26. Optimality-Theoretic Syntax

Gereon Müller (Universität Leipzig) July 22, 2012*

Contents

Abstract

This chapter lays out the basic assumptions and workings of optimality-theoretic syntax. After sketching central aspects of the model of grammar presupposed in optimality theory, I address four pieces of evidence for optimality-theoretic approaches to syntax in the first core section (viz., the ubiquity of constraint conflict, the existence of repair phenomena, the notion of default, and cross-linguistic variation). The second core section is devoted to three potential problems with optimality-theoretic syntax that have been identified in the literature (viz., complexity, ineffability, and optionality). In the third core section, the concept of optimization domain is subjected to scrutiny, and with it the question of whether syntactic optimization proceeds serially or in parallel. Finally, some conclusions are drawn concerning the prospects of optimality-theoretic syntax, particularly with respect to recent developments in the minimalist program.

1. 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). OT shares with most other grammatical theories the assumption that *constraints* are crucial in restricting the class of possible linguistic expressions (LEs) in natural languages; however, it differs in important ways from virtually all other grammatical theories in that it envisages a non-trivial interaction of constraints. More specifically, OT rests on four basic assumptions: First, constraints are universal (*universality*). Second, constraints are violable (*violability*). Third, constraints are ranked (*ranking*). And fourth, the wellformedness of an 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 (*competition*): LEs are *candidates*.[1] None of these assumptions is shared by standard grammatical theories like Chomsky's (1981) Government-Binding (GB) theory or Pollard & Sag's (1994) Head-Driven Phrase Structure Grammar. Taking GB theory as a typical example, we can first observe that here, not all constraints are universal (there are parameters and language-specific filters – but cf. 'third-factor' meta-constraints on constraints in recent work in the minimalist program, as in Chomsky (2007; 2008)). Second, constraints cannot be violated. Third, constraints are not ranked (i.e., all are equally important and do not interact).[2] Finally, the wellformedness of a linguistic expression LE (e.g., a sentence) can standardly fully be determined on the basis of LE's internal properties. External factors (i.e., the properties of other LEs) are irrelevant.

At the heart of OT is the concept of optimality of a candidate LE, which can be defined as in (1).

(1) *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 *violation profile*.

For now, we can assume that optimality equals grammaticality (or wellformedness). (1) introduces two additional concepts – that of a violation profile, and that of a candidate set. The violation profile determines which of two competing candidates is to be preferred. A concept of violation profile that in principle permits more than one candidate to be optimal is given in (2) (this is in contrast to Grimshaw (1997), which presupposes that only one candidate can be optimal in any given candidate set).

(2) *Violation profile*:
C_j has a better violation profile than C_i if there is a constraint Con_k such that (i) and (ii) hold:
a. C_j satisfies Con_k better than C_i.
b. There is no constraint Con_l ranked higher than Con_k for which C_i and C_j differ.

We can assume that a candidate C_j satisfies a constraint Con better than a candidate C_i if C_j violates Con less often than C_i. This includes, as a special case, the situation that C_j does not violate Con at all, whereas C_i does. Turning to candidate sets next, the basic task of this concept is to clarify what competes with what. Various different versions of the concept have been proposed for syntax. (3) lists some of the more widely adopted definitions.[3]

(3) *Candidate set*:
Two candidates are in the same candidate set iff
a. they have the same content words
b. they have the same words (see 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, where OT is combined with

Lexical Functional Grammar; cf. Choi (1999), Sells (2001a), Bresnan (2001), papers in Sells (2001b), and Butt & King (2012))

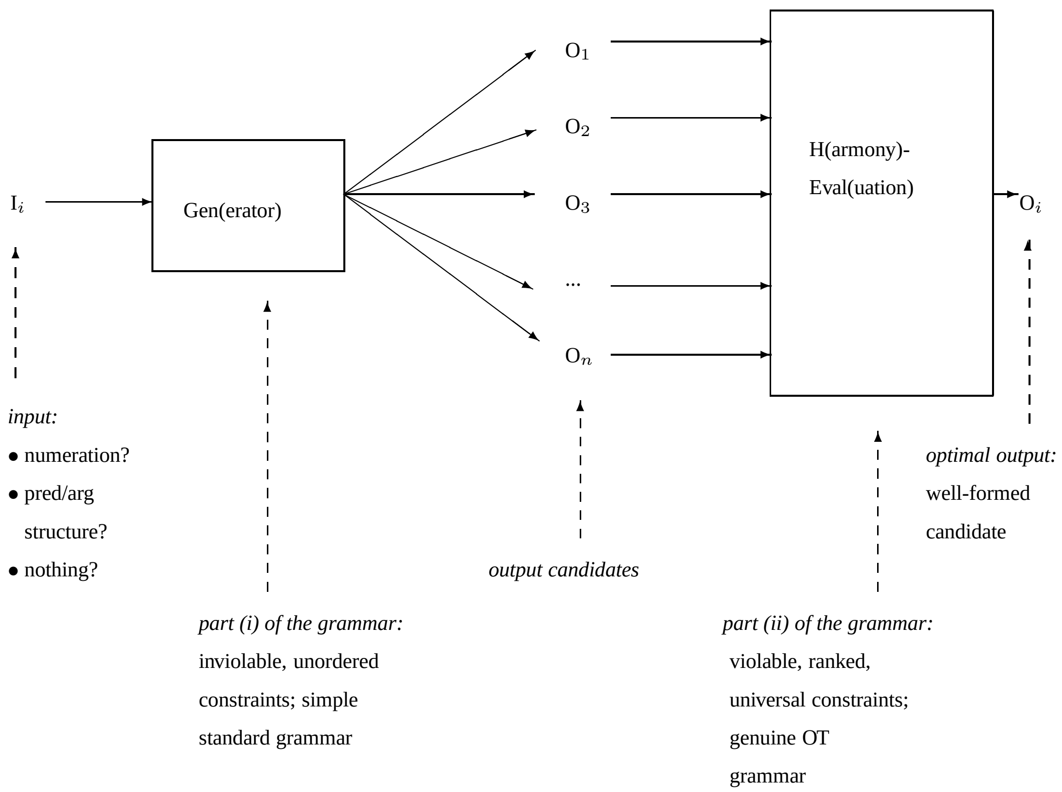
h. they have the same D-structure (see work in the GB tradition)

i. they have the same predicate/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"; see Legendre, Smolensky & Wilson (1998, 258)).

In order to be able to check candidate LEs against a set of violable and ranked constraints and resolve the competition by determining the optimal candidate in a candidate set, one must have the candidates first. In other words: The approach to syntax sketched so far presupposes that there is another, prior, component that generates the candidates.[4] The truly optimality-theoretic component of a grammar that selects a candidate with a best violation profile is often referred to as the H-EVAL ('Harmony Evaluation') part of the grammar; this component is fed by a simple standard grammar with inviolable and non-ranked constraints that is called GEN ('Generator'). The full structure of the syntax component of an OT grammar is given in figure 1.

Figure 1: Structure of an optimality-theoretic syntax component



It is clear what the H-EVAL component takes as its input (viz., the candidate set of competing output candidates generated by GEN); but, as indicated in figure 1, a major open question (in fact, arguably one of the biggest unresolved problems of OT syntax) is what GEN in turn takes as its input. For phonology, the standard OT assumption is that GEN creates outputs candidates on the basis of an input; i.e., inputs also define the candidate set (see Prince & Smolensky (2004)). Outputs then 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., underlying representations [URs]), and may even be identical. This seems hardly tenable for syntax (or for morphology) because it does not take into account the effect of *structure-building* operations: If outputs for H-EVAL are syntactic structures, and structures are generated by GEN, then where does the input structure come from if inputs are also syntactic structures? Consequently, it is at present completely unclear what the input in syntax should look like. Suggestions range from relatively poorly structured inputs (e.g., predicate/argument structures in Grimshaw (1997)) to extremely richly structured inputs (e.g., the 'index' of Legendre, Smolensky & Wilson (1998)); in fact, given that one task standardly attributed to the input is that of defining candidate sets, many of the proposals in (3) can also be viewed as proposals for concepts of inputs. What is more, it might be that there is no input in syntax at all.[5] In what follows, I will leave this issue undecided. I will continue to presuppose that inputs exist (but I will not presuppose any specific concept of input); I will equate candidates with outputs.

Standardly, two basic types of H-EVAL constraints can be distinguished in OT that often give rise to conflicts. On the one hand, there are *faithfulness constraints* that demand that input and output are identical with respect to some property. There are three basic subtypes: First, DEP ('Dependency') constraints (sometimes also referred to as FILL constraints, with subtle differences related to the overall organization of grammar that need not concern us here) state that there can be no items in the output that are not present in the input. Assuming, for instance, expletives to be absent in syntactic inputs, the occurrence of an expletive in an output will violate a DEP constraint. The same may hold for traces (or copies), assuming that syntactic inputs (whatever they ultimately look like) are unlikely to involve movement. Second, MAX ('Maximality') constraints (sometimes also referred to as PARSE constraints, with the same qualification as above) demand that all items that are present in the input are also present in the output. Thus, all kinds of deletion phenomena will incur violations of MAX constraints. Third, IDENT ('Identity') constraints prohibit the modification of items from input to output. Note that DEP, MAX, and IDENT constraints can be formulated for items of various complexity levels (e.g., feature values, features, feature bundles, lexical items, perhaps complex syntactic categories). Accordingly, MAX/DEP constraints for items with complexity *n* can often be reformulated as IDENT constraints at the next-higher complexity level *n+1*, and vice versa. E.g., deletion of a feature (a MAX violation) will give rise to a different lexical item bearing this feature in the input (an IDENT violation). Next to faithfulness constraints, there is a second basic type of H-EVAL constraint: *Markedness constraints* impose requirements on outputs that may necessitate a deviation from the input.[6]

Optimality-theoretic competitions are often illustrated by tables, so-called *tableaux*. The basic principle is illustrated in (4). There are three constraints A, B, and C, with A ranked higher than B, and B ranked higher than C (A $\gg$ B $\gg$ C). The candidate set contains five candidate outputs O_1–O_5 (typically, there are many more, but let us focus on these five for now). Violations incurred by a candidate are marked by a star (*). A decisive violation of some constraint that is responsible for eliminating the candidate by classifying it as suboptimal is here accompanied by an exclamation mark (!); this is strictly speaking redundant and is accordingly sometimes left out in tableaux. Finally, an optimal candidate is identified by the so-called pointing finger: ☞. Given the constraint violations induced by the candidates, and given the ranking of the three constraints A $\gg$ B $\gg$ C, O_1 turns out to have the (sole) best violation profile in T_1 (see definition (2)), and is therefore predicted to be optimal (see definition (1)).

T_1: The basic principle

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

Consider next the issue of cross-linguistic variation. An assumption that is not made in most minimalist approaches, but virtually everywhere else in syntactic theory (including GB theory) is that languages differ with respect to their grammars (i.e., not just the make-up of lexical items). Grammatical differences between languages are often assumed not to be completely arbitrary; this is then captured by assuming some kind of principled variation, or parametrization. Parametrization in optimality theory is simply viewed as *constraint reranking*. Thus, suppose that the ranking of constraints B and C is reversed in T_1, with the violation profile of the competing outputs remaining identical. In that case, O_3 (rather than O_1) is predicted to be optimal. This is shown in tableau T_2.

