to the inflectional probes. Moreover, the NP counts as closer to the pronoun than DP_{dat}. Thus Agree(NP,infl) instantiates [gen_{infl}:fem] on the pronoun (see O_1 in T_{14}). Having undergone Agree, the NP and the inflection are inactive. Hence, Agree can next affect the pronoun's root probes and DP_{dat}. This values [$gen_{\surd}$:masc] on the pronoun (see O_1 in T_{15}). Finally, MC can be satisfied by movement of the possessor DP to SpecD (this optimization is skipped).

T_{14}: Evaluation of gender inflection: Agree

Input: $[_{DP_1}$ $D_{[*case:dat*],[*gen_{\surd}:\square*],[*gen_{infl}:\square*],[\bullet EPP\bullet]}$ $[_{NP}$ $N_{[gen:fem]}$ $DP_{2[case:\square],[gen:masc]}$]]	MLC	AC	MC	LR
☞ O_1: $[_{DP_1}$ $D_{[*case:dat*],[*gen_{\surd}:\square*],[gen_{infl}:fem],[\bullet EPP\bullet]}$ $[_{NP}$ $N_{[gen:fem]}$ $DP_{2[case:\square],[gen:masc]}$]]		**	*	
O_2: $[_{DP_1}$ $D_{[*case:dat*],[gen_{\surd}:fem],[*gen_{infl}:\square*],[\bullet EPP\bullet]}$ $[_{NP}$ $N_{[gen:fem]}$ $DP_{2[case:\square],[gen:masc]}$]]	*!	**	*	

T_{15}: Evaluation of root's gender and possessor's case: Agree

Input: $[_{DP_1}\ D_{[*case:dat*],[*gen_{\surd}:\square*],[gen_{infl}:fem],[\bullet EPP \bullet]}\ [_{NP}\ N_{[gen:fem]}\ DP_{2[case:\square],[gen:masc]}\ ]]$	MLC	AC	MC	LR
☞ O_1: $[_{DP_1}\ D_{[case:dat],[gen_{\surd}:masc],[gen_{infl}:fem],[\bullet EPP \bullet]}\ [_{NP}\ N_{[gen:fem]}\ DP_{2[case:dat],[gen:masc]}\ ]]$			*	
O_2: $[_{DP_1}\ DP_{2[case:\square],[gen:masc]}\ D_{[*case:dat*],[gen_{infl}:fem],...}\ [_{NP}\ N_{[gen:fem]}\ t_2\ ]]$		*!*		*

Less local optimization:

Suppose optimization applied to phrases. An optimal DP will always involve raising of DP_{dat}. But with DP_{dat} raised, DP_{dat} and NP are equally close to the pronoun. Then the inflectional probe can receive value [masc], deriving (22-b) (see O_2 in T_{16}): Thus the approach overgenerates.

T_{16}: Phrasal optimization: wrong result

Input: $D_{[*case:dat*],[*gen_{\surd}:\square*],[*gen_{infl}:\square*],[\bullet EPP\bullet],...}$ + $[_{NP}\ N_{[gen:fem]}\ DP_{2[case:\square],[gen:masc]}]$	MLC	AC	MC	LR
☞ O_1: $[_{DP_1}\ DP_{2[case:dat],[gen:masc]}\ D_{[case:dat],[gen_{\surd}:masc],...}\ [_{NP}\ N_{[gen:fem]}\ t_2\ ]]$				
☛ O_2: $[_{DP_1}\ DP_{2[case:dat],[gen:masc]}\ D_{[case:dat],[gen_{infl}:masc],...}\ [_{NP}\ N_{[gen:fem]}\ t_2\ ]]$				
O_3: $[_{DP_1}\ D_{[case:dat],[gen_{\surd}:masc][gen_{infl}:fem],[\bullet EPP\bullet]}\ [_{NP}\ N_{[gen:fem]}\ DP_{2[case:dat],[gen:masc]}\ ]]$			*!	
O_4: $[_{DP_1}\ DP_{2[case:dat],[gen:masc]}\ D_{[case:dat],[*gen_{\surd}:\square*],...}\ [_{NP}\ N_{[gen:fem]}\ t_2\ ]]$		*!		

Conclusion

(24) Results:

a. Extremely local optimization in syntax seems viable.

b. Extremely local optimization in syntax is supported empirically:

(i) There are indeterminacies in rule application (Agree vs. Merge) in the minimalist program that need to be resolved.

