Optimality-Theoretic Syntax

Gereon Müller (Universität Leipzig), WiSe 2012/2013

Contents

1. Straßenverkehrsordnung

(1) *Vorfahrtsregeln*

a. V(ERKEHRS)-POL(IZIST) (StVO § 36):
Die Zeichen und Weisungen der Polizeibeamten auf der Kreuzung sind zu befolgen. ≫ ("Sie gehen allen anderen Anordnungen und sonstigen Regeln vor.")

b. BL(AULICHT)-EIN(SATZHORN) (StVO §§ 35, 38):
Fahrzeuge des Rettungsdienstes und der Polizei dürfen blaues Blinklicht zusammen mit einem Einsatzhorn verwenden; dies ordnet an: "Alle übrigen Verkehrsteilnehmer haben sofort freie Bahn zu schaffen." ≫

c. L(ICHT)-ZEI(CHEN) (StVO § 37):
An Kreuzungen bzw. Einmündungen bedeuten Grün: "Der Verkehr ist freigegeben"; Gelb: "Vor der Kreuzung auf das nächste Zeichen waren"; Rot: "Halt vor der Kreuzung". ≫ ("Lichtzeichen gehen Vorrangregeln [und] vorrangregelnden Verkehrsschildern ... vor.")

d. V(ERKEHRS)-ZEI(CHEN)(A) (StVO § 39):
Verkehrszeichen auf einem Fahrzeug ist Folge zu leisten. Sie gelten auch, wenn das Fahrzeug sich bewegt. ≫ ("Sie gehen den Anordnungen der ortsfest angebrachten Verkehrszeichen vor.")

e. V(ERKEHRS)-ZEI(CHEN)(B) (StVO § 39):
Ortsfesten Verkehrszeichen ist Folge zu leisten. ≫ ("Regelungen durch Verkehrszeichen gehen den allgemeinen Verkehrsregeln vor.")

f. S(TRASSE) V(OR) F(ELDWEG) (StVO § 8):
Fahrzeuge, die aus einem Feld- oder Waldweg auf eine andere Straße kommen, haben Vorfahr zu gewähren. ≫

g. R(ECHTS) V(OR) L(INKS) (StVO § 8):
An Kreuzungen und Einmündungen hat die Vorfahrt, wer von rechts kommt.

T_1: V-ZEI ≫ RVL

I: A(L,◇), B(R,∇)	V-POL	BL-EIN	LI-ZEI	V-ZEI	SVF	RVL
☞ O_1: A vor B						*
O_2: B vor A				*!		
O_3: A und B				*!		*

T_2: LI-ZEI ≫ V-ZEI

I: A(L,◇,*rot*), B(R,∇,*grün*)	V-POL	BL-EIN	LI-ZEI	V-ZEI	SVF	RVL
O_1: A vor B			*!			*
☞ O_2: B vor A				*		
O_3: A und B			*!	*		*

T_3: Bl-Ein ≫ Li-Zei

I: A(L,◇,*rot*,⊕), B(R,∇,*grün*)	V-Pol	Bl-Ein	Li-Zei	V-Zei	SvF	RvL
☞ O_1: A vor B			*			*
O_2: B vor A		*!		*		
O_3: A und B		*!	*	*		*

T_4: V-Pol ≫ *andere Regeln*

I: A(L,∇,*rot*,↑), B(R,◇,*grün*,↓)	V-Pol	Bl-Ein	Li-Zei	V-Zei	SvF	RvL
☞ O_1: A vor B			*	*		*
O_2: B vor A	*!					
O_3: A und B	*!		*	*	*	

Bemerkung:

↓ ein durch den Verkehrspolizisten erteiltes Haltegebot

↑ durch den Verkehrspolizisten erteilte Fahrtfreigabe

T_5: V-Pol ≫ V-Real

I: A(L,◇,*rot*,↓), B(R,∇,*grün*,↓)	V-Pol	V-Real	Bl-Ein	Li-Zei	V-Zei	SvF	RvL
O_1: A vor B	*!			*			*
O_2: B vor A	*!				*		
O_3: A und B	*!	*		*	*		*
☞ O_4: Ø		*					

2. Model of Grammar

Optimality Theory (OT) has been developed since the early nineties, by Alan Prince, Paul Smolensky, John McCarthy and others. At first, the focus was mainly on phonology; but the approach has since been extended to morphology, syntax, semantics, and pragmatics. The most comprehensive (and best) exposition of the theory is still Prince & Smolensky (1993; 2004). Early groundbreaking work in syntax includes Grimshaw (1997), Pesetsky (1998), and Legendre, Smolensky & Wilson (1998). Introductions include Kager (1999) (with little material on syntax), Müller (2000b) (in German), Legendre (2001), and McCarthy (2002) (with quite a bit on syntax).

(2) *Basic assumptions of Optimality Theory*:

a. *Universality*:
Constraints are universal.

b. *Violability*:
Constraints are violable.

c. *Ranking*:
Constraints are ranked.

d. *Competition*:
The wellformedness of a linguistic expression (e.g., a sentence) LE cannot solely be determined on the basis of LE's internal properties. Rather, external factors (more precisely, the competition of LE with other linguistic expressions) determine whether LE is grammatical or not. LEs are *candidates*.

Note:
LE stands for a grammatical unit that is subject to an optimization procedure deciding on its wellformedness. LE is the basic unit of a grammatical domain (phonology, morphology, syntax, semantics); e.g.: the sentence in syntax (but see below).

(3) *Basic assumptions of grammatical theories that do without competition (e.g., Chomsky's (1981) theory of Government and Binding)*:

a. Not all constraints are universal (parameters, language-specific filters – but cf. 'third-factor' meta-constraints on constraints in recent work in the minimalist program, as in Chomsky (2007; 2008)).

b. Constraints cannot be violated.

c. Constraints are not ranked (all are equally important – but cf., e.g., Subjacency vs. ECP violations).

d. The wellformedness of a linguistic expression (e.g., a sentence) LE can fully be determined on the basis of LE's internal properties. External factors (i.e., the properties of other LE's) are irrelevant.

(4) *Optimality*:
A candidate C_i is optimal with respect to some constraint ranking $<Con_1 \gg Con_2 \gg ... \gg Con_n>$ iff there is no other candidate C_j in the same *candidate set* that has a better *constraint profile*.

(5) *Constraint profile*:
C_j has a better constraint profile than C_i if there is a constraint Con_k such that (i) and (ii) hold:

a. C_j satisfies C_k better than Con_i.

b. There is no constraint Con_l that is ranked higher than Con_k, and for which C_i and C_j differ.

Note:
C_j satisfies a constraint Con better than C_i if C_j violates Con less often than C_i. (This includes the case where C_j does not violate Con at all, whereas C_i does.)

(6) *Candidate set* (possible definitions in syntax):
Two candidates are in the same candidate set iff

a. they have the same content words

b. they have the same words (Chomsky's (1995) numeration)

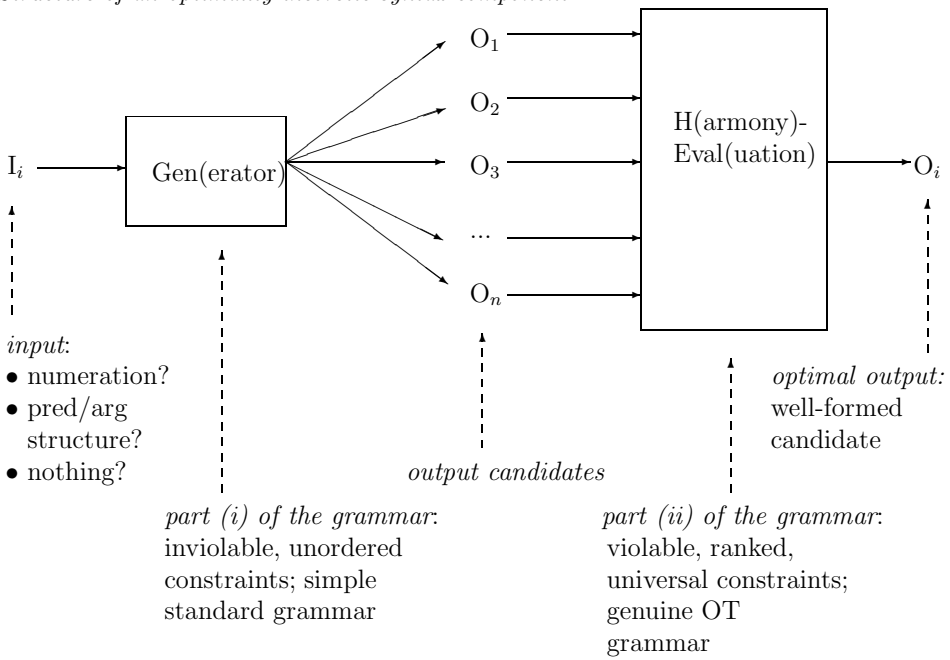
c. they have the same meaning

d. they have the same content words and the same meaning

e. they have the same words and the same meaning

f. they have the same content words and a "sufficiently similar" meaning
g. they have the same f-structure (see work in OT-LFG; Sells (2001a), Bresnan (2001), and papers in Sells (2001b))
h. they have the same D-structure (see work in the GB tradition)
i. they have the same predicte/argument structures and the same logical forms
j. they have an identical *index* (a "target predicate-argument structure, with scopes indicated for variables; operators mark scope"; Legendre, Smolensky & Wilson (1998, 258)).

Structure of an optimality-theoretic syntax component



(7) *The input/output distinction*:
a. Standard OT assumption: Gen creates outputs candidates on the basis of an input; i.e., inputs also define the candidate set.
b. Outputs differ from their underlying input in various ways (giving rise to faithfulness violations; see below), but inputs are standardly assumed to be of roughly the same type as outputs (e.g., URs in phonology), and may even be identical. This seems hardly tenable for syntax. (If outputs are syntactic structures, and structures are generated by Gen, then where does the input structure come from if inputs are also syntactic structures?)
c. Consequently, it is completely unclear what the input in syntax should look like; perhaps there is no input in syntax at all. See Heck et al. (2002), where it is argued that the two basic motivations for inputs – (a) defining candidate sets and (b) providing information for faithfulness constraints – are either unavailable (a) or irrelevant (b) in syntax (because syntax, unlike, phonology, is an information-preserving system, with, e.g., subcategorization information present on a verb throughout the derivation).
d. Anyway, from now on: candidates = outputs.

Types of constraints:
Standardly, two basic types of H-Eval constraints can be distinguished that often give rise to conflicts:

- *Faithfulness constraints* demand that input and output are identical with respect to some property (no addition of items in the output: Dep; no deletion of items in the output: Max; no change of items in the output: Ident).
- *Markedness constraints* impose requirements on outputs that may necessitate a deviation from the input.

Note:
Optimality-theoretic competitions are often illustrated by tables, so-called *tableaux*.

T_6: *The basic principle*

	A	B	C
☞ O_1			*
O_2			**!
O_3		*!	
O_4	*!		
O_5		*!	*

Generalization:
The optimal output is the candidate that has its first star furthest to the right in a tableau.

(8) *Cross-linguistic variation*:
Assumption (not made in minimalist approaches, but virtually everywhere else): Languages differ with respect to their grammars. Grammatical differences between languages = parametrization. Parametrization in optimality theory: *constraint reranking*.

Note:
Optimality theory was developed out of so-called "harmonic grammar" approaches → theory of neural networks. (Further reading: Prince & Smolensky (2004, ch. 10), Smolensky & Legendre (2006, part I). Main innovation: Quality before quantity; no number of violations of a lower-ranked constraint can outweigh a single violation of a higher-ranked constraint.
Caveat:
OT has introduced a means to undermine the irrelevance of constraint violation quantity as

T_7: Parametrization

	A	C	B
O_1		*!	
O_2		**!	
☞ O_3			*
O_4	*!		
O_5		*!	*

T_8: Irrelevance of constraint violation numbers as such

	A	B	C
☞ O_1			****
O_2			*****!**
O_3		*!	
O_4	*!		
O_5		*!	*

such: *(reflexive) local conjunction* (see Smolensky (1996; 2006); and Legendre, Smolensky & Wilson (1998), Fischer (2001), Aissen (1999; 2002), Keine & Müller (2008; 2009), and Keine (2009) for some syntactic applications of local conjunction).