T_2: Parametrization

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

OT was developed out of so-called "harmonic grammar" approaches, which are instantiations of a more general theory of neural networks.[7] The main innovation of OT is that quality comes before quantity, in the sense that no number of violations of a lower-ranked constraint can outweigh a single violation of a higher-ranked constraint. This property is encoded in the definition of violation profile in (2); it is illustrated by the abstract competition in T_3: Even though candidates O_3 and O_4 each incur only one constraint violation in total (and O_5 only two), O_1, with four constraint violations all in all, emerges as optimal because its violations only concern the lowest-ranked constraint C. Quantity does become relevant when quality cannot decide between candidates; thus, O_2 is blocked by O_1 because it incurs more violations of the highest-ranked constraint on which the two candidates differ.

T_3: Irrelevance of constraint violation numbers as such

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

However, there is a caveat. In some versions of OT, a means has been introduced that undermines the irrelevance of constraint violation quantity as such, viz., *local conjunction* of constraints (see Smolensky (1996; 2006)). Local conjunction can be defined as in (4).

(4) *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. It may be that $\text{Con}_1 = \text{Con}_2$. (Local conjunction is reflexive.)

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

Given local conjunction, the situation can arise that the joint violation of two low-ranked constraints B, C may in fact outweigh the violation of a higher-ranked constraint A (because the complex constraint $B\&_{\text{D}}C$ derived from local conjunction may be ranked higher than A). Moreover, local conjunction can be reflexive (see (4-c)); this means that multiple violations of a single constraint may also suffice to outweigh the violation of a higher-ranked constraint. This is illustrated in T_4, which differs minimally from T_3 in that C^4 (the result of iterated reflexive local conjunction applying to C which is violated when C is violated four times or more) is present, and which produces a different winner (viz., O_3).

T_4: A consequence of reflexive local conjunction

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

T_4 should make it clear that local conjunction is far from harmless. This conclusion is reinforced by the observation that an unrestricted system of local conjunction automatically leads to a non-finite set of constraints, which is otherwise unheard of in grammatical theory. Still, it can be noted that local conjunction (reflexive or otherwise) has given rise to a number of insightful analyses of various grammatical phenomena; so there is in fact quite a bit of empirical evidence for it (see, for instance, Legendre, Smolensky & Wilson (1998) on displacement, Fischer (2001) on quantifier scope, Aissen (1999; 2003a) and Keine & Müller (2008; 2009) on differential argument encoding, and Keine (2010) on eccentric instances of case and agreement).

In what follows, I will highlight central aspects of OT syntax, focussing on issues where OT syntax substantially differs from other syntactic theories. In doing so, I will first discuss types of empirical evidence that would seem to support an OT perspective (section 2); after that I turn to kinds of data that may qualify as problematic for an OT perspective (section 3). Section 4 then addresses a topic that strikes me as potentially very important for future work in OT syntax, particularly when compared to recent developments in the minimalist program, viz., the issue of optimization domains. Finally, in section 5 I turn to the prospects for OT syntax as a framework for syntax.

2. Evidence for OT Analyses in Syntax

Central pieces of evidence for OT analyses come from the following four domains: (i) constraint conflict, (ii) repair phenomena; (iii) default contexts ('emergence of the unmarked'), and (iv) cross-linguistic variation by constraint reranking. I will address these issues in turn.

2.1 Constraint Conflict

Here the profile of the empirical evidence looks as follows. The facts show that two general and far-reaching constraints are well motivated, independently of one another. 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. Simple *wh*-movement in English is a case in point; consider (5).

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

Any grammar of English will recognize (something like) (6-a) and (6-b) as two plausible constraints: On the one hand, in simple questions, a *wh*-phrase moves to a clause-initial position (SpecC, e.g.); on the other hand, a direct object shows up in the immediate vicinity of the verb that it is an object of.

(6) a. WH-CRITERION (WH-CRIT):
Wh-items are in SpecC$_{\text{[wh]}}$.
b. θ-ASSIGNMENT (θ-ASSIGN):
Internal arguments of V are c-commanded by V.

In (5), (6-a) and (6-b) cannot both be satisfied, and the well-formedness of (5-a) suggests that it is (6-b) that has to give in the case of conflict. This conclusion cannot be drawn in standard models of grammar (that do not envisage constraint violability), though. The consequence here has to be that 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). In contrast, in OT, both constraints can be assumed to hold, but they are ranked as in (7).

(7) *Ranking*:
WH-CRIT $\gg$ θ-ASSIGN

The competition underlying (5) is illustrated in T_5.[8]

T_5: Simple wh-question formation in English

Input: John, bought, which, book, v, T, $C_{[+wh]}$	WH-CRIT	θ-ASSIGN
☞ O_1: ... which book John bought		*
O_2: ... John bought which book	*!	

Note that 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 to leave a trace (t) or a copy. On this view, the role of θ-ASSIGN can be taken over by the more general constraint ECONOMY (see Grimshaw (1997), Legendre, Smolensky & Wilson (1998), Ackema & Neeleman (1998), among others, for versions of this constraint). Given that OT, like other grammatical theories, strives for maximally simple and elegant constraints, this would seem to be a step in the right direction.[9]

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

The main conclusion concerning the role of constraint conflict is summed up in the following quote.

> 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 (2006a, 94)

2.2 Repair Phenomena

With repair phenomena, the profile of the empirical evidence is this: The facts suggest that some well-formed 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 qualify as even worse, in a sense to be made precise). Consider the distribution of resumptive pronouns in English, as indicated by the examples in (9).

(9) 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

The insertion of resumptive pronouns may (often) be viewed as a repair phenomenon, i.e., as a last resort operation that can only take place if a well-formed sentence cannot otherwise be generated (see Shlonsky (1992) and Hornstein (2001), among others). Here, a resumptive pronoun is possible if movement is blocked by an island constraint (the Complex NP Constraint, in the case at hand; see Ross (1967)); compare (9-b) (movement) with (9-d) (resumption). If movement is possible, resumption is blocked; cf. (9-ac).[10]

The insertion of a resumptive pronoun (which, by assumption, is not part of the input) violates a DEP faithfulness constraint, but is required by a higher-ranked markedness constraint. OT analyses of resumptive pronouns that employ this general logic have been developed in Pesetsky (1998), Legendre, Smolensky & Wilson (1998), and Salzmann (2006). Let us look at what a (simplified) account of the pattern in (9) could look like. Suppose that there is a constraint like REL-CRIT in (10-a) that triggers displacement in relative clauses; and that there is an island constraint like CNPC in (10-b).[11] Furthermore, there is a DEP constraint blocking insertion of resumptive pronouns in outputs. Following Chomsky (2000; 2001), this constraint may be referred to as INCLUSIVENESS; see (10-c).

(10) 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.

Suppose next that the ranking is as in (11).[12]

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

This accounts for the pattern in (9). T_6 shows two things. First, the highest-ranked REL-CRIT is not violable in an optimal output (i.e., relative operator movement is obligatory). And second, a resumptive pronoun that violates INCL is blocked if movement is possible (i.e., compatible with CNPC).

T_6: 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)	*!		

In contrast, T_7 illustrates that if movement would have to violate CNPC, resumption becomes optimal: INCL is violable as a last resort.

T_7: 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)	*!		

There are many more such instances of repair phenomena in syntax, and some of them have been given OT analyses that reflect the basic mechanism just presented. The first repair phenomenon to receive an optimality-theoretic account is *do*-support in English; see Grimshaw (1997) (Grimshaw shows that insertion of an expletive *do*, which violates a faithfulness constraint, is only possible in *wh*-contexts and negation contexts, where (partially different sets of) higher-ranked constraints conspire so as to make verb raising obligatory, and an auxiliary that can raise is not available), and also Grimshaw (2010) and references cited there. Other phenomena include the so-called *Ersatz-infinitive* (Infinitivus pro Participio) in German and the opposite phenomenon of Participium pro Infinitivo in Swedish (see Wiklund (2001), Schmid (2005), Vogel (2009a)); R-pronouns in German, Dutch and (Middle) English (see Müller (2000a)); *wh*-scope marking in German and Hungarian (see Müller (1997)); expletives in SpecC and SpecT (see Müller (2000b), Grimshaw (2006)); repair-driven quantifier raising in VP ellipsis contexts in English (as identified in Fox (2000); see Heck & Müller (2000; 2003)); repair-driven intermediate movement steps as required by the PIC of Chomsky (2001; 2008) (see Heck & Müller (2000; 2003)); repair-driven multiple *wh*-movement in German sluicing constructions (as identified in Merchant (2001); see Heck & Müller (2000; 2003)); and so on. For all these phenomena, the idea that a repair or last resort effect is involved looks fairly natural, and has been widely pursued in various types of syntactic theories. However, as pointed out by Grimshaw (2010), theories that lack the concept of constraint violability and constraint ranking (like virtually all current non-OT approaches) "generally appeal to the last resort idea by word and not by deed"; "the words 'last resort' are employed but the concept plays no role in the analysis."

Consider briefly three of these phenomena. The German Ersatz-infinitive pattern is given in (12).

(12) 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

The perfect auxiliary *haben* normally selects a past participle; it is incompatible with an infinitival form of its dependent main verb; see (12-a) vs. (12-b). However, when the dependent verb of the perfect auxiliary is itself (used as) a modal verb that further subcategorizes for another verb (which in turn regularly shows up as an infinitive), it has to take on the infinitival form (see (12-d)); the expected participial form cannot show up (see (12-c)). In addition, the change to the "Ersatz"-form is obligatorily accompanied by a reversal of order (at least in Standard German): VP-Aux becomes Aux-VP. The gist of an OT analysis will then consist in postulating the interaction of a faithfulness constraint demanding selection requirements to be respected (the perfect auxiliary selects a past participle) on the one hand, and a higher-ranked markedness constraint banning configurations where past participle modals embed verbal categories on the other hand; the optimal violation profile will then (ideally) automatically emerge as one with a reversal of word order.

Next, consider the case of repair-driven quantifier raising as it is documented in English VP ellipsis constructions as in (13) (deletion is indicated by crossing out words).

(13) 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]] (∃∀, ∀∃)

The observation to be explained here is that whereas (13-ac) are scopally ambiguous, (13-b) is not. Suppose, following Fox (2000), that scope reversal from the linear order requires quantifier raising (QR), and that there is an economy constraint blocking QR if the same interpretation is reached without it (in particular, QR of an object quantifier is thus blocked if the subject is a proper name, as in the second (CP_2) conjunct in (13-b) and in the first (CP_1) conjunct in (13-c). Furthermore, note that VP ellipsis obeys strict parallelism: What happens in one conjunct must also happen in the other one. Finally, suppose, again following Fox, that the two conjuncts are generated one after the other, in a bottom-up, right-to-left fashion: CP_2 is then optimized before CP_1 is.[13] On this view, the ranking of the parallelism requirement above scopal economy will produce the pattern in (13). Both constraints can be fulfilled by both optimizations (applying first to CP_2, and then to CP_1) in (13-a). In (13-b), parallelism is not yet an issue in CP_2 (for CP_1 does not yet exist); so QR is blocked in CP_2. Subsequently, parallelism becomes relevant during CP_1 optimization, and since CP_2 cannot be changed anymore, it blocks scope reversal in CP_1 even though this would be semantically non-vacuous. Finally, in (13-c), QR may apply in CP_2 (since it is semantically non-vacuous), and if it does, parallelism will force it to also apply later in CP_1, even though it is not semantically motivated there – in other words, it is repair-driven.

The final repair phenomenon to be considered here is repair-driven multiple *wh*-movement in German sluicing constructions. Assume that sluicing is analyzed as *wh*-movement to a SpecC position, accompanied by deletion of the TP sister of an interrogative C.