(ii) They can be resolved in a principled way by assuming constraint violability and constraint ranking, i.e., standard optimality theory (the harmonic serialism version of Prince & Smolensky (2004)).

(iii) The evidence from argument encoding patterns and prenominal dative possessors suggests that optimization is extremely local, affecting the single operation: Less local optimization produces wrong results because differences that can be detected in the derivation may be lost at the phrase (hence: phase, clause, sentence) level.

Ackema, Peter & Ad Neeleman (1998): Optimal Questions, *Natural Language and Linguistic Theory* 16, 443–490.
Adger, David (2003): *Core Syntax*. Oxford University Press, Oxford, New York.
Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller & Philipp Weisser (2015): Ergatives Move Too Early. On an Instance of Opacity in Syntax, *Syntax* 18, 343–387.
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.
Caballero, Gabriela & Sharon Inkelas (2013): Word Construction: Tracing an Optimal Path Through the Lexicon, *Morphology* 23, 103–143.
Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.
Chomsky, Noam (1993): A Minimalist Program for Syntactic Theory. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 1–52.
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.

de Vries, Mark (2005): Possessive Relatives and (Heavy) Pied-Piping, *Journal of Comparative Germanic Linguistics* 9, 1–52.

Epstein, Sam, Hisatsugu Kitahara & Daniel Seely (2009): The Application of Transfer is Case-Sensitive and Deducible. Ms., University of Michigan, Keio University and Eastern Michigan University).

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 & 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 (2009): Local Modelling of Global Case Splits. Master's thesis, Universität Leipzig.

Haider, Hubert (1988): Zur Struktur der deutschen Nominalphrase, *Zeitschrift für Sprachwissenschaft* 7, 32–59.

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 (2003a): Derivational Optimization of Wh-Movement, *Linguistic Analysis* 33, 97–148. (Volume appeared 2007).

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

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 (2013): Extremely Local Optimization. In: H. Broekhuis & R. Vogel, eds., *Linguistic Derivations and Filtering*. Equinox, Sheffield, pp. 135–166.

Heck, Fabian & Gereon Müller (2016): On Accelerating and Decelarating Movement: From Minimalist Preference Principles to Harmonic Serialism. In: G. Legendre, M. Putnam, H. de Swart & E. Zaroukian, eds., *Optimality-Theoretic Syntax, Semantics, and Pragmatics*. Oxford University Press, Universität Leipzig, pp. 78–110.

Kimper, Wendell (2016): Positive Constraints and Finite Goodness in Harmonic Serialism. In: J. McCarthy & J. Pater, eds., *Harmonic Grammar and Harmonic Serialism*. Equinox, London, pp. 221–235.

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.

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 (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 (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.

McCarthy, John, Wendell Kimper & Kevin Mullin (2012): Reduplication in Harmonic Serialism, *Morphology* 22, 173–232.

Müller, Gereon (1998): *Incomplete Category Fronting*. Kluwer, Dordrecht.

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 (2018): Inflectional Morphology in Harmonic Serialism. Book ms., Universität Leipzig. To appear in the series *Advances in Optimality Theory*, London: Equinox.

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

Murphy, Andrew (2017): Cumulativity in Syntactic Derivations. PhD thesis, Universität Leipzig.

Pater, Joe (2016): Universal Grammar with Weighted Constraints. In: J. McCarthy & J. Pater, eds., *Harmonic Grammar and Harmonic Serialism*. Equinox, Sheffield, pp. 1–46.

Pesetsky, David (1982): Paths and Categories. PhD thesis, MIT, Cambridge, Mass.

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.

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.

Rezáč, Milan (2004): Elements of Cyclic Syntax: Agree and Merge. PhD thesis, University of Toronto.

Sigurðsson, Halldór (2002): To be an Oblique Subject: Russian vs. Icelandic, *Natural Language and Linguistic Theory* pp. 691–724.

Sternefeld, Wolfgang (2003): Syntax. Eine merkmalsbasierte Analyse. Book ms., Universität

Tübingen.
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.
Zifonun, Gisela (2004): Dem Vater sein Hut – Der Charme des Substandards und wie wir ihm gerecht werden, *Deutsche Sprache* 03, 97–126.