(9) *Local Conjunction*:

a. Local conjunction of two constraints Con_1, Con_2 with respect to a local domain D yields a new constraint $\text{Con}_1\&_{\text{D}}\text{Con}_2$ that is violated iff there are two separate violations of Con_1 and Con_2 in a single domain D.

b. Universal ranking: $\text{Con}_1\&_D\text{Con}_2 \gg \{\text{Con}_1, \text{Con}_2\}$

c. If $\text{Con}_1 = \text{Con}_2$, local conjunction is reflexive.

d. Notation: $\text{B}^2 = \text{B}\&\text{B}$, $\text{B}^3 = \text{B}^2\&\text{B}$, etc.

T_9: A consequence of reflexive local conjunction

	C^4	A	B	C
O_1	*!			****
O_2	*!			*******
☞ O_3			*	
O_4		*!		
O_5			*!	*

3. Evidence for OT Analyses in Syntax

Evidence for OT analyses:

- constraint conflict
- repair phenomena

- default contexts ('emergence of the unmarked')
- cross-linguistic variation by constraint re-ranking

3.1 Constraint Conflict

Profile of the empirical evidence:
The facts show that two general and far-reaching constraints are well motivated, independently of each other. However, in some contexts the two constraints may end up being in conflict, with the evidence suggesting that one may selectively, and systematically, be violated in favour of the other. In standard approaches to grammar, this state of affairs automatically gives rise to an undesirable consequence: One of the two constraints must be abandoned; or there has to be an explicit exception clause in the definition of one of the constraints; or the application of one of the two constraints has to be relegated to some other (typically more abstract) level of representation; etc. In an OT grammar, the constraint conflict can be systematically resolved by constraint ranking.

(10) *Wh-movement in English*:

a. I don't know [$_{\text{CP}}$ which book John bought]

b. *I don't know [$_{\text{CP}}$ John bought which book]

(11) *Two plausible constraints*:

a. Wh-Criterion (Wh-Crit):
Wh-items are in $\text{SpecC}_{[\text{wh}]}$.

b. θ-Assignment (θ-Assign):
Internal arguments of V are c-commanded by V.

Consequence for standard models of grammar:
Either θ-Assign does not hold; or the constraint is enriched by an exception clause ("does not hold for *wh*-items"); or both constraints hold, but not at the same level of representation (Wh-Crit may hold for surface representations or S-structure, θ-Assign may hold for an abstract level of predicate argument structure or D-structure).

Consequence in OT:
Both constraints hold, but they are ranked.

(12) *Ranking*:
Wh-Crit $\gg$ θ-Assign

T_{10}: Simple wh-question formation in English

Input: John, bought, which, book, v, T, $C_{[+\text{wh}]}$	Wh-Crit	θ-Assign
☞ O_1: ... which book John bought		*
O_2: ... John bought which book	*!	

Note:
The displacement of the *wh*-item can be analyzed in terms of a syntactic movement transformation that moves the *wh*-item from its base position into the target SpecC position.

Movement may be assumed leave a trace (t) or a copy. The role of θ-ASSIGN can then be taken over by the general constraint ECONOMY (Grimshaw (1997), Legendre, Smolensky & Wilson (1998), Ackema & Neeleman (1998), etc.).

(13) ECONOMY:
Traces (copies) are prohibited.

Note:
Arguably, (13) can (and should) be derived from more general constraints and their interaction (see Grimshaw (2001; 2006); Steddy & Samek-Lodovici (2009) for an application of the underlying logic to universal constraints on DP-internal order of D, Num, A and N).

Observation:
(i) All syntactic constituents violate alignment constraints (given dichotomies like HEAD-LEFT/HEAD-RIGHT, COMP-LEFT/COMP-RIGHT). More structure implies more violations of alignment constraints. Movement is structure-building (Grimshaw (2001)).
(ii) All chains are trivial (single-membered) in the input. Movement gives rise to non-trivial (multi-membered) chains. This implies a violation of faithfulness (IDENT/UNIQUENESS; see Grimshaw (2006)).

To sum up:

> In conclusion, whether UG constraints conflict or not is an empirical issue. If they do, and they do appear to do so, a formally precise theory of their interaction becomes necessary for a proper understanding of grammar because simultaneous satisfaction of all constraints ceases to be a viable definition of grammaticality.
>
> Samek-Lodovici (2006, 94)

3.2 Repair Phenomena

Profile of the empirical evidence:
The facts suggest that some grammatical complex LE exhibits properties that are not normally permitted in the grammar. It seems that, in the case at hand, these properties are permitted as a last resort (given that all alternatives are even worse, in a sense to be made precise).

3.2.1 Do-Support

Lit.:
Chomsky (1957; 1991)

(14) *'Do'-Einsetzung im Englischen bei Negation*:
a. *Mary not left
b. Mary did not leave

(15) *Keine 'do'-Einsetzung im Englischen bei fehlender Negation*:
a. Mary left
b. *Mary did leave

(16) *Lösung in Chomsky (1957, 39 & 62)*:
a. *Affix Hopping*:
Let *Af* stand for any of the affixes *past*, *S*, *en*, *ing*. Let *v* stand for any M(odal) or V(erb), or *have* or *be* (i.e., for any non-affix in the phrase *Verb*). Then:
$Af + v \rightarrow v + Af\#$,
where # is interpreted as word boundary.
Replace + by # except in the context $v - Af$. Insert # initially and finally.
b. T_{not}:
T_{not} adds *not* after the second segment of the string.
c. *Do Support*:
$\#Af \rightarrow \#\text{do}+Af$.
d. Assumption: T_{not} applies before *Affix Hopping* (bleeding), *Do Support* applies after *Affix Hopping* (counter-feeding).

Intuition: Affix Hopping wird blockiert bei Nicht-Adjazenz von Af und v (nach Negationseinsetzung). In diesem Fall, und nur in diesem, erfolgt dann *Do* Support (weil Af am linken Rand geblieben ist).

(17) *Keine 'do'-Einsetzung bei Negation und* have:
a. *Basis*:
Mary – S+have – en+come
b. T_{not}:
Mary – S+have+not – en+come
c. *Affix Hopping*:
Mary – have+S+not – en+come
d. *Do Support*:
– (Kann nicht applizieren, weil es kein allein stehendes Affix gibt).
e. *Realisierung*:
Mary has not come.

(18) *'Do'-Einsetzung bei Negation ohne Auxiliar*:
a. *Basis*:
Mary – S – come
b. T_{not}:
Mary – S+not – come
c. *Affix Hopping*:
– (Kann nicht applizieren, weil Neg-Insertion zu Nicht-Adjazenz von *Af* und *v* geführt hat.)
d. *Do-Support*:
Mary – do+S+not come
e. *Realisierung*:
Mary did not come.

Rekonstruktion bei Speas (1995)

(19) *Beschränkungsordnung*:
ECP $\gg$ LETZT-AUS $\gg$ ÖKON

T_{11}: Negation und 'do'-Einsetzung bei Speas

Kandidaten	ECP	LETZT-AUS	"ÖKON
☞ K_1: Mary did_1 not t_1 leave		*	*
K_2: Mary t_1 not $left_1$	*!		*

T_{12}: Verbot der 'do'-Einsetzung ohne Negation bei Speas

Kandidaten	ECP	LETZT-AUS	ÖKON
K_1: Mary did_1 t_1 leave		*!	*
☞ K_2: Mary t_1 $left_1$			*

Rekonstruktion von Chomsky (1991):
"Um die korrekten Ergebnisse zu erzielen, muss das Prinzip des 'letzten Auswegs' so interpretiert werden, dass Prinzipien der UG angewendet werden, wann immer das möglich ist, und dass sprachspezifische Regeln nur benutzt werden, um eine D-Struktur-Repräsentation zu "retten", die ansonsten kein Ergebnis liefert" (Chomsky (1991, 427)). (Vgl. auch die Logik bei Adgers (2003) Erklärung über die "Pronouncing Tense Rule" (PTR).)

(20) *Unabhängig blockiert*:
a. *Mary $left_1$ not t_1
b. *Mary $left_1$ t_1

Bemerkung:
Die Analyse hat Konsequenzen für die Definition der Kandidatenmenge.

Rekonstruktion bei Grimshaw (1997)

(21) a. LEX-ÖKON ("Bewegungsökonomie für lexikalische Köpfe", "No-Lex-Mvt"):
Bewegung von lexikalischen Köpfen ist verboten (X^0-$Spur_{lex}$ ist nicht erlaubt).
b. KASUS ("Case"; Chomsky (1981)):
Der Kopf einer NP-Kette muss in einer Kasusposition sein.
c. OB-KOPF ("Obligatorische Köpfe", "Ob-Hd"):
Eine Projektion hat einen (nicht-leeren) Kopf.
d. SUBJEKT ("Erweitertes Projektionsprinzip", "Subj"; Chomsky (1982; 1995)):
Der höchste A-Spezifikator eines Satzes muss durch ein Argument gefüllt sein.
e. VOLL-INT ("Vollständige Interpretation", "Full Interpretation"; Chomsky (1986b)):
Expletiveinsetzung ist verboten.
f. ÖKON ("Stay", s.o.):
Bewegung ist verboten (Spur ist nicht erlaubt).

(22) *Beschränkungsordnung*:
LEX-ÖKON ≫ KASUS ≫ OB-KOPF ≫ SUBJEKT ≫ VOLL-INT ≫ ÖKON

(23) VOLL-INT:
Die lexikalisch-konzeptuelle Struktur muss respektiert werden.

Grundannahme:
Die Größe der Satzstruktur ist variabel.

(24) a. *[$_{NegP}$ Not [$_{VP}$ Mary left]]
b. *[$_{NegP}$ $Mary_1$ not [$_{VP}$ t_1 left]]
c. *[$_{IP}$ $Mary_1$ [$_I$ –] [$_{NegP}$ not [$_{VP}$ t_1 left]]]
d. *[$_{IP}$ $Mary_1$ [$_I$ $left_2$] [$_{NegP}$ not [$_{VP}$ t_1 t_2]]]
e. [$_{IP}$ $Mary_1$ [$_I$ did] [$_{NegP}$ not [$_{VP}$ t_1 leave]]]

T_{13}: Negation und 'do'-Einsetzung bei Grimshaw

Kandidaten	LEX-ÖKON	KA-SUS	OB-KOPF	SUB-JEKT	VOLL-INT	ÖKON
K_1: [Neg [$_{VP}$ NP_1 V]]				*!		
K_2: [NP_1 Neg [$_{VP}$ t_1 V]]		*!				*
K_3: [NP_1 – [t_1 Neg [$_{VP}$ t_1 V]]]		*!	*			**
K_4: [NP_1 V_2 [t_1 Neg [$_{VP}$ t_1 t_2]]]	*!					***
☞ K_5: [NP_1 did_2 [t_1 Neg [$_{VP}$ t_1 V]]]					*	**

(25) a. [$_{VP}$ Mary left]
b. *[$_{IP}$ $Mary_1$ [$_I$ –] [$_{VP}$ t_1 left]]
c. *[$_{IP}$ $Mary_1$ [$_I$ $left_2$] [$_{VP}$ t_1 t_2]]
d. *[$_{IP}$ $Mary_1$ [$_I$ did] [$_{VP}$ t_1 leave]]

T_{14}: Verbot der 'do'-Einsetzung ohne Negation bei Grimshaw

Kandidaten	LEX-ÖKON	KA-SUS	OB-KOPF	SUB-JEKT	VOLL-INT	ÖKON
☞ K_1: [$_{VP}$ NP_1 V]						
K_2: [NP_1 – [$_{VP}$ t_1 V]]		*!	*			*
K_3: [NP_1 V_2 [$_{VP}$ t_1 t_2]]	*!					**
K_4: [NP_1 did_2 [$_{VP}$ t_1 V]]					*!	*

Gemeinsamkeit:
Die Rolle des ECP in Speas' Analyse und die Rolle von KASUS in Grimshaws Analyse sind äquivalent; es sind die entscheidenden Auslöser von *do*-Einsetzung in Negationskontexten. Ebenso äquivalent sind LETZT-AUS bei Speas und VOLL-INT bei Grimshaw: Beide Beschränkungen bestrafen die Verwendung eines expletiven Verbs *do*; sie sind in (Auxiliar-freien) Negationskontexten von optimalen Kandidaten verletzbar. LEX-ÖKON ist bei Speas implizit, bei Grimshaw explizit angenommen; dasselbe gilt auch für SUBJEKT (denn bei Annahme der Hypothese der prädikats-internen Subjekte muss auch Speas

gewährleisten, dass ein Satz wie **Not Mary left* ausgeschlossen ist). Schließlich hat auch ÖKON in beiden Analysen dieselbe Funktion, nämlich interessanterweise – für die betrachteten Fälle – gar keine.