(14) a. Irgendwer hat irgendwas geklaut, aber Kirke weiß nicht mehr [$_{CP}$ wer$_1$ was$_2$
someone has something stolen but Kirke knows not more who what
~~C t_1 t_2 geklaut hat~~]
stolen has

b. *Irgendwer hat irgendwas geklaut, aber Kirke weiß nicht mehr [$_{CP}$ wer$_1$ ~~C t_1 was$_2$~~
someone has something stolen but Kirke knows not more who what
~~geklaut hat~~]
stolen has

The interesting observation is that (assuming that all alternative parses, e.g., via *wh*-scrambling, can be excluded) (14-a) instantiates a case of multiple *wh*-movement, which is not available outside of sluicing contexts in German and would therefore seem to qualify as a repair phenomenon. An OT analysis can postulate an interaction of a general recoverability requirement that precludes a *wh*-phrase from being deleted, and a second constraint (or set of constraints) ensuring that only one *wh*-phrase can undergo movement to the specifier of an interrogative C in German; as shown by the contrast in (14-a) vs. (14-b), the first constraint outranks the second one, leading to multiple *wh*-movement in the case of conflict.

2.3 Default Contexts

The notion of default is a core concept in linguistics. The profile of the empirical evidence looks as follows: The data suggest that there is a concept like "unmarked case" ("default", "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.[14] In OT, an unmarked case signals the presence of a constraint C that is ranked very low, and that is typically rendered inactive by higher-ranked, conflicting constraints. However, when these latter constraints do not distinguish between the candidates, C becomes decisive; this state of affairs is usually referred to as the *emergence of the unmarked*.

As an example, consider the following empirical generalization: In the unmarked case, a DP bears nominative case in German; i.e., nominative is the default case. Default nominative shows up in all contexts in which the regular rules of case government do not apply. This includes the contexts in (15). (15-a) instantiates a construction in which an appositive DP introduced by *als* ('as') further specifies a DP. In principle, there is an option (not shown here) for the appositive DP to show up with the same case as the other DP (here, the genitive), via a process of case agreement (see Fanselow (1991)). However, if this option is not chosen, the appositive DP receives nominative case, as a default. A second context involves infinitival constructions with a (case-less) PRO subject (see again Fanselow (1991)); it looks as though there is no possible case-government or case-agreement source for the DP *einer* in (15-b); so it receives default nominative case. The third example in (15-c) involves left dislocation. As in the first context, there is an option of case agreement, but if this option is not chosen, the left-dislocated item bears default nominative case. Finally, (15-d) is an instance of a predicative use of *und* ('and'), which here connects a subject with an infinitival VP (see Sailer (2002)). Standardly, subjects in German bear nominative case in the presence of finite T, and accusative case if embedded under exceptional case marking (AcI; *accusativus cum infinitivo*) verbs. Since neither context is present in (15-d), there is a resort to default case.

(15) a. die Ehrung des Kanzlers als großer Politiker/ *großen Politiker
the homage to the chancellor as great politician$_{nom}$ great politican$_{acc}$
b. Wir baten die Männer [$_{\text{CP}}$ PRO einer nach dem anderen/ *einen nach dem anderen durch die Sperre zu gehen]
we asked the men one$_{nom}$ after the other one$_{acc}$ after the other through the barricade to go
c. Der Kaiser/ *Den Kaiser, dem verdanken wir nichts
the emperor$_{nom}$ the emperor$_{acc}$ him owe we nothing
d. Der/ *Den und ein Buch lesen? (Dass ich nicht lache!)
he$_{nom}$ him$_{acc}$ and a book read that I not laugh

The examples in (16) show that the nominative in (15-a-d) is indeed a default case; it is is overriden in all contexts in which rules of case-government apply (accusative assignment by, perhaps, the category v in unmarked object case contexts in (16-a), genitive assignment by a V which is lexically specified for this case in (16-b), dative assignment by, perhaps, an applicative functional head in a double object construction in (16-c)).

(16) 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/ dem Mann das Buch geben
that we the man$_{nom}$ the man$_{dat}$ the book$_{acc}$ give

The distribution of cases in (15) and (16) can (partially) be accounted for by the system of case-related constraints in (17), accompanied by the ranking in (18).[15]

(17) a. GEN(ITIVE) CONSTRAINT (GEN):
The object of a verb that is lexically marked as governing genitive case bears genitive.
b. ACC(USATIVE) CONSTRAINT (ACC):
The object of a transitive verb bears accusative case.
c. NOMINATIVE CONSTRAINT (NOM):
A DP bears nominative case.

(18) *Ranking*:
GEN $\gg$ ACC $\gg$ NOM

The competition in a typical case-government context is illustrated in tableau T_8: Nominative case is blocked on the object because a higher-ranked constraint demands accusative here.

In contrast, consider T_9: Here, all higher-ranked case-related constraints are satisfied vacu-

T_8: Accusative government

Input: dass, getroffen, habe, 1.Sg./Agent, 3.Sg./Patient	GEN	ACC	NOM
☞ O_1: dass ich$_{nom}$ ihn$_{acc}$ getroffen habe			*
O_2: dass ich$_{nom}$ er$_{nom}$ getroffen habe		*!	
O_3: dass mich$_{acc}$ ihn$_{acc}$ getroffen habe			**

ously, so the low-ranked constraint NOM springs into action and ensures nominative case on the subject of *und*.

T_9: Nominative as the unmarked case

Input: und, ein, Buch, lesen, 3.Sg./Agent/Dem	GEN	ACC	NOM
O_1: Den$_{acc}$ und ein Buch lesen ?			*!
☞ O_2: Der$_{nom}$ und ein Buch lesen ?			
O_3: Dem$_{dat}$ und ein Buch lesen ?			*!

Note that if nominative (or absolutive in ergative alignment patterns) is inherently a default case across languages, free reranking of the constraints in (18) must be blocked in some way (which may then be related to the more primitive feature structures of the cases; see, e.g., Wunderlich (1997), Kiparsky (1999; 2001)). Then again, a look at English may already suggest that other cases may also act as the default case in a language (the accusative in the case at hand).

There are many other default phenomena in natural languages, and most of them can arguably be treated straightforwardly in the same way, as an instanace of emergence of the unmarked.[16]

2.4 Cross-Linguistic Variation

The approach to case sketched in the preceding subsection relies on a system of ranked constraints *demanding* the realization of individual cases. Interestingly – and this reveals a more general pattern of OT analyses –, similar effects can be obtained under a system of ranked constraints *prohibiting* the realization of individual cases, as long as these constraints are accompanied by an inherently highest-ranked (or, as part of GEN, inviolable) constraint that states that all DPs have case. Such an approach is pursued by Woolford (2001), and it may serve to illustrate the simple way that cross-linguistic variation can be handled by reranking in OT. Here are Woolford's (2001) background assumptions. First, there are (ordered) markedness constraints that block the realization of cases; see (19-abc). Second, there are faithfulness constraints that demand the realization of case specifications in the input (i.e., the realization of lexical, inherent case). Case faithfulness constraints come in two versions: a general one that covers both intransitive and (as we will see, irrelevantly) transitive contexts in (19-d), and a more specific one for transitive contexts only in (19-e). Third, nominative (= absolutive) and accusative are assumed to be structural cases; but both dative and ergative (as well as genitive) are considered inherent cases (that must be specified on a verb). Finally, it is presupposed that every DP must be case-marked; at least for present purposes, this may be viewed as a requirement imposed by GEN.

(19) 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$_{\rm trans}$:
Realize a case feature specified on transitive V in the input.

Against this background, Woolford shows that cross-linguistic variation with respect to lexical ('quirky') case on subjects can be easily derived. The distribution of lexically case-marked subjects in Icelandic, Japanese, and English follows directly from the rankings assumed in (20).

(20) a. *Ranking in Icelandic*:
FAITH-LEX$_{tr}$ $\gg$ FAITH-LEX $\gg$ *DAT $\gg$ *ACC $\gg$ *NOM
b. *Ranking in Japanese*:
FAITH-LEX$_{tr}$ $\gg$ *DAT $\gg$ FAITH-LEX $\gg$ *ACC $\gg$ *NOM
c. *Ranking in English*:
*DAT $\gg$ FAITH-LEX$_{tr}$ $\gg$ FAITH-LEX $\gg$ *ACC $\gg$ *NOM

(20-a) correctly predicts that lexically specified case-marking on subjects in the input will always be realized in an optimal output in Icelandic (i.e., in both transitive and intransitive contexts), even if this implies a violation of a fairly high-ranked case markedness constraint like *DAT; see (21-a) (intransitive context) and (21-b) (transitive context).

(21) a. Bátnum hvolfdi
boat$_{dat}$ capsized
b. Barninu batnadhi veikin
child$_{dat}$ recovered from disease$_{nom}$

The competition underlying (21-a) is illustrated in T$_{10}$.

T$_{10}$: 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$_{[+dat]}$		*!		*	

The competition underlying the transitive context in (21-b) is illustrated in T$_{11}$. Note that given the verb's lexical dative specification for the external argument, the fact that the remaining (internal) argument must receive nominative case comes for free. This is an instance of emergence of the unmarked that is completely parallel to the approach to default case specified in the previous subsection (despite the move from demanding case realization to prohibiting case realization): An accusative (or dative, genitive, etc.) realization on the internal argument would fatally violate a markedness constraint (*ACC, etc.) that is ranked higher than the one violated by the optimal output (viz., *NOM).[17]

Furthermore, note that the more specific version of the FAITH-LEX constraint (viz., FAITH-LEX$_{\rm trans}$) is strictly speaking not yet needed for Icelandic – if it were absent, FAITH-LEX as

T_{11}: 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}	*!	*		*	*

such would suffice to exclude O_3 in T_{11}. The situation is different in Japanese, where quirky case can only show up on subjects of transitive clauses; see (22-ab).

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

A minimal reranking of *DAT and FAITH-LEX (see (20-b)) yields the Japanese pattern. T_{12} shows how lexical dative is now blocked in intransitive contexts (assuming a dative case specification on the verb).

T_{12}: 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} $V_{[+dat]}$			*	*!	

In contrast, high-ranked FAITH-LEX$_{\mathrm{trans}}$ still ensures lexically marked dative case on subjects in transitive clauses; see T_{13}.

T_{13}: 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}	*!		*	*	*

Turning finally to English, it is clear that the ranking in (20-c), with the prohibition against dative case outranking all case faithfulness constraints, ensures that there can be no lexical case on subjects in this language: Even if there is an inherent dative specification on a verb, high-ranked *DAT will not let it become optimal in the output.[18] To sum up, reranking of violable constraints offers a promising approach to parametrization of grammars. I have exemplified this with a tiny empirical domain, viz., cross-linguistic variation with respect to subjects that bear lexically marked case. Of course, there is a lot more to be said about case, and about the cross-linguistic variation encountered in this area, from an OT perspective; see Aissen (1999; 2003a), 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. Furthermore, case is by no means the only empirical domain in which parametrization by constraint reranking has proven successful; see, e.g., Legendre, Smolensky & Wilson (1998), Ackema & Neeleman (1998), and Müller (1997) on cross-linguistic variation in *wh*-movement, Grimshaw (1997) and Vikner (2001a;b) on cross-linguistic variation in verb movement, and Samek-Lodovici (1996; 2005), Costa (1998), Choi (1999), Büring (2001), Engdahl et al. (2004), Gutiérrez-Bravo (2007), and Gabriel (2010) for cross-linguistic variation in the placement of subjects, direct objects and indirect objects.

With *n* constraints, there is the logical possibility of *n!* (*n* factorial) rerankings. If no additional assumptions are made, *n!* therefore defines the possible number of grammars that can be created on the basis of a set of *n* constraints. This property of OT is accordingly often referred to as *factorial typology*. It is clear that the number of possible grammars for natural languages can become quite large this way (e.g., with a mere 12 constraints, free reranking produces more than 479 million grammars; with 13 constraints, it's already more than 6.2 billion, and so forth). In view of this (and although it is a priori far from clear whether the large number of grammars generated on the basis of a small set of constraints should in fact be viewed as problematic), strategies have been devised to narrow down the options for reranking. One such attempt relies on fixed subhierarchies of constraints, i.e., pairs of constraints whose ranking must be invariant across languages. We have encountered one such case above (see (4-b)): Local conjunction of two constraints A and B gives rise to a constraint C (= $A\&_D B$) that inherently outranks the individual constraints of which it is composed. Another restriction on free reranking follows from the concept of *harmonic alignment* (see below). In some cases, the fixed order of related constraints that differ along some dimension may simply have to be stipulated (see, e.g., Baković (1995; 1998)). Moreover, it turns out that in quite a number of cases, reranking of two constraints does not actually produce an extensionally different grammar because exactly the same candidates are predicted to be optimal under two rankings. In the case at hand (i.e., concerning the five constraints in (19) that played a role in the licensing of lexically case-marked subjects), factorial typology as such would predict not 3, but 120 different grammars. Some of the variation will be empirically innocuous, and not lead to extensionally different grammars; e.g., all rankings in which FAITH-LEX outranks FAITH-LEX$_{\text{trans}}$ will be such that the actual position of FAITH-LEX$_{\text{trans}}$ is irrelevant for the outcome.[19] Other rerankings will give rise to peculiar language types that may not be attested; e.g., reversing the ranking of *DAT, *ACC, and *NOM will predict a language in which dative is the default case, and nominative is highly marked. To avoid grammars of this type, it can be assumed that, in the present context (and notwithstanding the above remarks on unmarked case in English), the order of case markedness constraints is invariantly *DAT $\gg$ *ACC $\gg$ NOM, and the order of case faithfulness constraints is invariantly FAITH-LEX$_{\text{trans}}$ $\gg$ FAITH-LEX. This will then ideally be derivable from the internal make-up of the constraint families (e.g., they might all be derivable from more basic abstract constraints by techniques like local conjunction and harmonic alignment; also see Wunderlich (2003), Stiebels (2002) for a somewhat different, but also principled, approach).

3. Problems for OT Analyses in Syntax

Of the problems for OT analyses that have been raised in the literature, three can be singled out as both potentially troublesome and highly illuminating. First, there is the issue of complexity of competition-based grammars (with possibly infinite candidate sets). Second, a problem arises that is practically unheard of in most other syntactic approaches, viz., that of deriving instances of ineffability (or absolute ungrammaticality). And finally, accounting for syntactic optionality remains an intricate issue in OT syntax to this day. As with the evidence in support of OT analyses, I address the issues one by one.[20]

3.1 Complexity

The potential problem here is very easy to grasp. Competition adds complexity; and because of the general option of recursion in syntax, candidate sets are not finite in most analyses. From the very beginning, Prince and Smolensky had anticipated this criticism. Here is their 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)

I have nothing to add to this statement, except for the observation that if there is a problem here, OT shares the problem with other competition-based theories of syntax (e.g., minimalist approaches like those of Chomsky (1993; 1995), Collins (1994), and Bošković (1997), which rely on transderivational constraints applying to candidate derivations in large (typically infinite) reference sets – note that versions of transderivational constraints are arguably still adopted in more recent minimalist analyses, e.g., those that rely on a constraint like Merge before Move; see section 5 below).

3.2 Ineffability (Absolute Ungrammaticality)

Basically, a sentence (more generally, any LE) can only qualify as ungrammatical in OT if there is some other sentence (or LE) that blocks it by being the optimal candidate. However, sometimes it is far from obvious what this other sentence (LE) should look like. Consider illicit extraction from an adjunct island, as in the German example in (23).

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

This is a clear case of ineffability, or absolute ungrammaticality. At least five different approaches to ineffability can be distinguished in OT syntax. In what follows, I introduce them in turn, based on the problem posed by (23).

3.2.1 The Generator

A first approach relocates the problem with (23) from the H-EVAL system to GEN (see, e.g., Pesetsky (1997)). One might simply assume that GEN contains constraints like (24) that preclude a generation of outputs like (23) in the first place.

(24) ADJUNCT CONDITION:
Movement must not cross an adjunct clause.

This way, the problem of accounting for ineffability is solved outside the OT system proper.

3.2.2 Empty Outputs

A second approach relies on the assumption that each candidate set contains a candidate that leaves the input completely unrealized. This candidate is the "empty output" or "null parse": Ø (see, e.g., Ackema & Neeleman (1998)). By definition, the empty output does not violate any faithfulness constraints; in fact, the *only* constraint that it violates is *Ø in (25).

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

T_{14} shows how the empty output (here, O_3) can become optimal, and successfully block both a candidate with *wh*-movement across an adjunct island as in (23) (O_1), and a candidate that fails to carry out *wh*-movement altogether (O_2).

T_{14}: 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			*

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

3.2.3 Bad Winners

A third kind of approach to ineffability assumes that the optimal candidate cannot be interpreted by other components of grammar (phonology, semantics), or by the interfaces with these components (see Grimshaw (1994) and Müller (1997), among others). Thus, one might posit that (without the null parse present in candidate sets), a candidate that leaves the *wh*-phrase in situ throughout the derivation might be optimal, as in T_{15}.

T_{15}: Ineffability and bad winners

	ADJUNCT CONDITION	WH-CRIT
O_1: was ... [nachdem er t V]	*!	
☞ O_2: – ... [nachdem er was V]		*

However, the optimal candidate O_2 of T_{15}, which corresponds to (26), might be semantically uninterpretable as a regular question.

(26) #Fritz ist eingeschlafen [$_{CP}$ nachdem er was gelesen hat] ?
Fritz is fallen asleep after he what read has

Arguably, this approach corresponds to 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).

3.2.4 Repair

The null hypothesis might be that there is in fact an optimal repair candidate for (23); i.e., that extraction from an adjunct island is blocked in favour of a repair strategy that exhibits properties which are not otherwise tolerated in the language. The sentences in (27-ab) are two potential repair candidates for (23) in German.

(27) 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

In (27-a), the function of the incriminating item is changed: The *wh*-pronoun *was* ('what') is reinterpreted as an indefinite pronoun *was* ('something', a short form of *etwas*). In (27-b), the form of the incriminating item is changed: Instead of a moved *wh*-pronoun, there is a *wh*-PP (*bei was*, roughly 'with respect to what') in the interrogative SpecC position outside the island, together with a resumptive pronoun *es* ('it') within it. Let us focus mainly on the first alternative here (the theoretical issues raised by the other repair approach are more or less identical). Suppose that there is a faithfulness constraint like (28).

(28) IDENT([wh]):
A feature [+wh] in the input must not be changed to [–wh] in the output.

(27-a) would then, as a last resort, violate IDENT([wh]) by reinterpreting the *wh*-pronoun as an indefinite pronoun, as shown in T_{16}.

T_{16}: Ineffability and repair

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

However, unfortunately the repair approach to ineffability does not work (at least not in the way just sketched). The problem is that the "repair" strategy is also available outside island contexts, e.g., with successive-cyclic *wh*-movement from a declarative clause embedded by a bridge verb; see (29-a) (where *wh*-movement is possible) vs. (29-b) (where the indefinite interpretation of *was* is also available). Similarly, the *wh*-PP/resumptive pronoun strategy is possible without an island being present; see (29-c).

(29) a. Was glaubt Fritz [$_{CP}$ dass er t lesen sollte] ?
what thinks Fritz that he read should
b. Fritz glaubt [$_{CP}$ dass er was lesen sollte]
Fritz thinks that he what (= something) read should
c. Von was glaubt Fritz [$_{CP}$ dass er es lesen sollte] ?
of what thinks Fritz that he it read should

As shown in T_{17} for the indefinite interpretation of *was*, considering (27-a) or (27-b) to be repair options does not make the right predictions: O_3 is blocked by O_1 in T_{17}, but it must be an optimal candidate.

To conclude, clauses with *wh*-indefinites 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 clause; see Koster (1986), Cinque (1990), Sternefeld (1991),

T_{17}: A wrong prediction

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

Barbiers (2002), Gallego (2007). However, there is a more fine-grained version of the repair approach that allows one to both have the cake and eat it. It is based on neutralization.

3.2.5 Neutralization

The fifth and final approach to ineffability to be discussed here centers around the concept of input neutralization. The main premise is that there can be 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) may converge on a single optimal candidate. Approaches of this type have been developed by Legendre et al. (1995), Legendre, Smolensky & Wilson (1998), Keer & Baković (2004), Keer & Baković (2001), Vogel (2001), and Wilson (2001), among others. For the case at hand, consider first a competition with a transparent context and a *wh*-item that bears the feature [+wh] in the input, as in T_{18}.

T_{18}: Transparent contexts without neutralization: 'was$_{[+wh]}$' in the input

Input: was$_{[+wh]}$, ...	ADJUNCT CONDITION	WH-CRIT	IDENT([wh])
☞ O_1: was$_{[+wh]}$ ... [dass er t V]			
O_2: – ... [– dass er was$_{[+wh]}$ V]		*!	
O_3: – ... [dass er was$_{[-wh]}$ V]			*!
O_4: was$_{[-wh]}$ ... [dass er t V]			*!

O_1, which leaves the +-value of the *wh*-feature intact and moves the *wh*-phrase to SpecC, emerges as optimal. O_3, which changes the value from + to –, fatally violates IDENT([WH]). Consider next a competition with a transparent context where the *wh*-item is an indefinite (i.e., [–wh]) in the input, as in T_{19}. Again, the faithful candidate wins – changing the feature value does not lead to an improved behaviour with respect to higher-ranked constraints. In both competitions, there is a further output O_4 that applies movement of a [–wh] phrase to SpecC. As it stands, O_4 has the same violation profile as O_3 with respect to the three constraints given here. However, it is suboptimal because it violates ECONOMY (see (8)) in addition without contributing to a better behaviour with respect to any other constraint. As a matter of fact, O_4 is not expected to be grammatical under any reranking of the four pertinent constraints; the technical expression for such a state of affairs is that O_4 is *harmonically bounded* by O_3 (see Prince & Smolensky (2004)).[21] Harmonically bounded candidates are predicted to be universally unavailable.

This solves the problem with the pure repair approach: Both strategies (*wh*-movement, *wh*-indefinite) can survive in transparent contexts because they go back to minimally different inputs, and thus different competitions. However, in opaque contexts where a locality constraint like the ADJUNCT CONDITION becomes active and distinguishes between candidates, neutralization takes place. Under present assumptions, (30-a) and (30-b) compete with one another, both with a [+wh]-specification in the input and with a [–wh]-specification in the input.

T_{19}: Transparent contexts without neutralization: 'was$_{[-wh]}$' in the input

Input: was$_{[-wh]}$, ...	ADJUNCT CONDITION	WH-CRIT	IDENT([wh])
O_1: was$_{[+wh]}$ ... [dass er t V]			*!
O_2: – ... [– dass er was$_{[+wh]}$ V]		*!	*
☞ O_3: – ... [dass er was$_{[-wh]}$ V]			
O_4: was$_{[-wh]}$ ... [dass er t V]			

(30) 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

If there is a [+wh]-specification in the input, as in T_{20}, the two higher-ranked constraints ADJUNCT CONDITION and WH-CRIT eliminate the faithful candidates O_1, O_2 (hence, (30-a)), and the unfaithful candidate O_3 becomes optimal (i.e., (30-b); note that the harmonically bounded output O_4 is ignored here and in the following tableau).