3.2.2 Resumptive Pronouns

(26) *Resumptive pronouns in English*

a. (the man) who(m) I saw t
b. *(the man) who(m) I don't believe the claim that anyone saw t
c. *(the man) who(m) I saw him
d. ?(the man) who(m) I don't believe the claim that anyone saw him

Observation:
The insertion of resumptive pronouns may (often) be viewed as a repair phenomenon, i.e., a last resort, if a well-formed sentence cannot otherwise be generated (Shlonsky (1992), Hornstein (2001)). The insertion of a resumptive pronoun (which, by assumption, is not part of the input) violates a faithfulness constraint, but is required by a higher-ranked markedness constraint. (See Pesetsky (1998), Legendre, Smolensky & Wilson (1998), and Salzmann (2006) for OT analyses of resumptive pronouns.)

(27) a. REL-CRITERION (REL-CRIT):
Relative pronouns are in SpecC of a relative clause.
b. COMPLEX NP CONDITION, CNPC):
A moved item must not be separated from its trace by an intervening DP.
c. INCLUSIVENESS (INCL, a DEP constraint):
Every element of the output must be present in the input.

(28) *Ranking*:
REL-CRIT ≫ CNPC ≫ INCL

T_{15}: Trace vs. resumptive pronouns; transparent context

Input: I, who(m), saw, $C_{[rel]}$, the, man	REL-CRIT	CNPC	INCL
☞ O_1: the man who(m) I saw t			
O_2: the man who(m) I saw him			*!
O_3: the man I saw who(m)	*!		

T_{16}: Trace vs. resumptive pronoun, opaque CNPC context

Input: anyone, who(m), saw, I, do, not, believe, the, claim, that, $C_{[rel]}$, the man	REL-CRIT	CNPC	INCL
O_1: the man who(m) I don't believe the claim that anyone saw t		*!	
☞ O_2: the man who(m) I don't believe the claim that anyone saw him			*
O_3: the man I don't believe the claim that anyone saw who(m)	*!		

3.2.3 Other Repair Phenomena

Further cases of repair phenomena:

- Ersatz-infinitive in German, Participium pro Infinitivo in Swedish (Wiklund (2001)) → Schmid (2005)
- R-pronouns in German, Dutch and (Middle) English (Riemsdijk (1978)) → Müller (2000a)
- *wh*-scope marking in German and Hungarian → Müller (1997)
- expletives in SpecC and SpecT → Müller (2000b), Grimshaw (2006)
- repair-driven quantifier raising in VP ellipsis contexts in English (as in Fox (2000)) → Heck & Müller (2000; 2003)
- repair-driven intermediate movement steps as required by the PIC (Chomsky (2001; 2008)) → Heck & Müller (2000; 2003)
- repair-driven multiple *wh*-movement in German sluicing constructions (Merchant (2001)) → Heck & Müller (2000; 2003)

(29) *Ersatz-infinitive*:

a. dass sie das *gewollt* hat
that she that wanted.PART has
b. *dass sie das hat *wollen*
that she that want.INF has
c. *dass sie das Lied singen *gewollt* hat
that she the song sing wanted.PART has
d. dass sie das Lied hat singen *wollen*
that she the song has sing want.INF

(30) *Repair-driven quantifier raising*:

a. [$_{CP_1}$ Some boy admires every teacher], [and [$_{CP_2}$ some girl does [$_{VP}$ ~~admire every teacher~~] too]] (∃∀, ∀∃)
b. [$_{CP_1}$ Some boy admires every teacher], [and [$_{CP_2}$ Mary does [$_{VP}$ ~~admire every teacher~~] too]] (∃∀, *∀∃)
c. [$_{CP_1}$ Mary admires every teacher], [and [$_{CP_2}$ some boy does [$_{VP}$ ~~admire every teacher~~] too]] (∃∀, ∀∃)

(31) *Repair-driven multiple wh-movement*:

a. Irgendjemand hat irgendetwas geerbt, aber der Fritz weiß nicht mehr
someone has something inherited but the Fritz knows not more
[$_{CP}$ wer$_1$ was$_2$ ~~C$_{[\Delta]}$ t$_1$ t$_2$~~ ~~geerbt~~ ~~hat~~]
who what inherited has

b. *Irgendjemand hat irgendetwas geerbt, aber der Fritz weiß nicht mehr
someone has something inherited but the Fritz knows not more
[$_{CP}$ wer$_1$ ~~C$_{[\Delta]}$ t$_1$~~ ~~was$_2$~~ ~~geerbt~~ ~~hat~~]
who what inherited has

3.3 Default Contexts

Profile of the empirical evidence:
The empirical evidence suggests that there is a concept like "unmarked case" ("default case", "elsewhere case"): Some linguistic property P of LEs counts as the unmarked case if it shows up whenever something else (that is incompatible with P) is not explicitly required. In standard conceptions of grammar, the theoretical implementation of this concept is far from unproblematic. (Whenever it seems to be unproblematic, as in approaches to syntax that envisage blocking (Williams (1997), Fanselow (1991)), or in Distributed Morphology (Halle & Marantz (1993)), this is due to the fact that the approach is in fact based on competition and candidate sets, too.) In OT, an unmarked case signals the activity of a constraint that is ranked very low, and that is typically rendered inactive by higher-ranked, conflicting constraints → *emergence of the unmarked*.

Empirical generalization:
In the unmarked case, a DP bears nominative case in German; nominative is the default case.

(32) *Nominative as an unmarked case*:

a. *Constructions with 'als' that do not have case agreement*:
die Ehrung des Kanzlers als großer Politiker/ *großen Politiker
the homage to the chancellor as great politician$_{nom}$ great politican$_{acc}$

b. *Infinitival constructions with (case-less) PRO*:
Wir baten die Männer [$_{CP}$ PRO einer nach dem anderen/ *einen nach
we asked the men one$_{nom}$ after the other one$_{acc}$ after
dem anderen durch die Sperre zu gehen]
the other through the barricade to go

c. *Left dislocation without case agreement*:
Der Kaiser/ *Den Kaiser, dem verdanken wir nichts
the emperor$_{nom}$ the emperor$_{acc}$ him owe we nothing

d. *Predicate 'und'-constructions*:
Der/ *Den und ein Buch lesen? (Dass ich nicht lache!)
he$_{nom}$ him$_{acc}$ and a book read that I not laugh

(33) *Blocked nominative in contexts with case government*:

a. dass ich *er/ ihn getroffen habe
that I he$_{nom}$ him$_{acc}$ met have

b. dass man *der Mann/ des Mannes gedachte
that one the man$_{nom}$ the man$_{gen}$ remembered

c. dass wir *der Mann/ den Mann das Buch lesen sehen
that we the man$_{nom}$ the man$_{acc}$ the book$_{acc}$ read see

(34) *Case-related constraints*:

a. Gen(itive) Constraint (Gen):
The object of a verb that is lexically marked as governing genitive case bears genitive.
(A subcase of a more general constraint demanding faithfulness to lexical case specifications.)

b. Acc(usative) Constraint (Acc):
The object of a transitive verb bears accusative case.

c. Nominative Constraint (Nom):
A DP bears nominative case.

(35) *Ranking*:
Gen ≫ Acc ≫ Nom

T_{17}: Accusative government

Input: dass, getroffen, habe, 1.Sg./Agent, 3.Sg./Patient	Gen	Acc	Nom
☞ O_1: dass ich ihn getroffen habe			*
O_2: dass ich er getroffen habe		*!	
O_3: dass mich ihn getroffen habe			**

T_{18}: Nominative as the unmarked case

Input: und, ein, Buch, lesen, 3.Sg./Agent/Dem	Gen	Acc	Nom
O_1: Den und ein Buch lesen ?			*!
☞ O_2: Der und ein Buch lesen ?			
O_3: Dem und ein Buch lesen ?			*!

Note:
If nominative (absolutive in ergative alignment patterns) is inherently a default case across languages, free reranking in (35) must be blocked in some way (that may be related to the more primitive feature structures of the cases; see, e.g., Wunderlich (1997), Kiparsky (1999; 2001)).

Observation:
Instead of a system of constraints demanding case realization, one can just as well

have a system of constraints prohibiting case realization, accompanied by an inherently highest-ranked (or inviolable) constraint that states that all DPs have case. See Woolford (2001).

Background assumptions (Woolford (2001)):

1. There are (ordered) markedness constraints that block the realization of cases.
2. There are faithfulness constraints that demand the realization of case specifications in the input (lexical, inherent case).
3. Nominative/absolutive and accusative are structural cases; dative and ergative (and genitive) are inherent cases (that must be specified on a verb).
4. Every DP must be case-marked (perhaps part of Gen).

(36) a. *Dat ("*Dative"):
Avoid dative case.
b. *Acc ("*Accusative"):
Avoid accusative case.
c. *Nom ("*Nominative"):
Avoid nominative case.
d. Faith-Lex:
Realize a case feature specified on V in the input.
e. Faith-Lex$_{\text{trans}}$:
Realize a case feature specified on transitive V in the input.

(37) a. *Ranking in Icelandic*:
Faith-Lex$_{tr}$ ≫ Faith-Lex ≫ *Dat ≫ *Acc ≫ *Nom
b. *Ranking in Japanese*:
Faith-Lex$_{tr}$ ≫ *Dat ≫ Faith-Lex ≫ *Acc ≫ *Nom
c. *Ranking in English*:
*Dat ≫ Faith-Lex$_{tr}$ ≫ Faith-Lex ≫ *Acc ≫ *Nom

(38) *Quirky case in Icelandic*:
a. Bátnum hvolfdi
boat$_{dat}$ capsized
b. Barninu batnadhi veikin
child$_{dat}$ recovered from disease$_{nom}$

T_{19}: Intransitive V in Icelandic; inherent dative

Candidates	Faith-Lex$_{tr}$	Faith-Lex	*Dat	*Acc	*Nom
☞ O$_1$: DP$_{dat}$ V$_{[+dat]}$			*		
O$_2$: DP$_{nom}$ V$_{[+dat]}$		*!			*
O$_3$: DP$_{acc}$ V$_{[+dat]}$		*!		*	

T_{20}: Transitive V in Icelandic; inherent dative on DP$_{ext}$

Candidates	Faith-Lex$_{tr}$	Faith-Lex	*Dat	*Acc	*Nom
☞ O$_1$: DP$_{dat}$ V$_{[+dat]}$ DP$_{nom}$			*		*
O$_2$: DP$_{dat}$ V$_{[+dat]}$ DP$_{acc}$			*	*!	
O$_3$: DP$_{nom}$ V$_{[+dat]}$ DP$_{acc}$	*!	*		*	*

(39) *Quirky case in Japanese*:
a. Akatyan-ga/*-ni moo arukeru
baby$_{nom/dat}$ already walk can
b. Taroo-ni eigo-ga hanaseru
Taro$_{dat}$ English$_{nom}$ speak can

T_{21}: Intransitive V in Japanese; no inherent dative

Candidates	Faith-Lex$_{tr}$	*Dat	Faith-Lex	*Acc	*Nom
O$_1$: DP$_{dat}$ V$_{[+dat]}$		*!			
☞ O$_2$: DP$_{nom}$ V$_{[+dat]}$			*		*
O$_3$: DP$_{acc}$ V$_{[+dat]}$			*	*!	

T_{22}: Transitive V in Japanese; inherent dative on DP$_{ext}$

Candidates	Faith-Lex$_{tr}$	*Dat	Faith-Lex	*Acc	*Nom
☞ O$_1$: DP$_{dat}$ V$_{[+dat]}$ DP$_{nom}$		*			*
O$_2$: DP$_{dat}$ V$_{[+dat]}$ DP$_{acc}$		*		*!	
O$_3$: DP$_{nom}$ V$_{[+dat]}$ DP$_{acc}$	*!		*	*	*

Note:
For more on case in OT, see Kiparsky (1999), Wunderlich (2000; 2003), Stiebels (2000; 2002), Woolford (2001), Lee (2003), de Hoop & Malchukov (2008), Swart (2007), Keine & Müller (2008; 2009), and references cited in these works.

4. Problems for OT Analyses in Syntax

Problems for OT Analyses:

- complexity of competition-based grammars (potentially infinite candidate sets)
- ineffability (absolute ungrammaticality)
- optionality

4.1 Complexity

Observation:
Because of the general option of recursion in syntax, candidate sets are not finite in most analyses.

Reaction:

> This qualm arises from a misapprehension about the kind of thing that grammars are. It is not incumbent upon a grammar to compute, as Chomsky has emphasized repeatedly over the years. A grammar is a function that assigns structural descriptions to sentences; what matters formally is that the function is well-defined. The requirements of explanatory adequacy (on theories of grammar) and descriptive adequacy (on grammars) constrain and evaluate the space of the hypotheses. Grammatical theorists are free to contemplate any kind of formal device in pursuit of these goals; indeed, they *must* allow themselves to range freely if there is to be any hope of discovering decent theories. Concomitantly, one is not free to impose arbitrary additional meta-constraints (e.g. 'computational plausibility') which could conflict with the well-defined basic goals of the enterprise. In practice, computationalists have always proved resourceful. All available complexity results for known theories are stunningly distant from human processing capacities ... yet all manner of grammatical theories have nonetheless been successfully implemented in parsers, to some degree or another, with comparable efficiency. ... There are neither grounds of principle nor grounds of practicality for assuming that computational complexity considerations, applied directly to grammatical formalisms, will be informative.
>
> Prince & Smolensky (1993, 197; 2004, 233)

Also cf.:

> To avoid what has been a continuing misunderstanding, it is perhaps worth while to reiterate that a generative grammar is not a model for a speaker or hearer. It attempts to characterize in the most neutral possible terms the knowledge of the language that provides the basis for actual use of language by a speaker-hearer. When we speak of a grammar as generating a sentences with a certain structural descritpion, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentences has a certain derivation with respect to a a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These question belong to the theory of language use – the theory of performance.
>
> Chomsky (1965, 9)

Note:
If there is a problem here, OT shares the problem with other competition-based theories of syntax (e.g., early minimalist approaches like that of Chomsky (1993), which rely on transderivational constraints applying to candidate derivations in large (typically infinite) reference sets).

4.2 Ineffability (Absolute Ungrammaticality)

Observation:
Basically, a sentence can only qualify as ungrammatical if there is some other sentence that blocks it by being the optimal candidate. However, sometimes it is far from obvious what this other sentence should look like.

(40) *Adjunct islands in German*:
*Was ist Fritz eingeschlafen [$_{CP}$ nachdem er t gelesen hat] ?
what is Fritz fallen asleep after he read has

4.2.1 The Generator

Assumption:
Gen contains constraints like (41) that preclude a generation of outputs in the first place.

(41) Adjunct Condition:
Movement must not cross an adjunct clause.

4.2.2 Empty Outputs

Assumption:
Each candidate set contains a candidate that leaves the input completely unrealized. This candidate is the "empty output" or "null parse": Ø. By definition, the empty output does not violate any faithfulness constraints; the only constraint that it violates is (42).

(42) *Ø ("Avoid Null Parse"):
The input must not be completely unrealized.

T_{23}: Ineffability and empty outputs

	Adjunct Condition	Wh-Crit	*Ø
O_1: was ... [nachdem er t V]	*!		
O_2: – ... [nachdem er was V]		*!	
☞ O_3: Ø			*

Note:
The constraint *Ø defines a strict upper bound in constraint rankings: Constraints that outrank *Ø are not violable by optimal outputs.

4.2.3 Bad Winners

Assumption:
The optimal candidate cannot be interpreted by other components of grammar (phonology, semantics), or by the interfaces with these components.

T_{24}: Ineffability and bad winners

	Adjunct Condition	Wh-Crit
O_1: was ... [nachdem er t V]	*!	
☞ O_2: – ... [nachdem er was V]		*

(43) *Semantic uninterpretability of the optimal wh-in-situ candidate*:
#Fritz ist eingeschlafen [$_{CP}$ nachdem er was gelesen hat] ?
Fritz is fallen asleep after he what read has

Note:
This approach arguably chimes in well with recent trends in minimalist syntax to attribute much of the work standardly done by syntactic constraints to interface requirements; see Chomsky (2007; 2008), and particularly Boeckx (2009).

4.2.4 Repair

Hypothesis:
There is in fact an optimal repair candidate for (40).

(44) *Two potential repair candidates*:

a. Fritz ist eingeschlafen [$_{CP}$ nachdem er was gelesen hat] (= etwas)
Fritz is fallen asleep after he something read has

b. Bei was ist Fritz eingeschlafen [$_{CP}$ nachdem er es gelesen hat] ?
with respect to what is Fritz fallen asleep after he it read has

(45) MAX([wh]):
A feature [+wh] in the input must be realized in the output.

T_{25}: Ineffability and repair

	ADJUNCT CONDITION	WH-CRIT	MAX([wh])
O_1: was$_{[+wh]}$ ... [nachdem er t V]	*!		
O_2: – ... [nachdem er was$_{[+wh]}$ V]		*!	
☞ O_3: – ... [nachdem er was$_{[-wh]}$ V]			*

(46) *Problem*:

a. *Long wh-movement*:
Was glaubt Fritz [$_{CP}$ dass er t lesen sollte] ?
what thinks Fritz that he read should

b. *Wh-indefinite*:
Fritz glaubt [$_{CP}$ dass er was lesen sollte]
Fritz thinks that he what (= something) read should

c. *Optional wh-argument plus resumptive pronoun*:
Von was glaubt Fritz [$_{CP}$ dass er es lesen sollte] ?
of what thinks Fritz that he it read should

T_{26}: A wrong prediction

	ADJUNCT CONDITION	WH-CRIT	MAX([wh])
☞ O_1: was$_{[+wh]}$ ... [dass er t V]			
O_2: – ... [– dass er was$_{[+wh]}$ V]		*!	
O_3: – ... [dass er was$_{[-wh]}$ V]			*!

Conclusion:
Wh-indefinite clauses are not repair forms; they are available even if long *wh*-movement is permitted. (Similar conclusions hold in the case of sentences with optional *wh*-argument generation in the matrix clauses; see Koster (1986), Cinque (1990), Barbiers (2002), Gallego (2007).)

4.2.5 Neutralization

Assumption:
There are two competitions based on minimally differing inputs (e.g., inputs that differ only with respect to some feature value). These input differences can then be neutralized by some high-ranked markedness constraint in the output; i.e., two different competitions (based on two candidate sets) converge on a single optimal candidate.

T_{27}: Transparent contexts without neutralization: 'was$_{[+wh]}$' in the input

Input: was$_{[+wh]}$, ...	ADJUNCT CONDITION	WH-CRIT	MAX([wh])
☞ O_1: was$_{[+wh]}$ ... [dass er t V]			
O_2: – ... [– dass er was$_{[+wh]}$ V]		*!	
O_3: – ... [dass er was$_{[-wh]}$ V]			*!

T_{28}: Transparent contexts without neutralization: 'was$_{[-wh]}$' in the input

Input: was$_{[-wh]}$, ...	ADJUNCT CONDITION	WH-CRIT	MAX([wh])
O_1: was$_{[+wh]}$ ... [dass er t V]			*!
O_2: – ... [– dass er was$_{[+wh]}$ V]		*!	*
☞ O_3: – ... [dass er was$_{[-wh]}$ V]			

Note:
A further output O_4 that applies movement of a [–wh] phrase to SpecC in T_{28} has the same constraint profile as O_3 with respect to the three constraints given here. However, it is suboptimal (in fact, *harmonically bounded*, i.e., not expected as grammatical under any reranking) because it violates ECONOMY in addition without contributing to a better behaviour with respect to any other constraint.

(47) *Competing outputs*:

a. *Was ist Fritz eingeschlafen [$_{CP}$ nachdem er t gelesen hat] ?
what is Fritz fallen asleep after he read has

b. Fritz ist eingeschlafen [$_{CP}$ nachdem er was gelesen hat]
Fritz is fallen asleep after he what (= something) read has

T_{29}: Island contexts with neutralization, unfaithful: 'was$_{[+wh]}$' in the input

Input: 'was$_{[+wh]}$, ...	ADJUNCT CONDITION	WH-CRIT	MAX([wh])
O_1: was$_{[+wh]}$ ... [nachdem er t V]	*!		
O_2: – ... [nachdem er was$_{[+wh]}$ V]		*!	
☞ O_3: – ... [nachdem er was$_{[-wh]}$ V]			*

T_{30}: Island contexts with neutralization, faithful: 'was$_{[-wh]}$' in the input

Input: 'was$_{[-wh]}$, ...	Adjunct Condition	Wh-Crit	Max([wh])
O_1: was$_{[+wh]}$ ... [nachdem er t V]	*!		*
O_2: – ... [nachdem er was$_{[+wh]}$ V]		*!	*
☞ O_3: – ... [nachdem er was$_{[-wh]}$ V]			

Note:
For more on ineffability, see Fanselow & Féry (2002a), Legendre (2009), Vogel (2009).

4.3 Optionality

Observation:
In general, only one candidate should be optimal. (However, note that the definition of optimality in (4) is in principle compatible with there being more than one winner, in contrast to, e.g., the one in Grimshaw (1997).) What about situations where it looks as though several outputs can co-exist as optimal?

(48) *Complementizer deletion in English*:
a. I think – John will leave
b. I think that John will leave

(49) *Partial wh-movement in German*:
a. Wen glaubst du [$_{CP}$ dass man t einladen sollte] ?
whom think you that one invite should
b. Was glaubst du [$_{CP}$ wen man t einladen sollte] ?
what think you whom one invite should

(50) *Wh-movement in French*:
a. Qui as-tu vu t ?
b. – Tu as vu qui ?

(51) *Extraposition in German*:
a. dass eine Frau [die ich mag] zur Tür reingekommen ist
that a woman whom I like to the door in come is
b. dass eine Frau t zur Tür reingekommen ist [die ich mag]
that a woman to the door in come is whom I like

(52) *Free word order in German*:
a. dass keiner den Fritz gesehen hat
that no-one$_{nom}$ the Fritz$_{acc}$ seen has
b. dass den Fritz keiner gesehen hat
that the Fritz$_{acc}$ no-one$_{nom}$ seen has

(53) *Optionality of two candidates C_i, C_j* (cf. Müller (2003b)):
a. *Pseudo-optionality*:
C_i, C_j belong to different candidate sets and do not interact.
b. *True optionality*:
C_i, C_j have an identical constraint profile.
c. *Ties*:
C_i, C_j differ only on two (or more) constraints that are tied. Ties can be interpreted in various ways:
(i) ordered global tie
(ii) ordered local tie
(iii) conjunctive local tie
(iv) disjunctive local tie
(v) disjunctive global tie
d. *Neutralization*:
C_i, C_j belong to different candidate sets, but interact nonetheless
e. *Stochastic optimality theory*

4.3.1 Pseudo-Optionality

Assumption:
Candidate sets are defined in such a way that there is little competition.