T_{20}: Island contexts with neutralization, unfaithful: 'was$_{[+wh]}$' in the input

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

However, O_3 (= (30-b)) also emerges as optimal in the minimally different context where the input specification is [–wh] to begin with; see T_{21}.

T_{21}: Island contexts with neutralization, faithful: 'was$_{[-wh]}$' in the input

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

Thus, the difference in the input between T_{20} and T_{21} is neutralized in the output.[22] As before, the question arises of whether a sentence like (30-b) must then be assumed to have two (or, perhaps, many more) possible sources; and as before, the standard answer given in OT is that input optimization may compare the two optimal candidates, and filter out one of them. Much more will eventually have to be said about absolute ungrammaticality in OT syntax, but I will leave it at that.[23]

3.3 Optionality

In general, only one candidate should be optimal in a given candidate set.[24] Thus, the question arises of what to do about situations where it looks as though several outputs can co-exist as optimal. Let us go through a number of potentially relevant phenomena. Consider first complementizer deletion in English. The example pair in (31-ab) shows that a complementizer *that* can be left out in declarative object clauses (at least under certain circumstances, which include the right choice of matrix predicate).

(31) a. I think – John will leave
b. I think that John will leave

Next, (32-ab) show that German has an alternation between so-called partial *wh*-movement (where a *wh*-phrase moves to an intermediate declarative SpecC position, and the final SpecC position of the interrogative clause is filled by a scope marker *was* ['what']), and standard long-distance *wh*-movement. At least in certain varieties of German (32-a) and (32-b) are both perfectly legitimate, unmarked ways of producing exactly the same kind of question.

(32) a. Wen glaubst du [$_{\text{CP}}$ dass man t einladen sollte] ?
whom think you that one invite should
b. Was glaubst du [$_{\text{CP}}$ wen man t einladen sollte] ?
what think you whom one invite should

A third example involves *wh*-movement in French. In certain contexts (viz., root questions), and with certain *wh*-items (viz., arguments), *wh*-movement is optional in this language; cf. (33-ab).

(33) a. Qui as-tu vu t ?
b. – Tu as vu qui ?

Extraposition is also an operation that often applies optionally; compare the example with an in-situ relative clause modifying the the head of the subject in German in (34-a) with the minimally different example in (34-b), where the relative clause has undergone extraposition.

(34) 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

It is entirely unproblematic to continue this list with more examples from many more languages. For present purposes, it may do to give one more relevant example: In free word order languages, there are many contexts where various orders can co-exist; this is shown for the optional permutation of subject and object in Korean in (35) (from Choi (1999, 172)).

(35) a. Swuni-nun Inho-lul manna-ss-e
Swuni$_{top}$ Inho$_{acc}$ meet-PAST-DECL
b. Inho-lul Swuni-nun manna-ss-e
Inho$_{acc}$ Swuni$_{top}$ meet-PAST-DECL
'Swuni met Inho.'

Various kinds of approaches to optionality can be distinguished in OT syntax. The taxonomy of analysis types in (36) is based on Müller (2003b).

(36) *Analyses of optionality of two candidates C_i, C_j*:
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 violation profile.

c. *Ties*:
C_i, C_j differ only with respect to 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*

I cannot go through all these different strategies for dealing with optionality in detail here (see Müller (2003b) for a comprehensive exposition covering all but the stochastic analyses). In the following subsections, I will confine myself to pseudo-optionality and neutralization, true optionality, two kinds of ties, and finally stochastic optimality theory.

3.3.1 Pseudo-Optionality and Neutralization

The basic assumption underlying pseudo-optionality analyses is that instances of optionality are only apparent: The two optimal candidates are winners of two different competitions. To achieve this, candidate sets must be defined in such a way that there is little competition. Suppose, for instance, that for the example pairs in (33), (34), and (35), a movement-inducing feature can optionally be present in the input; if it is present, it triggers a movement operation that creates the different word order (*wh*-movement, extraposition, and scrambling, respectively). (For (31) and (32), invoking different lexical material may suffice to generate two separate competitions.) Assuming that candidate sets are (at least partly) defined by input identity, the candidate with movement will be the optimal output of the candidate set that has the relevant feature in the input, and the candidate without movement will be the optimal output of the candidate set that lacks this feature in the input. This gives rise to an obvious problem, though: If there is not much competition, this weakens the overall theory and increases the problem of accounting for ineffability. To see this, consider, e.g., the case of partial *wh*-movement in German. Whereas the data in (32) show that partial and long-distance *wh*-movement can co-exist, the examples in (37-ab) show that the distribution of the two construction types is not fully identical. If there is negation in the matrix clause, partial *wh*-movement ceases to be possible, while long-distance *wh*-movement is much more acceptable for many speakers. From an optimality-theoretic perspective, this strongly suggests that partial *wh*-movement and long-distance *wh*-movement do belong to one and the same competition after all, with the latter option blocking the former one in certain island contexts. Here and henceforth, I will refer to instances of optionality that breaks down in certain contexts as *alternations*.

(37) a. ?Wen glaubst du nicht [$_{\text{CP}}$ dass man t einladen sollte] ?
whom think you not that one invite should

b. *Was glaubst du nicht [$_{\text{CP}}$ wen man t einladen sollte] ?
what think you not whom one invite should

The same problem shows up with the co-existence of *wh*-movement and *wh*-in situ in French (cf. (33)). As shown in (38-ab), in embedded clauses only the former strategy is available (the same asymmetry arises in (embedded or matrix) contexts where the *wh*-phrase is an adjunct; see

Aoun et al. (1981)). Again, this alternation suggests that the two construction types are part of the same competition after all, which excludes a pseudo-optionality approach.

(38) 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

Of course, there is one potential way out of this dilemma (see Legendre et al. (1995), Keer & Baković (2004), Keer & Baković (2001)): As with ineffability, one can adopt a neutralization approach. On this view, each of the two optional variants in alternations like those in (32) and (33) is a faithful winning candidate of one competition, and a fatally unfaithful losing candidate of the other competition. In contexts like those in (37) and (38), where optionality breaks down, the two separate competitions converge on a single output that is faithful (with respect to the relevant input property defining the candidate set) in one case and unfaitful (with respect to this property) in the other case – the input difference is neturalized in the output. Basically, in this kind of approach, there is no relevant difference between the ineffability problem and the optionality problem: The relation between, say, (29-a) and (29-b) emerges as an instance of optionality in the same way that the relation between, say, (31-a) and (31-b) does. As with ineffability, the neutralization approach to optionality presupposes that the generator (GEN) is sufficiently powerful to create strongly unfaithful candidates; and as before, the issue of input optimization arises.

3.3.2 True Optionality

Here the assumption is that two (or more) candidates can in fact have the same (optimal) violation profile; given an appropriate definition of optimality that does not presuppose that there is a single, unique winner (see (1)), they can then both be optimal. Approaches of this type have been pursued by Grimshaw (1994) and Vikner (2001a), among others. For instance, Grimshaw (1994) suggests deriving the optionality of English complementizer drop in (31) from an identical violation profile of the two candidates. This approach solves the alternation problem (i.e., it straightforwardly captures a selective breakdown of optionality.) However, adopting such an approach proves very hard (or indeed impossible) in practice: Because of faithfulness constraints for lexical items and features, and because of symmetrical markedness constraints that invariably incur violations for all pieces of structure present in a candidate (like ALIGN-X-LEFT being accompanied by ALIGN-X-RIGHT), there will *always* be constraints where two non-surface-identical candidates differ. Given a low ranking of these constraints, they may not be active in the sense that they can determine an optimal output's properties; but they will suffice to create a distinct profile profile of two candidates. Along these lines, analyses involving true optionality have been widely criticized (see, e.g., Keer & Baković (2001), Grimshaw (1999), and, indeed, Grimshaw (1997, 411) already); and they do not seem to be regularly pursued anymore.

3.3.3 Ties

The central idea behind ties is that two (or more) constraints are equally important, i.e., "tied." If two candidates differ only with respect to a tie of constraints, they can both be optimal, even if their violation profiles are not completely identical. In what follows, I will render a tie of two constraints A and B as "A∘B". Various concepts of tie have been proposed in the literature. A basic distinction is between what can be called "global tie" and what can be called "local tie". Global ties are abbreviations for multiple constraint rankings; local ties are essentially special constraint types. I am aware of at least five distinct concepts of tie (two of them global, three local) that can be shown to be both conceptually different, and empirically incompatible, viz.: ordered global ties (see, e.g., Sells et al. (1996), Ackema & Neeleman (1998), Schmid (2001; 2005); and Prince & Smolensky (2004) for the original concept); disjunctive global ties (see Müller (1999)); ordered local ties (see, e.g., Pesetsky (1997; 1998)); conjunctive local ties (see, e.g., Prince & Smolensky (1993; 2004), Legendre et al. (1995), Legendre, Smolensky & Wilson (1998), Müller (1997), Tesar (1998), and Legendre (2001)); and disjunctive local ties (see Broihier (1995)). Still, in the abstract tableau T_{22}, with the ranking A $\gg$ B$\circ$C $\gg$ D, all these concepts of constraint tie turn out to make identical predictions: O_1 and O_2 are both optimal, whereas O_3 and O_4 are blocked as suboptimal.

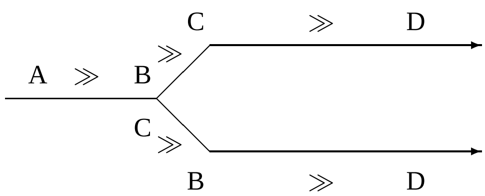
T_{22}: Constraint tie: B$\circ$C

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

As noted, the basic distinction is between *global* ties and *local* ties. Global ties can be viewed as abbreviations for the simultaneous presence of different constraint rankings in a language; they thus essentially correspond to the multiple-grammar approach to (temporary) optionality in syntax as it has been proposed in studies in historical linguistics (see Kroch (2001) and references cited there). In contrast, local ties can be viewed as special constraint types.[25] The most widespread concepts of tie in the literature are arguably *ordered global ties* and *conjunctive local ties*.

Ordered global ties essentially work like this: A constraint ranking that exhibits an ordered global tie of two constraints B$\circ$C is underspecified; it is an abbreviation that encodes the simultaneous presence of two hierarchies that exhibit the rankings B $\gg$ C and C $\gg$ B, respectively. A candidate is grammatical if it qualifies as optimal under one of the possible resolutions of the tie. With global ties (ordered or not), the optimal outputs of one candidate set may have a different violation profile below the tie. This is illustrated for ordered global ties in (39): An output is optimal if it is optimal with respect to at least one of the two rankings (A $\gg$ B $\gg$ C $\gg$ D or A $\gg$ C $\gg$ B $\gg$ D).

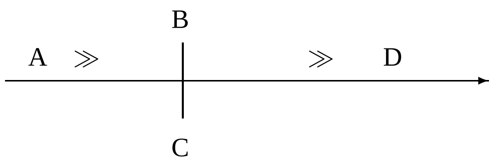
(39) *Diagram of an ordered global tie* B$\circ$C



As a concrete example, consider the optionality of *wh*-movement and scope marking in German in (32), and its breakdown (alternation) in (37). Each movement step may be assumed to violate ECONOMY (see above), and scope marker insertion may be assumed to violate a DEP constraint blocking expletive insertion (like FULL-INT(erpretation) in Grimshaw (1997)); this presupposes that expletives are never part of an initial input (see Hornstein (2001) for a comparable assumption concerning numerations in minimalist syntax). Assuming these two constraints to be part of an ordered global tie, optionality may emerge. However, in (say) negative island contexts, the partial movement output may violate a locality constraint ranked above the tie more often than the long-distance movement candidate does (given the approach in Müller (1997), the reason might be that the former candidate incurs two violations of the locality constraint, one in syntax and one in logical form, whereas the latter candidate incurs only one violation, in syntax).

In contrast, conjunctive local ties are not abbreviations for the simultaneous presence of more than one ranking. They basically act like ordinary constraints, and there is no resolution of the tie into suborders. Rather, the two constraints are merged into a single constraint that is interpreted via logical conjunction: A candidate violates a conjunctive local tie if it violates a constraint that is part of this tie, and multiple violations add up. With local ties, two outputs can only be optimal if they have an identical behaviour with respect to constraints that are ranked below the tie. Otherwise, a breakdown of optionality is predicted. (Thus, a somewhat less severe version of the central problem for true optionality approaches persists.) The working of a conjunctive local tie is illustrated in (40).

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



Taking again the optionality (and alternation) of *wh*-movement in German as an example, a conjunctive local tie approach might rely on the assumption that ECONOMY and DEP (FULL-INT) are locally tied (like B and C in (40)). The locality constraint underlying negative islands might then be ranked either above or below the tie.

Whereas the data in (32) and (37) do not differentiate between these two (and other) approaches to ties per se, an argument is brought forward in Müller (1997) to the effect that more complex data favour the conjunctive local tie approach: In cases where there are *two* intervening SpecC domains between the base position and the ultimate SpecC[+wh] target position, three outputs can emerge as optimal; see (41).

(41) a. Wann$_1$ meinst du [$_{CP}$ t_1'' dass sie gesagt hat [$_{CP}$ t_1' dass sie t_1 kommen würde]] ?
when think you that she said has that she come would

b. Was$_1$ meinst du [$_{CP}$ wann$_1$ (dass) sie gesagt hat [$_{CP}$ t_1' dass sie t_1 kommen würde]] ?
[+wh] think you when that she said has that she come would

c. Was$_1$ meinst du [$_{CP}$ was$_1$ sie gesagt hat [$_{CP}$ wann$_1$ (dass) sie t_1 kommen würde]] ?
[+wh] think you [+wh] she said has when that she come would

(41-a) incurs three violations of ECONOMY (and no violations of DEP); (41-b) incurs two violations of ECONOMY and one violation of DEP; and (41-c) incurs one violation of ECONOMY and two violations of DEP. A conjunctive global tie permits all three outputs to be optimal (because they all incur three violations of the single merged constraint ECONOMY$\circ$DEP); an ordered global tie ceteris paribus makes the wrong prediction that (41-b) should be suboptimal (because the ranking ECONOMY $\gg$ DEP will favour maximal use of scope markers, as in (41-c), and the reverse ranking DEP $\gg$ ECONOMY will favour maximal use of movement, as in (41-a)). That said, there is also conflicting evidence from other empirical domains which would seem to suggest that (ordered) global ties form the superior concept. As present, it is an open question which version of tie (if any) is to be preferred in OT syntax. (Of course, several concepts of tie may also co-exist.)

Interestingly, there is a version of the concept of ordered global tie which has received some attention in more recent years even though the close connection is usually not made explicit: The concept shows up in stochastic approaches to optimality theory.

3.3.4 Stochastic Optimality Theory

Stochastic optimality-theoretic analyses of phonological phenomena have been developed in Anttila (1997), Boersma & Hayes (2001), and Hayes (2001). Syntactic applications include Aissen (2003a;b) (on optionality with differential object marking and with DP-internal possessor placement, respectively), Bresnan, Dingare & Manning (2001) (on optionality in passive formation), and Bresnan, Deo & Sharma (2007) (on types of inflection of the verb *be*, including negation, in varieties of English). The basic observation is that quite often, the constructions that co-exist as optional and may participate in an alternation (with one selectively blocking the other in certain contexts) are not equally *frequent*, or equally *unmarked* (or, for that matter, equally *well formed* – i.e., they may exhibit different degrees of acceptability). For instance, the positioning of possessors in English DPs, while often (though not always) optional, also often illustrates clear preferences for one or the other option that can be detected by checking relative frequency in corpora, and also by consulting native speaker intuitions. Preferences are indicated by $>$ in (42) (and ?*/* signals cases of (near-) complete ungrammaticality of one output, i.e., the breakdown of optionality).

(42) 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

Evidently, placement of a possessor on animacy and definiteness scales (which are independently motivated, see Hale (1972), Silverstein (1976)) plays an important role in their DP-internal positioning. Aissen (2003b) sets out to derive the pattern in (42) – both the preferences for positioning in cases of optionality, and the categorical unavailability of some of the options. To this end, she first assumes that the underlying animacy and definiteness hierarchies can be used as primitives to generate sequences of constraints with a fixed internal order (sometimes called "subhierarchies"), via a process of *harmonic alignment* of scales (see Prince & Smolensky (2004), and Aissen (1999) for an influential application in syntax). Harmonic alignment is defined as in (43) (cf. Prince & Smolensky (2004, 161)).

(43) *Harmonic Alignment*:
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 $\succ$ X/b $\succ$ ... $\succ$ X/z
b. H_Y: Y/z $\succ$ ... $\succ$ Y/b $\succ$ Y/a

The *constraint alignment* is the pair of constraint hierarchies C_X, C_Y:

a. C_X: *X/z $\gg$ ... $\gg$ *X/b $\gg$ *X/a
b. C_Y: *Y/a $\gg$ *Y/b $\gg$ ... $\gg$ *Y/z

Thus, given an animacy scale [human $>$ animate $>$inanimate] and a definiteness scale [pronoun $>$ name $>$ definite DP $>$ indefinite DP], harmonic alignment of these scales with the binary scale [SpecN $>$ CompN] will automatically produce the four constraint subhierarchies in (44). (Here, SpecN and CompN are abbreviations for prenominal placement and post-nominal placement of the possessor in a DP, respectively, with the former realized by genitive *'s* and the latter by an *of*-PP.) Note that the order within a subhierarchy is universally fixed. This derives varying degrees of markedness of certain options. For instance, from (44-a(i)) it follows that a pre-nominal inanimate possessor will always violate a higher-ranked constraint in the subhierarchy (and therefore qualify as more marked) than a pre-nominal animate possessor.

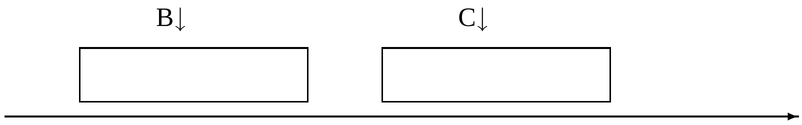
(44) a. (i) *SpecN/inanimate $\gg$ *SpecN/animate $\gg$ *SpecN/human
(ii) *CompN/human $\gg$ *CompN/animate $\gg$ *CompN/inanimate
b. (i) *SpecN/indef $\gg$ *SpecN/def $\gg$ *SpecN/name $\gg$ *SpecN/pron
(ii) *CompN/pron $\gg$ *CompN/name $\gg$ *CompN/def $\gg$ *CompN/indef

Given that DP-internal possessors must be placed either in a pre-nominal or in a postnominal position, the constraints in (44-a(i)) and the constraints (44-a(ii)) impose conflicting requirements on outputs (e.g., *SpecN/inanimate requires inanimate possessors to show up post-nominally, whereas *CompN/inanimate requires inanimate possessors to show up pre-nominally); as do the constraints in (44-b(i)) and (44-b(ii)). In a standard OT system without ties, interleaving of the two hierarchies in (44-a(i)) and (44-a(ii)), and the two hierarchies in (44-b(i)) and (44-b(ii)) will determine a single optimal output of each input. As regards (44-a), if *CompN/inan $\gg$ *SpecN/inan, all possessors will be realized pre-nominally; if *SpecN/human $\gg$ *CompN/human, all possessors will be realized post-nominally; otherwise, mixed patterns will result which, however, respect implicational generalizations (e.g., if an animate (non-human) possessor is realized pre-nominally, a human possessor also has to be realized pre-nominally; or if an animate (non-human) possessor is realized post-nominally, an inanimate possessor also has to be realized post-nominally). Similar conclusions apply in the case of (44-b). Furthermore, by *locally conjoining* the members of two similar subhierarchies (e.g., (44-a(i)) and (44-b(i)), both precluding pre-nominal possessor placement) in an order-preserving way, a two-dimensional picture arises: In the case at hand, the highest-ranked constraint then is the one blocking a pre-nominal placement of inanimate indefinite possessors (*SpecN/inanimate & *SpecN/indef), the lowest-ranked constraint bans a pre-nominal placement of human pronominal possessors (*SpecN/human & *SpecN/pron); and whereas there is no fixed ranking between, say *SpecN/inanimate & *SpecN/-def and *SpecN/animate & SpecN/indef, the ranking of *SpecN/inanimate & *SpecN/def and *SpecN/inanimate & *SpecN/name is fixed again (as is the ranking between *SpecN/animate & SpecN/indef and *SpecN/animate & SpecN/def).

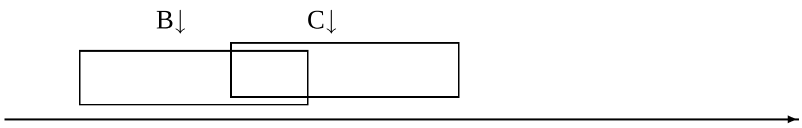
In view of the (partial) optionality visible in (42), this approach does not yet seem correct. In principle, one could now assume ties to derive optionality where it occurs. However, Aissen (2003b) does not pursue this approach because it does not offer a way to integrate the finding that in the cases of optionality in (42), one of the two options is typically more frequent (and less marked) than the other one. This state of affairs can be derived by adopting a stochastic OT approach.

The basic idea of stochastic OT is that 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. A categorical vs. overlapping application domains of constraints is illustrated in (45-ab), where the boxes signal the possible application domains of constraints on an abstract continuum (→) from “more strict” to “less strict” (interpreted from left to right).

(45) a. *Categorical order of application domains of constraints*:

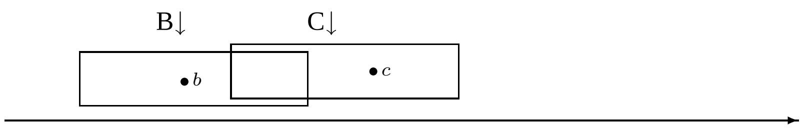


b. *Overlapping order of application domains of constraints*:

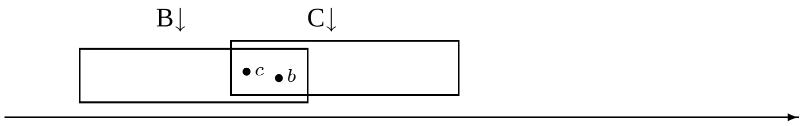


Here is how the approach works technically: A candidate is evaluated at an *evaluation time*; it is well formed if it is optimal at that point. For an evaluation, an arbitrary point is chosen in the application domain of a constraint. A constraint B is ranked higher than another constraint C at a given 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 above 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. So far, this is basically identical to the concept of ordered global tie. However, in addition to permitting an account of optionality, the new system also captures *preferences*: The choice of an 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-à-vis C points at a given evaluation time is, the more the construction favoured by B is going to be preferred over the construction favoured by C; similarly, the more frequent the construction favoured by B will be in corpora. This is illustrated in (46).