Problem:
If there is not much competition, this weakens the overall theory and increases the problem of accounting for ineffability.

(54) *Competition of partial and long-distance wh-movement in German*:
a. ?Wen glaubst du nicht [$_{CP}$ dass man t einladen sollte] ?
whom think you not that one invite should
b. *Was glaubst du nicht [$_{CP}$ wen man t einladen sollte] ?
what think you not whom one invite should

(55) *Competition of wh-movement and wh-in-situ in French*:
a. Je me demande [qui C tu as vu t]
I ask myself whom dyou have seen
b. *Je me demande [– (que) tu as vu qui]
I ask myself that you have seen whom

4.3.2 True Optionality

Assumption:
Two (or more) candidates can in fact have the same (optimal) constraint profile.

Problem:
This proves very hard (or impossible) in practice, e.g., because of the existence of faithfulness constraints.

4.3.3 Ties

Assumption:
Two (or more) constraints are equally important, i.e., tied. Candidates that differ only with respect to these ties can all be optimal.

T_{31}: Constraint tie: B∘C

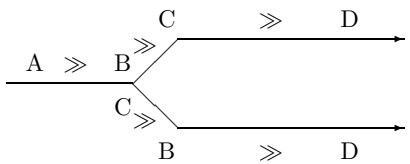
	A	B ¦ C	D
☞ O_1		¦ *	
☞ O_2		* ¦	
O_3		*(!) ¦ *(!)	
O_4	*!	¦	

Note:
There are various different (and largely incompatible) concepts of tie in the literature. A basic distinction that can be made is one between *global* ties and *local* ties. Global ties are abbreviations for the simultaneous presence of different constraint rankings in a language. Local ties can be viewed as special constraint types. The most widespread concepts of tie in the literature are arguably *ordered global ties* and *conjunctive local ties*.

Observation:
With global ties, the optimal outputs of one candidate set may have a different constraint profile below the tie. With local ties, such a different behaviour with lower-ranked constraints leads to a breakdown of optionality.

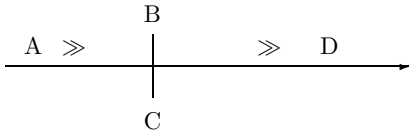
(56) *Ordered global tie*:
Suppose that $\Gamma = <\mathrm{Con}_1 \gg ... \gg \mathrm{Con}_n>$ is a partial constraint order in language L, and Con_i ($1 \leq i \leq n$) $\in \Gamma$ is a tie $\mathrm{Con}_{i_1} \circ ... \circ \mathrm{Con}_{i_n}$. Then, for every suborder O of the constraints in Con_i, Γ_O is a constraint order of language L, where Γ_O differs from Γ only in that Con_i is replaced by O.

(57) *Diagram of an ordered global tie* B∘C



(58) *Conjunctive local tie*:
Suppose that $\Gamma = <\mathrm{Con}_1 \gg ... \gg \mathrm{Con}_n>$ is a total constraint order in language L, and Con_i ($1 \leq i \leq n$) $\in \Gamma$ is a tie $\mathrm{Con}_{i_1} \circ ... \circ \mathrm{Con}_{i_n}$. Then, Con_i is violated by a candidate C iff there is a constraint Con_{i_j} that is violated by C.

(59) *Diagram of a conjunctive local tie* B∘C



4.3.4 Stochastic Optimality Theory

Refs.:
Anttila (1997), Boersma & Hayes (2001), Hayes (2001) (stochastic OT)
Aissen (2002; 2003), Bresnan, Dingare & Manning (2001), Bresnan, Deo & Sharma (2007) (syntactic applications)

Observation:
Often, the constructions that participate in an alternation are not equally frequent or equally unmarked (or, for that matter, equally "well formed" → degrees of acceptability).

(60) *Preferences with optionality in the positioning of English possessives*:
a. the result of the accident > the accident's result
b. Mary's sister > the sister of Mary
c. the boy's uncle > the uncle of the boy
d. the door of the building > the building's door
e. someone's shadow > the shadow of someone
f. the shadow of something > *something's shadow
g. her money > ?*the money of her

Observation:
Animacy and definiteness scales are independently motivated (Hale (1972), Silverstein (1976)). These hierarchies can be used as primitives to generate sequences of constraints (with a fixed internal order: subhierarchies), via *harmonic alignment* of scales (Prince & Smolensky (2004), Aissen (1999)).

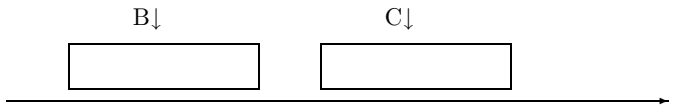
(61) *Harmonic Alignment* (Prince & Smolensky (2004)):
Suppose given a binary dimension D_1 with a scale X > Y on its elements {X,Y}, and another dimension D_2 with a scale a > b > ... > z on its elements {a,b,...,z}. The *harmonic alignment* of D_1 and D_2 is the pair of Harmony scales H_X, H_Y:
a. H_X: X/a ≻ X/b ≻ ... ≻ X/z
b. H_Y: Y/z ≻ ... ≻ Y/b ≻ Y/a
The *constraint alignment* is the pair of constraint hierarchies C_X, C_Y:
a. C_X: *X/z ≫ ... ≫ *X/b ≫ *X/a
b. C_Y: *Y/a ≫ *Y/b ≫ ... ≫ *Y/z

(62) *Constraint subhierarchies via animacy and definiteness scales*:
a. (i) *SpecN/inanimate ≫ *SpecN/animate ≫ *SpecN/human
(ii) *CompN/human ≫ *CompN/animate ≫ CompN/inanimate

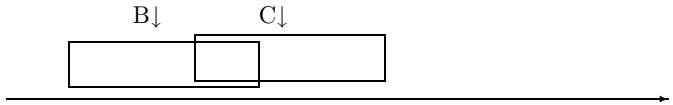
b. (i) *SpecN/indef ≫ *SpecN/def ≫ *SpecN/name ≫ *SpecN/pron
(ii) *CompN/pron ≫ *CompN/name ≫ *CompN/def ≫ *CompN/indef

Proposal:
Constraints are not necessarily categorically ordered with respect to each other. Rather, their application domains may *overlap*. An overlap of application domains gives rise to optionality.

(63) *Categorical order of application domains of constraints*:



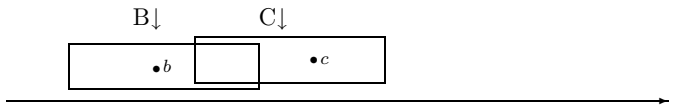
(64) *Overlapping order of application domains of constraints*:



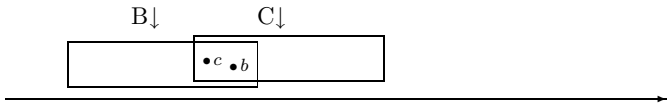
Assumption:
A candidate is evaluated at an *evaluation time*; it is well formed if it is optimal at that point. For an evaluation, an arbitrary point is chosen in the application domain of a constraint. A constraint B is ranked higher than another constraint C at a given evaluation time if the point chosen for B is above the point chosen for C. If the domains of B and C are categorically ordered, then the point for B is always going to be on top of the point for C, and there will be no optionality. However, if the domains of B and C overlap, optionality arises; the winning candidate is determined by whether the point chosen for B is above the point chosen for C or vice versa. (This is basically the concept of ordered hierarchical tie.)

Preferences:
The choice of evaluation point at a given evaluation time is free as such. However, the smaller the common domain of B and C is, the more likely it is that the point chosen for the higher-ranked constraint (say, B) is above the point chosen for the lower-ranked constraint (say, C). Accordingly, the more likely a higher position of B-points vis-a-vis C points at a given evlaution time is, the more the construction favoured by B is going to be preferred over the construction favoured by C; similarly, the more frequent B will be in corpora.

(65) *Typical result*: B ≫ C



(66) *Rare result*: C ≫ B



5. Optimization Domains

5.1 Introduction

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 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)

> 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 serialsim" 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, Smolensky & Wilson (1998, 285-286)

(67) *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)

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.

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.

(68) *Proposals for local optimization*:
a. *Minimal clause*:
Ackema & Neeleman (1998) on *wh*-movement in Czech; Müller (2003a) 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 (2000c) on secondary remnant movement; Heck & Müller (2000; 2003) 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 (2009) on ergative and accusative argument encoding patterns; Lahne (2008; 2009) on excluding SVO in ergative languages; Georgi (2009) on global case splits in Tauya.

5.2 Clauses as optimization domains

Ref.: Ackema & Neeleman (1998)

(69) *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 *wh*s 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 *wh*s), which would go against strict cyclicity.
>
> Ackema & Neeleman (1998, fn. 25)

(70) 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$_{32}$: 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$_2$]]	*!		

Notation:
O$_{11}$-O$_{13}$ in T$_{33}$ are descendants of O$_1$.

T_{33}: 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$_2$]]]]]]		*** *******!**	

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_{34}: 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.

5.3 Derivational Steps as Optimization Domains

Ref.: Heck & Müller (2007), 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.
(Note: In assuming structure-building and optimization to systematically alternate, this approach crucially differs from Broekhuis' *Derivations and Evaluations* model, where all structure-building precedes optimization.)

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.

5.3.1 Constraints, Features, and Operations

(71) *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∗].

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

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

(74) *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.)

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

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

(77) 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).

(78) Last Resort (LR):
Move of α and β follows Agree of α and β.

Comment:
The (perhaps less ordinary) treatment of Move in (78) as a binary operation rests on the assumption that Move is Merge (with β internal to α), which is binary.

5.3.2 Empirical Evidence for Extremely Local Optimization 1: Argument Encoding

(79) *Basic patterns of argument encoding*:

a. Accusative marking		b. Ergative marking	
$\boxed{\text{DP}_{ext}\text{-V}_i\ \text{DP}_{int}\text{-V}_i}$		$\boxed{\text{DP}_{ext}\text{-V}_i\ \text{DP}_{int}\text{-V}_i}$	
$\text{DP}_{ext}\text{-V}_t$	$\boxed{\text{DP}_{int}\text{-V}_t}$	$\boxed{\text{DP}_{ext}\text{-V}_t}$	$\text{DP}_{int}\text{-V}_t$
nom	acc	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 (80).

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

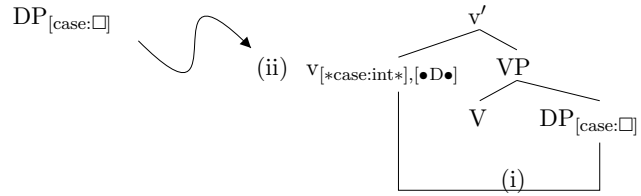
a. T bears a probe [∗case:ext∗] that instantiates a matching [case:ext] goal on DP.
b. v bears a probe [∗case:int∗] that instantiates a matching [case:int] goal on DP.

A 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:□]."

(81) *Stage Σ:*



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.

(82) 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.

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

(84) *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}.

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.