(46) a. *Typical result*: B $\gg$ C



b. *Rare result*: C $\gg$ B



Thus, by assuming that the constraints determining possessor placement may have both non-overlapping and overlapping (but typically non-identical) application domains in English, Aissen (2003b) succeeds in deriving both categorical ungrammaticality of some options (the composite constraints *SpecN/indef & *SpecN/inanimate and *CompN/pron & *CompN/hum properly outrank their respective antagonists), and preferences among the two basically optional placement strategies (e.g., in (42-a), the application domains of *SpecN/inanimate & *SpecN/def and *CompN/inanimate & *CompN/def overlap, with the likelihood of choosing a higher evaluation point in the former constraint's domain being greater than the likelihood of choosing a higher evaluation point in the latter constraint's domain).

More generally, since stochastic OT can be viewed as a way to assign preferences to options permitted by globally tied constraints (conceived of as constraints with overlapping application domains), it should in principle be possible to transfer all analyses that rely on (ordered) global ties to stochastic OT analyses; and indeed, it has often been noted for cases like those in (31)–(35) that one of the two options tends to be less marked than the other one (with markedness degrees subject to micro-variation, possibly even idiolectal variation).

4. Optimization Domains

The previous two sections have addressed syntactic evidence that supports an optimality-theoretic approach, and syntactic evidence that may turn out to be problematic. When discussing these issues, I have presented each syntactic analysis in some actual framework. Typically, this has been the one in which it was presented in the original literature; and typically, this has been a version of the Principles and Parameters (P&P) approach (see Chomsky (1981), Chomsky & Lasnik (1993)). However, OT is a theory of constraint interaction, not a theory of the basic building blocks that create (or license) LEs (sentences, in the case at hand) as such. So, while it may be true that much of the groundbreaking work in OT syntax has assumed a P&P perspective on syntactic candidates and the makeup of syntactic constraints of both the GEN and the H-EVAL components (see, e.g., Grimshaw (1997), Pesetsky (1998), and Legendre, Smolensky & Wilson (1998)), there is no intrinsic relation between the P&P approach and OT. Indeed, it would seem that most syntactic theories could be enriched by an OT component; and whereas theories like HPSG or TAG seem to have largely withstood the impact of OT, LFG in particulary seems to have embraced OT, at least for a while (see Choi (1999), Sells (2001a;b), Bresnan (2001), and Kuhn (2001), among many others). Against this background, one may ask whether optimization might also be compatible with minimalist approaches (see Chomsky (1993; 2001; 2008) and much related work). In this section, I will address the issue on the basis of the related issue of *optimization domains*.

4.1 Background

A fundamental question is whether optimization of a LE applies only once (so-called *harmonic parallelism*) or more than once (*harmonic serialism*). To some extent (but see below), this distinction also manifests itself in the similar distinction between a *representational* and a *deriva-*

tional organization of grammar. Whereas in classical rule-based generative phonology the concept of ordered application of rules is crucial (giving rise to feeding, bleeding, and opacity effects in the guise of counter-feeding and counter-bleeding), OT phonology can for the most part do without derivations (with potential problems arising in the area of opacity, though), and thus qualifies as an instance of harmonic parallelism. In fact, it still seems to be a widespread assumption (particularly, but not exclusively, among those who work outside OT) that OT is inherently representational, and characterized by harmonic parallelism. However, this assessment is most certainly incorrect, as the following quote from Prince & Smolensky (1993; 2004) makes clear:

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

As a matter of fact, having first addressed the issue in McCarthy (2000), John McCarthy has recently come to embrace an approach to OT phonology that relies on harmonic serialism; see McCarthy (2008; 2010) and much related recent work (though, crucially, not McCarthy (2007)). The same goes for syntax: There is no deep reason why OT syntax should have to be strictly representational, and qualify as an instance of harmonic parallelism. The following quote makes it clear that there is no fundamental obstacle to reconciling OT with the derivational approach to syntax envisaged in the minimalist program.

> 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)

What is more, Legendre, Smolensky & Wilson (1998) point out that there are actually two ways to reconcile a derivational approach to syntax with OT – either via standard, parallel optimization of full derivations (also cf. the role of candidate chains and so-called *rLUMSeq*s in McCarthy (2007)), or via serial optimization. In the latter case, another issue becomes relevant: In classic transformational grammar (e.g., Chomsky (1965)), syntactic transformations applying to the output of the base component effect derivational steps where the input and the output have roughly the same size, exactly as in phonology. For instance, a *wh*-movement transformation may reorder a *wh*-phrase with respect to the rest of the clause, but the transformation does not per se create additional structure (many transformations are structure-preserving). Things are different in the minimalist program, where the operations of the "base" component and of the "transformational" component are systematically interspersed; syntactic structures start with two lexical items and grow throughout the derivation by iterated application of (external or internal) Merge. From a serial OT perspective, this implies that iterated optimization in syntax cannot apply to objects of (roughly) the same size (as it still is the case with serial optimization in phonology, which involves no structure-building) – the optimal output of one optimization procedure will have to be smaller than the optimal output of the next optimization procedure (assuming there is more than one). This in turn means that we have to introduce a second fundamental difference in optimization options: Optimization may be *parallel* (i.e., apply once) or *serial* (i.e., apply more than once); and optimization may be *global* (applying to the full LE) or *local* (applying also to smaller domains). Whereas serial optimization in phonology is typically global (phonology *restricts* the shape of LEs – words or morphemes – but it does not *create* them),[26] serial optimization in minimalist syntax must be local (the LEs created by Merge grow successively).[27]

Given that syntactic optimization can be both serial and local, the question arises of how the local domain that optimization applies to is defined. (47) lists a number of options.

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

As noted, the standard assumption OT syntax is that the whole sentence is subject to a single, parallel global optimization procedure (Grimshaw (1997), Pesetsky (1998), and 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)). In contrast, serial global optimization of whole sentences is proposed in Wilson (2001) and Heck (1998; 2001). Finally, serial *local* optimization in syntax is closely related to developments in the minimalist program. Conceptually, there are trade-offs. An argument for small optimization domains might be this: The smaller the optimization domain is, the more the complexity of the overall system is reduced (i.e., there is a reduction of the size of candidate sets). Furthermore (as noted by a reviewer), smaller optimization domains are also likely to require *simpler* constraints than larger domains. On the other hand, an argument for larger optimization domains might be that the larger the optimization domain is, the less often optimization procedures have to be carried out. Assuming that iterated optimization in small domains is ultimately cheaper than single optimization of extremely large domains, one might perhaps make a case that local optimization is conceptually preferable. It is also worth noting that there is evidence outside of language for optimization of small domains (see, e.g., Gigerenzer et al. (1999) on "fast and frugal" *decision-making*, which relies on the availability of very little information).

However, ultimately empirical arguments are needed to decide whether optimization domains should be viewed as small or large (possibly global). Such arguments are of the following type: If the ranked constraints have access to more/less structure, a wrong winner is predicted, ceteris paribus. All the options in (47) have been pursued, and arguments for the specific notion of optimization domain chosen in each case have typically taken this form. Let me list a few examples: (i) The *minimal clause* is identified as the optimization domain for syntax in Ackema & Neeleman's (1998) study of *wh*-movement in Czech, and in an analysis of extraction from verb-second clauses in German that I develop in Müller (2003a). (ii) Arguments for the *phase* as the optimization domain are presented in Fanselow & Ćavar's (2001) investigation of MeN-deletion in Malay, and in Müller (2000a; 2002), studies that deal with R-pronouns in German. (iii) Next, the *phrase* (XP) is argued to be the syntactic optimization domain in the cross-linguistic study of reflexivization (including long-distance reflexivization) documented in Fischer (2004; 2006); similarly, the approach to secondary remnant movement in Müller (2000c) and the approach to *wh*-movement, superiority, quantifier raising, and sluicing developed in Heck & Müller (2000; 2003) make crucial reference to the phrase as the domain of optimization. (iv) Finally, empirical arguments for optimization of individual *derivational steps* are given in Heck & Müller (2007; 2010) (based on gender agreement with dative possessors in German DPs and expletives in German verb-second clauses), in Müller (2009) (based on ergative vs. accusative argument encoding patterns), in Lahne (2008; 2009) (based on the incompatibility of SVO order and ergative systems of argument encoding), and in Georgi (2009) (based on global case splits in Tauya). In what follows, I will address two arguments for serial, local optimization in a bit more detail.

4.2 Clauses as Optimization Domains

Ackema & Neeleman (1998) are concerned with multiple question formation in typologically different systems. Based on earlier work by a number of authors, they identify two general possibilities for the analysis of multiple *wh*-movement as it can be found in Slavic languages, viz. *wh*-cluster formation (all *wh*-phrases are adjoined to one, which then undergoes movement) on the one hand and multiple separate movement on the other. Czech can be shown to follow the latter strategy. This must also be the case when multiple *wh*-movement applies long-distance, as in (48) (note that the sequence of fronted *wh*-phrases can be interrupted by parentheticals and the like, as in (48), which is incompatible with *wh*-cluster formation).

(48) [$_{VP}$ 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 analysis is based on the three constraints in (49). Both Q-MARK and Q-SCOPE trigger *wh*-movement. Q-MARK requires movement to a designated specifier of a functional head, whereas Q-SCOPE can be satisfied via movement to a local VP-adjoined position. STAY is a gradient version of ECONOMY that minimizes the length of movement paths (with no movement at all emerging as the ideal option).

(49) 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.

The ranking Q-SCOPE $\gg$ STAY $\gg$ Q-MARK in Czech predicts multiple separate *wh*-movement to VP-adjunction sites in matrix questions, as opposed to local *wh*-cluster formation and movement of the *wh*-cluster to a specifier position (as it is predicted for Bulgarian-type languages where Q-MARK is also ranked high). The reason is that separate *wh*-movements involve shorter movement paths if the target position is in the same clause. However, if the ultimate target is outside the minimal clause, and long-distance *wh*-movement is called for (as in (48)), the analysis requires local optimization in order to predict the right outcome. Here is what Ackema and Neeleman have to say in the footnote where they tackle the potential problem:

> 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)

Thus, optimization first applies to the embedded CP; see T_{23}. O_1 (with separate movement) is optimal because O_3 (with *wh*-cluster formation and movement to a clause-initial specifier) incurs fatal violations of STAY.

T_{23}: 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$_2$]]	*!		

The optimal output O_1 is subsequently taken as the sole input for optimization of the matrix CP; all competing candidates are descendants of O_1 (this is signalled by labelling them O_{11}, O_{12}, etc.). Again, *wh*-movement has to apply, and again, separate movement emerges as optimal because it involves shorter movement paths and thereby minimizes STAY violations. (Q-MARK is now violated, but this is harmless given the ranking.) This is shown in T_{24}.

T_{24}: 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]]]]]]		*** *******!**	

If there had not been serial, local optimization of CPs, but rather parallel, global optimization of the whole sentence as it is standardly assumed, a wrong winner would have been predicted. In the words of Ackema and Neeleman: "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." This is illustrated in T_{25}. The output that would wrongly predicted to be optimal (due to fewer nodes crossed in the course of movement overall) is marked by ☛.