(85) *Rankings*:

a. Accusative patterns:
(MLC $\gg$) AC $\gg$ MC
b. Ergative patterns:
(MLC $\gg$) MC $\gg$ AC

T_{35}: Accusative pattern, step 1 (Σ as input): Agree

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

T_{36}: Accusative pattern, step 2: Merge

Input: $[_{v'}\ v_{[\bullet D\bullet]} \ldots DP_{[case:int]} \ldots]$ Workspace = $\{DP_{[case:\square]}, T_{[*case:ext*]}, \ldots\}$	MLC	AC	MC
☞ O_{21}: $[_{vP}\ DP_{[case:\square]}\ [_{v'}\ v \ldots DP_{[case:int]} \ldots]]$			

T_{37}: Accusative pattern, step 3: Merge

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

Note:
Agree in T_{42} is just local enough to be in accordance with the PIC in Chomsky (2001,

T_{38}: Accusative pattern, step 4: Agree

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

T_{39}: Ergative pattern, step 1 (Σ as input): Merge

Input: $[_{v'}\ v_{[*case:int*],[\bullet D\bullet]} \ldots DP_{[case:\square]} \ldots]$ Workspace = $\{DP_{[case:\square]}, T_{[*case:ext*]}, \ldots\}$	MLC	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_{40}: Ergative pattern, step 2: Agree (with DP_{ext})

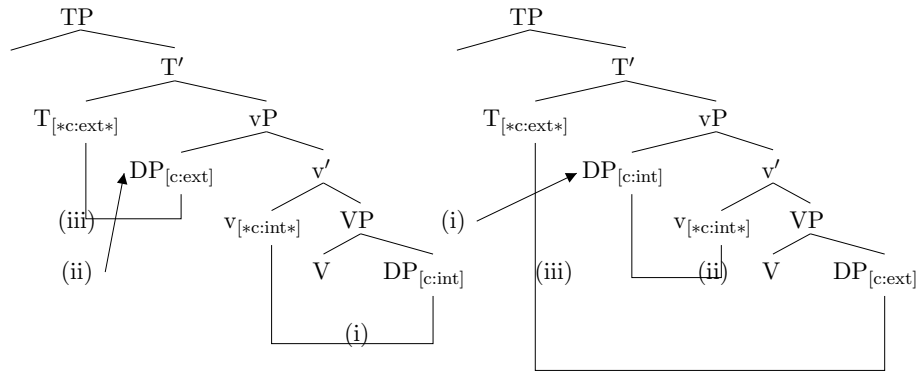
Input: $[_{v'}\ DP_{[case:\square]}\ [_{v'}\ v_{[*case:int*]} \ldots DP_{[case:\square]} \ldots]]$ Workspace = $\{T_{[*case:ext*]}, \ldots\}$	MLC	MC	AC
☞ O_{11}: $[_{vP}\ DP_{[case:int]}\ [_{v'}\ v \ldots DP_{[case:\square]} \ldots]]$			
O_{12}: $[_{vP}\ DP_{[case:\square]}\ [_{v'}\ v \ldots DP_{[case:int]} \ldots]]$	*!		

14)). (Also, DP_{ext} does not intervene, given the definition of the MLC).

(86) *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))

(87) a. *Agree before Merge: accusative* 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

T_{41}: Ergative pattern, step 3: Merge

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

T_{42}: 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 ... $DP_{[case:ext]}$ ... $]]]$			

T_{43}: 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 ... $DP_{[case:int]}$ ... $]]$	*!		

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.

5.3.3 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 (88)), but everything can be transferred to the other features as well.

(88) *Gender agreement with dative possessors in German:*

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

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

Analysis:
(i) DP_{dat} is merged as a complement of the possessee (de Vries (2005)) and undergoes [•EPP•]-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.

T_{44}: 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 ... $DP_{[case:int]}$ ... $]]$	*!		

Consequence:
It follows that the pronoun has a dual role: It bears [*gen:□*] probes that trigger Agree and an [•EPP•]-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.

(89) MLC ≫ AC ≫ MC ≫ LR (Ranking for German)

Analysis:
Suppose we want to derive (88-a). We enter the derivation at stage Σ. Due to AC ≫ 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_{45}). 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 [$gen_{\sqrt{}}$:masc] on the pronoun (see O_1 in T_{46}). Finally, MC can be satisfied by movement of the possessor DP to SpecD (this optimization is skipped).

T_{45}: Evaluation of gender inflection: Agree

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

T_{46}: Evaluation of root's gender and possessor's case: Agree

Input: $[_{DP_1}$ $D_{[*case:dat*],[*gen_{\sqrt{}}:\square*],[gen_{infl}:fem],[\bullet EPP\bullet]}$ $[_{NP}$ $N_{[gen:fem]}$ $DP_{2[case:\square],[gen:masc]}$ $]]$	MLC	AC	MC	LR
☞ O_1: $[_{DP_1}$ $D_{[case:dat],[gen_{\sqrt{}}:masc],[gen_{infl}:fem],[\bullet EPP\bullet]}$ $[_{NP}$ $N_{[gen:fem]}$ $DP_{2[case:dat],[gen:masc]}$ $]]$			*	
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, both DP_{dat} and NP are equally close to the pronoun. Then the inflectional probe can receive value [masc], deriving (88-b) (see O_2 in T_{47}): Thus the approach overgenerates.

T_{47}: Phrasal optimization: wrong result

Input: $D_{[*case:dat*],[*gen_{\surd}:\Box*],[*gen_{infl}:\Box*],[\bullet EPP\bullet],...}$ + $[_{NP}\ N_{[gen:fem]}\ DP_{2[case:\Box],[gen:masc]}]$	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: $[_{DP_1}\ DP_{2[case:dat],[gen:masc]}\ D_{[case:dat],[*gen_{\surd}:\Box*],...}\ [_{NP}\ N_{[gen:fem]}\ t_2\ ]]$		*!		

5.3.4 Conclusion

(90) *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.

Outlook:
There are certain repair phenomena that lend themselves to an optimality-theoretic solution but initially raise problems in a local optimization approach because they involve long dependencies.

(91) *Optimization procedures that are apparently non-local*:
- a. Resumptive pronouns (as in (26), see Legendre, Smolensky & Wilson (1998), Pesetsky (1998))
- b. Long-distance binding (see Fischer (2004))

Hypothesis:
These phenomena all involve issues of *morphological realization in chains*. In these cases (and only in these), optimization can be non-local: Chains provide *wormholes* for post-syntactic spell-out (Fischer (2004; 2006)).

6. Ubiquity of Optimization

6.1 Background

State of affairs:
As a common research program, optimality-theoretic syntax is not well.
(i) There are no regular workshops expressly devoted to optimality-theoretic syntax anymore. (There were such workshops on optimality-theoretic syntax between 1997 and 2002, originally initiated by Sten Vikner at Stuttgart University, and there were several such meetings in the US in the second half of the last decade.)
(ii) Very few optimality-theoretic syntax papers have appeared in leading journals over the last few years.
(iii) The few papers that have appeared in the last years do not seem to share common research goals, do not tackle similar questions, and regularly do not cite other recent work in optimality-theoretic syntax.
(iv) For some time, new edited volumes with a focus on optimality-theoretic syntax appeared regularly (see, e.g., Archangeli & Langendoen (1997), Dekkers et al. (2001), Fanselow & Féry (2002b), Legendre, Grimshaw & Vikner (1998), Müller & Sternefeld (2001), Sells (2001b)). This seems to have stopped. (Note also that the working paper volumes Vogel & Broekhuis (2006) and Broekhuis & Vogel (2008) on "Optimality Theory and Minimalism" both have only few contributions that might rightfully be subsumed under the label "optimality-theoretic syntax". And the book series *Advances in Optimality Theory*, edited by Ellen Woolford and Armin Mester (Equinox publishing), does not seem to have a single volume yet that would be (mainly) on syntax, let alone a syntax monograph.)
(v) All this is very different from the situation in phonology, morphology, semantics, and pragmatics (see particularly the work on bidirectional OT going back to Blutner (2000) and Jäger & Blutner (2000), and thriving to this day). That said:

The legacy of optimality-theoretic syntax:
Optimality-theoretic syntax lives on. Its key mechanisms are implicit in much recent (and not so recent) work in the Principles and Parameters tradition, and optimization procedures arguably form an indispensible part of the minimalist program.

6.2 Hidden Optimization

Observation:
Implicit optimization procedures (that must be construed with violable and ranked constraints) show up again and again in work in the Principles and Paramters tradition.

Some instances of hidden optimization:

- OblControl ≫ *Pron
  analysis of pronoun vs. PRO in Englisch gerunds via Avoid Pronoun in Chomsky (1981); see Müller (2000b)
- Top/pro ≫ *Pron
  analysis of pro vs. overt pronoun in pro-drop languages via Avoid Pronoun in Haegeman (1995); see Müller (2000b)
- Faith(Lex) ≫ OblControl
  analysis of lexical vs. structural control in German in von Stechow & Sternefeld (1988); see Müller (2000b)
- IP-Crit ≫ Wh-Crit
  analysis of complementizer-finality and the absence overt *wh*-movement in Japanese in Kayne (1994); see Müller (2000b)
- *Complex-Wh ≫ Wh-Real
  analysis of multiple *wh*-questions in German in Grewendorf (2001); see Müller (2000b)
- *ComplexHead ≫ Head-Real
  phonological realization in head chains in Roberts (1997); see Müller (2000b)
- N-Def, D-Def, HMC ≫ N-to-D ≫ *Dissoc, Full-Int (Swedish)
  D-Def, HMC ≫ N-to-D ≫ *Dissoc, Full-Int, N-Def (Danish)
  analysis of definiteness marking in Swedish and Danish DPs in Embick & Noyer (2001); see Heck, Müller & Trommer (2008)
- Minimality, Feature Matching ≫ Agree
  analysis of Agree in Haegeman & Lohndal (2008); see Lahne (2009)
- Check-F, Parse ≫ *Struc ≫ Stay
  analysis of strong and weak pronouns in Cardinaletti & Starke (1999); see Samek-Lodovici (2006)

An example: phonological realization in head chains in Roberts (1997)

(92) *Restructuring and clitic climbing in Italian*:
Gianni li_1 *vuole* tutti *leggere* t_1
Gianni sie will alle lesen

Generalization:
Infinitives are transparent for extraaction if the non-finite verb incorporates into the matrix verb (see Baker (1988)).

Assumption (Roberts (1997)):
Incorporation has in fact taken place in (92); however, it cannot easily be detected because a non-highest head chain member is PF-realized.

(93) a. $*[_X\ \text{W}_1\ \text{W}_2\ ]$, where W_n is a morphological word.

b. A head is spelled out in the highest position of its chain, *subject to [(93-a)]*.

The optimality-theoretic reconstruction is straightforward:

(94) a. *ComplexHead:
$*[_X\ \text{W}_1\ \text{W}_2\ ]$, where W_n is a morphological word.
b. Head-Real:
A phonologically realized head must not be c-commanded by a copy in its chain.
c. *Ranking*: *ComplexHead ≫ Head-Real

T_{48}: Restructuring and clitic climbing in Italian

Candidates	*ComplexHead	Head-Real
☞ O_1: Gianni li_1 vuole t_2 tutti $[_{AGR_SP}$ leggere_2 t_2 t_1]		*
O_2: Gianni li_1 vuole leggere_2 tutti $[_{AGR_SP}$ t_2 t_2 t_1]	*!	
O_3: Gianni li_1 vuole t_2 tutti $[_{AGR_SP}$ t_2 leggere_2 t_1]		**!

6.3 Optimization in the Minimalist Program

Observation:
At the heart of the minimalist program are *elementary operations* like Agree, Merge, Move, Delete, Transfer, Select. Given that each operation is supposed to apply as soon as its context for application is present (a general Earliness requirement on derivations), it is clear that there will be conflicts. These conflicts have to be resolved by postulating ranking and minimal violability of constraints.

(95) Merge before Move (Chomsky (1995; 2000; 2001; 2005); formulation based on Frampton & Gutmann (1999)):
Suppose that the derivation has reached stage Σ_n, and Σ_{n+1} is a legitimate instance of Merge, and Σ'_{n+1} is a legitimate instance of Move. Then, Σ_{n+1} is to be preferred over Σ'_{n+1}.

(96) *The classic effect*:
a. There_1 seems $[_{\text{TP}}$ t_1 to be $[_{\text{PP}}$ someone_2 in the room]]
b. $*\text{There}_1$ seems $[_{\text{TP}}$ someone_2 to be $[_{\text{PP}}$ t_2 in the room]]

(97) *Optimization of derivational steps: T' as input*:
a. $[_{\text{T'}}$ to be $[_{\text{PP}}$ someone_2 in the room]] →
b. $[_{\text{TP}}$ there $[_{\text{T'}}$ to be $[_{\text{PP}}$ someone_2 in the room]]] (Merge)
c. $*[_{\text{TP}}$ someone_2 $[_{\text{T'}}$ to be $[_{\text{PP}}$ t_2 in the room]]] (Move)

(98) *Avoidance of the effect if no expletive is present in the numeration*:
a. Someone_2 seems $[_{\text{TP}}$ t'_2 to be t_2 in the room]
b. $[_{\text{TP}}$ someone_2 to be t_2 in the room]

(99) *A potential problem for Merge before Move, part (i)*:
a. $*\text{John}_1$ expected $[_{\text{TP}}$ t_1 to be a proof_2 discovered] (Merge)

b. John$_1$ expected [$_{TP}$ a proof$_2$ to be t$_2$ discovered] (Move)

(100) *A potential problem for Merge before Move, part (ii)*:

a. *C [$_{T'}$ Was a proof$_2$ discovered] (Merge)

b. [$_{CP}$ C [$_{TP}$ A proof$_2$ was t$_2$ discovered] (Move)

Problem:
Why is Move not blocked by Merge (of the external argument or the complementizer) after the generation of T′ in (99) and (100)?

Solution:
Merge before Move in (95) demands that the preferred option be *legitimate*. Merge of the external argument *John*$_1$ in (99) violates the Theta-Criterion; and Merge of the complementizer in (100) violates subcategorization requirements of T.

(101) *Another potential problem for Merge before Move, part (i)*:

a. *It$_1$ seems [$_{CP}$ (that) t$_1$ was told John [$_{CP}$ that Bill left]] (Merge)

b. It$_1$ seems [$_{CP}$ (that) John$_2$ was told [$_{CP}$ that Bill left]] (Move)

(102) *Another potential problem for Merge before Move, part (ii)*:

a. *It$_1$'s fun [$_{CP}$ t$_1$ to [$_{vP}$ PRO$_2$ go to the beach]] (Merge)

b. It$_1$'s fun [$_{CP}$ PRO$_2$ to [$_{vP}$ t$_2$ go to the beach]] (Move)

c. *It$_1$ was decided [$_{CP}$ t$_1$ to [$_{vP}$ PRO$_2$ be executed at dawn]] (Merge)

d. It$_1$ was decided [$_{CP}$ PRO$_2$ to [$_{vP}$ t$_2$ be executed at dawn]] (Move)

Problem:
Why doesn't the first example in (101) and in (102) block the second one because of Merge before Move? Here, the expletive must be part of the numeration.

Chomsky's (2000) solution via lexical subarrays:

> Suppose ... that at each stage of the derivation a subset LA$_i$ is extracted, placed in active memory (the 'work space'), and submitted to the procedure L. When LA$_i$ is exhausted, the computation may proceed if possible. Or it may return to LA and extract LA$_j$, proceeding as before. The process continues until it terminates. Operative complexity in some natural sense is reduced, with each derivation accessing only part of the LA. If the subarray in active memory does not contain Expl, then Move can take place in the corresponding stage; if it does, Merge of Expl preempts Move. The next step is to determine the subarrays LA$_i$ that can be selected for active memory. LA$_i$ should determine a natural syntactic object ... the counterpart to a proposition. ... LA$_i$ can then be selected straightforwardly: LA$_i$ contains an occurrence of C or of v ... – exactly one occurrence if it is restricted as narrowly as possible.
>
> Chomsky (2000)

Thus, each LA$_i$ corresponds to a phase.

Consequence:
In *there*-constructions (where Merge before Move effects obtain), the expletive and the

DP always have to be in the same LA$_i$. In *it*-constructions (where there are no Merge before Move effects), the expletive and the DP do not have to be in the same LA$_i$; they can be in different subarrays.

Optimality-theoretic reconstruction:

(103) a. Merge Condition (MC) (neutral version; compare (76):
Merge (external Merge) applies if its context for application is met.

b. Move Condition (MoveC):
Move (internal Merge) applies if its context for application is met.

c. *Ranking*: MC ≫ MoveC.

T$_{49}$: Merge before Move: 'there'-constructions

Input: [$_{T'}$ to be [$_{PP}$ someone$_2$ in the room]]	MC	MoveC
☞ O$_1$: [$_{TP}$ there [$_{T'}$ to be [$_{PP}$ someone$_2$ in the room]]		*
O$_2$: [$_{TP}$ someone$_2$ [$_{T'}$ to be [$_{PP}$ t$_2$ in the room]]	*!	

T$_{50}$: Blocked Merge before Move: Intervention of higher-ranked constraints

Input: [$_{T'}$ to be a proof$_2$ discovered]	Θ-Crit	MC	MoveC
O$_1$: [$_{TP}$ John$_1$ to be a proof$_2$ discovered]	*!		*
☞ O$_2$: [$_{TP}$ a proof$_2$ to be t$_2$ discovered]		*	

Another example (Chomsky (2001)):
The Inclusiveness condition must be minimally violable in favour of the requirement that intermediate steps of successive-cyclic movement proceed via *edge feature insertion*: Edge features on phase heads are not part of the numeration.

7. Abstractness of Optimality-Theoretic Syntax

Example:
Legendre, Smolensky & Wilson (1998) on accounting for *wh*-island effects without relying on the concept of intervention (as in Rizzi (1990; 2004)).

Assumption:
(i) All instances of extraction from a CP (even a complement CP) are assumed to violate locality constraints (most phrases are local domains), but such constraint violation is possible if it is (a) minimal, and (b) forced by higher-ranked constraints.
(ii) The constraint that forces *wh*-movement (e.g., Wh-Crit) cannot play this role: It is ranked lower than the locality constraint (Loc) that an item (more precisely, an adjunct) violates when it undergoes extraction from an object CP.
(iii) However, a third constraint (Sel) that is highest-ranked (in effect: inviolable) demands that (lexically determined) selection requirements are respected. One such require-

ment is that a V that selects an interrogative CP finds a *wh*-item in the embedded SpecC (or C) position.

(104) *Wh-Islands: Optimality Theory*:

a. How$_1$ do you think [$_{CP}$ Mary fixed the car t$_1$] ?

b. *How$_1$ does she know [$_{CP}$ [$_{DP_2}$ which car] Mary fixed t$_2$ t$_1$] ?

Consequence:
(i) Extraction of the adjunct violates locality constraints on movement (LOC) in exactly the same way in (104-a) and (104-b): There is no intervention effect induced by *which car* in (104-b).
(ii) (104-b) is ungrammatical because it blocked by a competing candidate with a better constraint profile: (105) violates the constraint that would normally trigger *wh*-movement to the matrix clause (WH-CRIT), but since this constraint is ranked lower than the locality constraint violated with extraction from all CPs (LOC), (105) is the optimal candidate.

T_{51}: *Wh-islands for long wh-movement of adjuncts*

Kandidaten	SEL	LOC	WH-CRIT
C$_1$: how$_1$ C$_{[+wh]}$ ... V$_{[+wh]}$ [$_{CP}$ t$'_1$ ... t$_1$ ...]		*!	
☞ C$_2$: C$_{[+wh]}$ ... V$_{[+wh]}$ [$_{CP}$ how$_1$ ... t$_1$ ...]			*

(105) *Optimal candidate blocking wh-movement from wh-island*:
She knows [$_{CP}$ how$_1$ [$_{DP}$ which car$_2$] Mary fixed t$_2$ t$_1$]

Problem:
(105) does not look like a well-formed English sentence.

Solution:
Given the copy theory of movement, *wh*-in situ may involve multiple *wh*-movement with selective PF realization of copies, such that one of the traces (rather than its antecedent) is PF-realized in English (Pesetsky (2000), Fanselow & Ćavar (2001), Grewendorf (2001).

Problem:
Why is (104-a) possible after all?

Solution:
The matrix verb in (104-a) selects a declarative CP complement, and not an interrogative CP complement; but if the *wh*-phrase *how* stays in the embedded SpecC position, the embedded clause will have to be interpreted as a *wh*-clause. This would violate highest-ranking SEL.

Note:
A third candidate in which the *wh*-phrase stays in situ throughout the derivation must also be considered. As a matter of fact, as it stands, this output O$_3$ would qualify as optimal in

T_{52}: *Transparent declarative clauses*

Kandidaten	SEL	LOC	WH-CRIT
C$_1$: how$_1$ C$_{[+wh]}$ ... V$_{[-wh]}$ [$_{CP}$ t$'_1$ ... t$_1$ ...]		*	
☞ C$_2$: C$_{[+wh]}$ ... V$_{[-wh]}$ [$_{CP}$ how$_1$ ... t$_1$ ...]	*!		*

both T$_{51}$ and T$_{52}$. This problem is solved if it is assumed that Gen requires *wh*-elements to show up in non-trivial chains (with pronunciation a matter of PF realization, as before).

References

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.

Aissen, Judith (1999): Markedness and Subject Choice in Optimality Theory, *Natural Language and Linguistic Theory* 17, 673–711.

Aissen, Judith (2002): Bidirectional Optimization and the Problem of Recoverability in Head Marking Languages. Ms., University of California, Santa Cruz.

Aissen, Judith (2003): Harmonic Alignment in Morphosyntax. Ms., University of California, Santa Cruz.

Anttila, Arto (1997): Variation in Finnish Phonology and Morphology. PhD thesis, Stanford University.

Archangeli, Diana & Terence Langendoen, eds. (1997): *Optimality Theory. An Overview*. Blackwell, Oxford.

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

Barbiers, Sjef (2002): Remnant Stranding and the Theory of Movement. In: *Dimensions of Movement*. Benjamins, Amsterdam, pp. 47–67.

Blutner, Reinhard (2000): Some Aspects of Optimality in Natural Language Interpretation, *Journal of Semantics* 17, 189–216.

Boeckx, Cedric (2009): Merge Alpha: Movement and Filtering. Ms., ICREA/UAB, Barcelona.

Boersma, Paul & Bruce Hayes (2001): Empirical Tests of the Gradual Learning Algorithm, *Linguistic Inquiry* 32, 45–86.

Bresnan, Joan (2001): The Emergence of the Unmarked Pronoun. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 113–142.

Bresnan, Joan, Ashwini Deo & Devyani Sharma (2007): Typology in Variation: A Probabilistic Approach to 'be' and 'n't' in the 'Survery of English Dialects', *English Language and Linguistics* 11, 301–346.

Bresnan, Joan, Shipra Dingare & Christopher Manning (2001): Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi. In: *Proceedings of the LFG 2001 Conference*. CSLI Publications.

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.

Broekhuis, Hans & Ralf Vogel, eds. (2008): *Optimality Theory and Minimalism: Interface Theories*. Vol. 28, Linguistics in Potsdam, Universität Potsdam.

Cardinaletti, Anna & Michal Starke (1999): The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns. In: H. v. Riemsdijk, ed., *Clitics in the Languages of Europe*. Mouton de Gruyter, Berlin, pp. 145–235.

Chomsky, Noam (1957): *Syntactic Structures*. Mouton, The Hague and Paris.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

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

Chomsky, Noam (1991): Some Notes on Economy of Derivation and Representation. In: R. Freidin, ed., *Principles and Parameters in Comparative Grammar*. MIT Press, Cambridge, Mass., pp. 417–454.

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

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

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

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

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.

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

Collins, Chris (1994): Economy of Derivation and the Generalized Proper Binding Condition, *Linguistic Inquiry* 25, 45–61.

de Hoop, Helen & Andrej Malchukov (2008): Case-Marking Strategies, *Linguistic Inquiry* 39, 565–587.

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

Dekkers, Joost, Frank van der Leeuw & Jeroen van de Weijer, eds. (2001): *Optimality Theory. Phonology, Syntax, and Acquisition*. Oxford University Press, Oxford.

Embick, David & Rolf Noyer (2001): Movement Operations after Syntax, *Linguistic Inquiry* 32, 555–595.

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

Fanselow, Gisbert & Caroline Féry (2002a): Ineffability in Grammar. In: G. Fanselow & C. Féry, eds., *Resolving Conflicts in Grammars*. Buske, Hamburg, pp. 265–307.

Fanselow, Gisbert & Caroline Féry, eds. (2002b): *Resolving Conflicts in Grammar*. Special volume of *Linguistische Berichte*, Buske, Hamburg.

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 (2001): On the Integration of Cumulative Effects into Optimality Theory. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 151–173.

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.

Fox, Danny (2000): *Economy and Semantic Interpretation*. MIT Press, Cambridge, Mass.
Frampton, John & Sam Gutmann (1999): Cyclic Computation, *Syntax* 2, 1–27.
Gallego, Ángel (2007): Phase Theory and Parametric Variation. PhD thesis, Universitat Autónoma de Barcelona, Barcelona.
Georgi, Doreen (2009): Local Modelling of Global Case Splits. Master's thesis, Universität Leipzig.
Grewendorf, Günther (2001): Multiple Wh-Fronting, *Linguistic Inquiry* 32, 87–122.
Grimshaw, Jane (1997): Projection, Heads, and Optimality, *Linguistic Inquiry* 28, 373–422.
Grimshaw, Jane (2001): Economy of Structure in OT. Ms., Rutgers University.
Grimshaw, Jane (2006): Chains as Unfaithful Optima. Ms., Rutgers University. Also in *Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince*; ROA 844, 97-110.
Haegeman, Liliane (1995): *An Introduction to Government and Binding Theory*. 2 edn, Blackwell, Oxford.
Haegeman, Liliane & Terje Lohndal (2008): Negative Concord is Not Multiple Agree. Poster presentation, NELS 39, Cornell University.
Haider, Hubert (1988): Zur Struktur der deutschen Nominalphrase, *Zeitschrift für Sprachwissenschaft* 7, 32–59.
Hale, Ken (1972): A New Perspective on American Indian Linguistics. In: A. Ortiz, ed., *New Perspectives on the Pueblos*. University of New Mexico Press, Albuquerque, pp. 87–103.
Halle, Morris & Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.
Hayes, Bruce (2001): Gradient Well-Formedness in Optimality Theory. In: J. Dekkers, F. van der Leeuw & J. van de Weijer, eds., *Optimality Theory. Phonology, Syntax, and Acquisition*. Oxford University Press, Oxford, pp. 88–120.
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 & Jochen Trommer (2008): A Phase-Based Approach to Scandinavian Definiteness Marking. In: C. B. Chang & H. J. Haynie, eds., *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Cascadilla Proceedings Project, Somerville, MA, pp. 226–233.
Heck, Fabian, Gereon Müller, Ralf Vogel, Silke Fischer, Sten Vikner & Tanja Schmid (2002): On the Nature of the Input in Optimality Theory, *The Linguistic Review* 19, 345–376.
Hornstein, Norbert (2001): *Move. A Minimalist Theory of Construal*. Blackwell, Oxford.
Jäger, Gerhard & Reinhard Blutner (2000): Against Lexical Decomposition in Syntax. In: A. Wyner, ed., *Proceedings of IATL*. Vol. 15, University of Haifa, pp. 113–137.
Kager, René (1999): *Optimality Theory*. Cambridge University Press, Cambridge.
Kayne, Richard (1994): *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.
Keine, Stefan (2009): Impoverishment Effects on Agreement. Master's thesis, Universität Leipzig. To appear in the series *Linguistische Arbeiten*; Tübingen: Niemeyer.
Keine, Stefan & Gereon Müller (2008): Differential Argument Encoding by Impoverishment. In: M. Richards & A. Malchukov, eds., *Scales*. Vol. 86 of *Linguistische Arbeitsberichte*, Universität Leipzig, pp. 83–136.
Keine, Stefan & Gereon Müller (2009): Non-Zero/Non-Zero Alternations in Differential Object Marking. Ms., Universität Leipzig. To appear in *Proceedings of NELS* 39; longer version in *Proceedings of the Cologne Workshop on Transitivity*.
Kiparsky, Paul (1999): Analogy and OT: Morphological Change as Emergence of the Unmarked. Vortrag auf der 21. Jahrestagung der DGfS, Konstanz. Ms., Stanford University.
Kiparsky, Paul (2001): Structural Case in Finnish, *Lingua* 111, 315–376.
Koster, Jan (1986): The Relation Between Pro-Drop, Scrambling, and Verb Movements. Ms., Rijksuniversiteit Groningen.
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.
Lee, Hanjung (2003): Parallel Optimization in Case Systems. Ms., University of Minnesota, Twin Cities.
Legendre, Géraldine (2001): An Introduction to Optimality Theory in Syntax. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 1–27.
Legendre, Géraldine (2009): The Neutralization Approach to Ineffability in Syntax. In: C. Rice & S. Blaho, eds., *Modeling Ungrammaticality in Optimality Theory*. Advances in Optimality Theory, Equinox Publishing, London.
Legendre, Géraldine, Jane Grimshaw & Sten Vikner, eds. (1998): *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass.
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 (2002): *A Thematic Guide to Optimality Theory*. Cambridge University Press, Cambridge.
Merchant, Jason (2001): *The Syntax of Silence - Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press, Oxford.
Müller, Gereon (1997): Partial Wh-Movement and Optimality Theory, *The Linguistic Review* 14, 249–306.
Müller, Gereon (1998): *Incomplete Category Fronting*. Kluwer, Dordrecht.
Müller, Gereon (2000a): Das Pronominaladverb als Reparaturphänomen, *Linguistische Berichte* 182, 139–178.
Müller, Gereon (2000b): *Elemente der optimalitätstheoretischen Syntax*. Stauffenburg, Tübingen.
Müller, Gereon (2000c): 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 (2003a): 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 (2003b): Optionality in Optimality-Theoretic Syntax. In: L. Cheng & R. Sybesma, eds., *The Second GLOT International State-of-the-Article Book. The Latest in Linguistics*. Mouton de Gruyter, Berlin, pp. 289–321.
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 & Wolfgang Sternefeld (2001): The Rise of Competition in Syntax: A Synopsis. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 1–68.
Murasugi, Kumiko (1992): Crossing and Nested Paths. PhD thesis, MIT, Cambridge, Mass.
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.
Pesetsky, David (2000): *Phrasal Movement and Its Kin*. MIT Press, Cambridge, Mass.
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.
Riemsdijk, Henk van (1978): *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Foris, Dordrecht.
Rizzi, Luigi (1990): *Relativized Minimality*. MIT Press, Cambridge, Mass.
Rizzi, Luigi (2004): Locality and Left Periphery. In: L. Rizzi, ed., *The Structure of CP and IP. The Cartography of Syntactic Structures, vol. 2*. Oxford University Press.
Roberts, Ian (1997): Restructuring, Head Movement, and Locality, *Linguistic Inquiry* 28, 423–460.
Salzmann, Martin (2006): Variation in Resumption Requires Violable Constraints. A Case Study in Alemannic Relativization. In: H. Broekhuis & R. Vogel, eds., *Optimality Theory and Minimalism: Interface Theories*. Vol. 28, Linguistics in Potsdam, Potsdam, pp. 99–132.
Samek-Lodovici, Vieri (2006): Optimality Theory and the Minimalist Program. In: R. Vogel & H. Broekhuis, eds., *Optimality Theory and Minimalism: A Possible Convergence?*. Vol. 25, Linguistics in Potsdam, Potsdam, pp. 77–97.
Schmid, Tanja (2005): *Infinitival Syntax. Infinitivus Pro Participio as a Repair Strategy*. Benjamins, Amsterdam.
Sells, Peter (2001a): *Structure, Alignment and Optimality in Swedish*. Stanford Monographs in Linguistics, CSLI Publications, Palo Alto.
Sells, Peter, ed. (2001b): *Formal and Empirical Issues in Optimality Theoretic Syntax*. CSLI, Stanford.
Shlonsky, Ur (1992): Resumptive Pronouns as a Last Resort, *Linguistic Inquiry* 23, 443–468.
Sigurðsson, Halldór (2002): To be an Oblique Subject: Russian vs. Icelandic, *Natural Language and Linguistic Theory* pp. 691–724.
Silverstein, Michael (1976): Hierarchy of Features and Ergativity. In: R. Dixon, ed., *Grammatical Categories in Australian Languages*. Australian Institute of Aboriginal Studies, Canberra, pp. 112–171.

Smolensky, Paul (1996): On the Comprehension/Production Dilemma in Child Language, *Linguistic Inquiry* 27, 720–731.
Smolensky, Paul (2006): Harmonic Completeness, Local Constraint Conjunction, and Feature Domain Markedness. In: P. Smolensky & G. Legendre, eds., *The Harmonic Mind*. Vol. II, MIT Press, Cambridge, Mass., chapter 14, pp. 27–160.
Smolensky, Paul & Geraldine Legendre (2006): *The Harmonic Mind*. MIT Press, Cambridge, Mass.
Steddy, Sam & Vieri Samek-Lodovici (2009): On the Ungrammaticality of Remnant Movement in the Derivation of Greenberg's Universal 20. Ms., University College London.
Sternefeld, Wolfgang (2003): Syntax. Eine merkmalsbasierte Analyse. Book ms., Universität Tübingen.
Stiebels, Barbara (2000): Linker Inventories, Linking Splits and Lexical Economy. In: B. Stiebels & D. Wunderlich, eds., *Lexicon in Focus*. Akademie-Verlag, Berlin, pp. 211–245.
Stiebels, Barbara (2002): *Typologie des Argumentlinkings: Ökonomie und Expressivität*. Akademie Verlag, Berlin.
Swart, Peter de (2007): Cross-Linguistic Variation in Object Marking. PhD thesis, Nijmegen University. LOT Dissertations 168.
Vogel, Ralf (2009): Wh-Islands: A View from Correspondence Theory. In: C. Rice & S. Blaho, eds., *Modeling Ungrammaticality in Optimality Theory*. Advances in Optimality Theory, Equinox Publishing, London.
Vogel, Ralf & Hans Broekhuis, eds. (2006): *Optimality Theory and the Minimalist Program*. Vol. 25, Linguistics in Potsdam, Universität Potsdam.
von Stechow, Arnim & Wolfgang Sternefeld (1988): *Bausteine syntaktischen Wissens*. Westdeutscher Verlag, Opladen.
Wiklund, Anna-Lena (2001): Dressing Up For Vocabulary Insertion: The Parasitic Supine, *Natural Language and Linguistic Theory* 19, 199–228.
Williams, Edwin (1997): Blocking and Anaphora, *Linguistic Inquiry* 28, 577–628.
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 (2001): Case Patterns. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 509–543.
Wunderlich, Dieter (1997): Der unterspezifizierte Artikel. In: C. Dürscheid, K. H. Ramers & M. Schwarz, eds., *Sprache im Fokus*. Niemeyer, Tübingen, pp. 47–55.
Wunderlich, Dieter (2000): Optimal Case in Hindi. Ms., Universität Düsseldorf.
Wunderlich, Dieter (2003): The Force of Lexical Case: German and Icelandic Compared. Ms., Universität Düsseldorf. To appear in Kristin Hanson and Sharon Inkelas (eds.) The Nature of the Word: essays in honor of Paul Kiparsky. MIT Press 2003.
Zifonun, Gisela (2004): Dem Vater sein Hut – Der Charme des Substandards und wie wir ihm gerecht werden, *Deutsche Sprache* 03, 97–126.