T_{25}: 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]]]]		********	

The underlying logic is this: Two short separate movements may be better than a short movement (creating a *wh*-cluster) plus a longer movement of the cluster. E.g., 2+2=4 nodes may be crossed in the first case, and 1+5=6 nodes in the second. However, two medium-sized separate movements can still be worse than a short movement (creating a cluster) and a very long movement. E.g., 7+7=14 nodes may be crossed in the first case, and 1+10=11 nodes in the second.[28]

4.3 Derivational Steps as Optimization Domains

Heck & Müller (2007; 2010) suggest that the minimalist program and optimality theory can be fruitfully combined.[29] A basic assumption is that syntactic structure is built up derivationally and is subject to repeated local optimization: Structure-building and optimization apply in a cyclic interleaving fashion. More specifically, based on a given input, the minimalist operations Merge, Move, and Agree create various output candidates α_1,...,α_n: the candidate set M. M is subject to optimization. The optimal output α_i then serves as the input for the next cycle, and so on, until the final root node has been reached and the sentence is complete. Thus, in this approach every derivational step is subject to optimization.[30] The empirical arguments for extremely local optimization domains presented in Heck & Müller (2007; 2010) take the following form: 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. The conflict can be resolved by ranking the requirements: The highest-ranked requirement is satisfied immediately; lower-ranked ones are not yet satisfied 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. However, 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. This is the wrong result because sentences would be predicted to be well formed that aren't.

One of the relevant phenomena is the prenominal dative possessor construction German, which is fairly widespread but still considered substandard (see, e.g., Haider (1988), Zifonun (2004)). Here, a dative-marked possessor DP$_2$ shows up in SpecD of a matrix DP$_1$; and there is evidence that it has been moved there from the complement position of the noun (de Vries (2005), Chomsky (1970)). D$_1$ in turn is realized by a possessive pronoun with a dual role: The *root* of the pronoun agrees with DP$_{dat}$ (the *possessor*) with respect to [num(ber)] and [gend(er)]; and the *inflection* of the pronoun agrees with its complement NP (the *possessum*) with respect to [num], [gend], and [case]; see (50-a). A basic assumption is that the [gend] features of the possessive pronoun are not inherent; rather, they are determined in the course of the derivation, by Agree relations with gender-bearing nominals; so in principle, the possibility might arise that gender agreement were reversed, with the root of the pronoun agreeing with the possessum and the inflection agreeing with the possessor. This would yield (50-b), which needs to be excluded.[31]

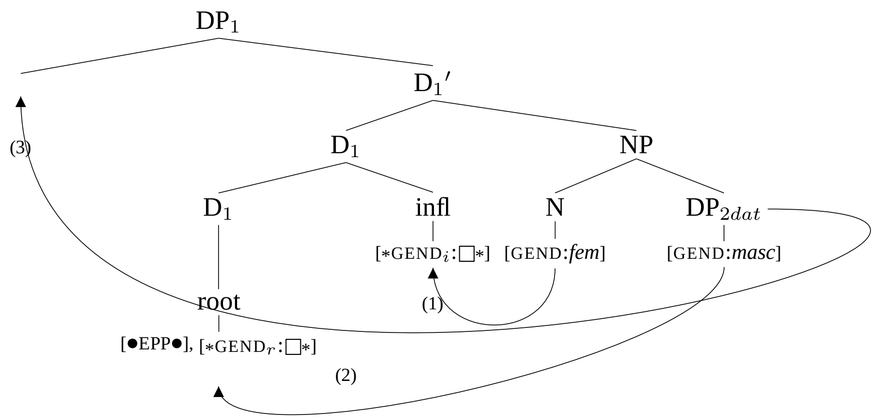
(50) a. [$_{\text{DP}_1}$ [$_{\text{DP}_2}$ dem Fritz] [$_{\text{D}'}$ [$_{\text{D}_1}$ sein-e] [$_{\text{NP}}$ [$_{\text{N}'}$ Schwester] t_2]]]
the.MASC.DAT Fritz his.MASC-FEM sister.FEM

b. *[$_{\text{DP}_1}$ [$_{\text{DP}_2}$ dem Fritz] [$_{\text{D}'}$ [$_{\text{D}_1}$ ihr-Ø] [$_{\text{NP}}$ [$_{\text{N}'}$ Schwester] t_2]]]
the.MASC.DAT Fritz her.FEM-MASC sister.FEM

"Fritz's sister"

The analysis is based on three constraints. First, the AGREE CONDITION (AC) demands an immediate valuation of so far unvalued features on an item if the structural context for Agree (roughly, m-command) is available. Second, the MERGE CONDITION (MC) requires structure-building operations (including movement, as an instance of internal Merge; see Chomsky (2008)) to take place immediately when the structural context for this operation is present. And third, the MINIMAL LINK CONDITION (MLC) states that all grammatical operations (like Agree and Merge) involve the smallest possible path between two items involved. By assumption, MLC is undominated (or belongs to GEN), and the ranking for German (or, at least, for derivational steps in the nominal domain in German) is AC $\gg$ MC. With this in mind, consider (51), which is the relevant stage (here called Σ stage) of the derivation of the prenominal dative possessor construction.[32]

(51) *The Σ Stage of the Derivation and the Subsequent Order of Operations*:



At stage Σ, various operations operation (signalled by (1), (2), and (3) in (51)) could in principle be carried out in the next step because the contexts for Agree and Move to apply are all met. However, given the ranking AC $\gg$ MC, gender valuation rather than movement has to apply next, and given the MLC, gender agreement must take place between the inflectional part of the possessive pronoun and the head noun of the construction, which minimizes path lengths for syntactic dependencies (operation (1) in (51)). This is shown in T_{26}. (Note that AC is still violated once by the optimal output O_1; the reason is that one gender feature of the pronoun is still unvalued even though the context for Agree to apply is present.)

The optimal output O_1 of this optimization is then used as the input for the next optimization procedure. As before, agreement is given preference to movement of DP_2 (because of AC $\gg$ MC), resulting in valuation of the remaining gender feature on the root of the possessive pronoun (operation (2) in (51)); see T_{27}.

Finally, movement can and must take place (operation (3) in (51)); O_{111} in T_{28} is the sole

T_{26}: Valuation of inflection's gender, step 1 (Σ as input): Agree (with possessum NP)

Input: $[_{D'_1}$ D$_{[*\text{GEND}_r:\square*],\ [*\text{GEND}_i:\square*],\ [\bullet\text{EPP}\bullet]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ DP$_{2[\text{GEND}:\textit{masc}]}$]]	MLC	AC	MC
☞O$_1$: $[_{D'_1}$ D$_{[*\text{GEND}_r:\square*],\ [\text{GEND}_i:\textit{fem}],\ [\bullet\text{EPP}\bullet]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ DP$_{2[\text{GEND}:\textit{masc}]}$]]		*	*
O$_2$: $[_{D'_1}$ D$_{[\text{GEND}_r:\textit{fem}],\ [*\text{GEND}_i:\square*],\ [\bullet\text{EPP}\bullet]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ DP$_{2[\text{GEND}:\textit{masc}]}$]]	*!	*	*
O$_3$: $[_{DP_1}$ DP$_{2[\text{GEND}:\textit{masc}]}$ D$_{[*\text{GEND}_r:\square*],\ [*\text{GEND}_i:\square*]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ t$_2$]]		**!	

T_{27}: Valuation of root's gender, step 2: Agree (with possessor DP)

Input: $[_{D'_1}$ D $_{[*\text{GEND}_r:\square*],\ [\text{GEND}_i:\textit{fem}],\ [\bullet\text{EPP}\bullet]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ DP$_{2[\text{GEND}:\textit{masc}]}$]]	MLC	AC	MC
☞O$_{11}$: $[_{D'_1}$ D$_{[\text{GEND}_r:\textit{masc}],\ [\text{GEND}_i:\textit{fem}],\ [\bullet\text{EPP}\bullet]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ DP$_{2[\text{GEND}:\textit{masc}]}$]]			*
O$_{12}$: $[_{DP_1}$ DP$_{2[\text{GEND}:\textit{masc}]}$ D$_{[\text{GEND}_r:\textit{masc}],\ [\text{GEND}_i:\textit{fem}]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ t$_2$]		*!	

remaining candidate (at least among those that have any chance of becoming optimal). The resulting order of operations is illustrated by numbered arrows in (51).

T_{28}: Possessor raising, step 3: Move

Input: $[_{D'_1}$ D$_{[\text{GEND}_r:\textit{masc}],\ [\text{GEND}_i:\textit{fem}],\ [\bullet\text{EPP}\bullet]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ DP$_{2[\text{GEND}:\textit{masc}]}$]]	MLC	AC	MC
☞O$_{111}$: $[_{DP_1}$ DP$_2$ D$_{[\text{GEND}_r:\textit{masc}],\ [\text{GEND}_i:\textit{fem}]}$ … …$[_{NP}$ N$_{[\text{GEND}:\textit{fem}]}$ t$_2$]]			

Suppose now that optimization applied to phrases, to phases, to clauses, or to full sentences – i.e., to any domain that is larger than the derivational step. 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; the *input* for optimization will then involve the full structure in (51), after movement to SpecD. Now, ceteris paribus, the unvalued gender feature on the inflectional part of the pronoun can receive value [masc], which derives (50-b). Thus, the approach overgenerates; see T$_{29}$, where O$_2$ in addition to O$_1$ is wrongly classified as optimal.

From a more general perspective, the argument presented here is a standard *counter-feeding* argument against strictly representational analyses (see Chomsky (1975), Kiparsky (1973)): Movement of DP$_2$ to SpecD$_1$ could in principle feed agreement of the inflectional part of D$_1$ with the possessor DP$_2$, but such movement comes too late in the derivation and therefore doesn't. Many arguments for serial local optimization are of this general type, involving either counter-feeding (where properties of the ultimate output representation suggest that an operation should have been able to apply even though it could not, the reason being that the context for application was not yet there at the crucial stage of the derivation) or counter-bleeding (where properties of the ultimate output representation suggest that an operation should not have been able to legimately apply even though evidently it could, the reason being that the properties that

T_{29}: DP (vP, CP, ...) optimization: wrong result

Input: $D_{[*\text{GEND}_r:\square*],\ [*\text{GEND}_i:\square*],\ [\bullet\text{EPP}\bullet]}$ $[_{\text{NP}}\ N_{[\text{GEND}:\textit{fem}]}\ DP2_{[\text{GEND}:\textit{masc}]}\ ]$	MLC	AC	MC	LR
☞O1: $[_{\text{DP}_1}\ DP2_{[\text{GEND}:\textit{masc}]}\ D_{[\text{GEND}_r:\textit{masc}],\ [\text{GEND}_i:\textit{fem}]} \cdots$ $\ldots[_{\text{NP}}\ N_{[\text{GEND}:\textit{fem}]}\ t_2\ ]]$				
☛O2: $[_{\text{DP}_1}\ DP2_{[\text{GEND}:\textit{masc}]}\ D_{[\text{GEND}_r:\textit{fem}],\ [\text{GEND}_i:\textit{masc}]} \cdots$ $\ldots[_{\text{NP}}\ N_{[\text{GEND}:\textit{fem}]}\ t_2\ ]]$				
O3: $[_{\text{DP}_1}\ D_{[\text{GEND}_r:\textit{masc}],\ [\text{GEND}_i:\textit{fem}],\ [\bullet\text{EPP}\bullet]} \cdots$ $\ldots[_{\text{NP}}\ N_{[\text{GEND}:\textit{fem}]}\ DP2_{[\text{GEND}:\textit{masc}]}\ ]]$			*!	
O4: $[_{\text{DP}_1}\ DP2_{[\text{GEND}:\textit{masc}]}\ D_{[*\text{GEND}_r:\square*],\ [\text{GEND}_i:\textit{fem}]} \cdots$ $\ldots[_{\text{NP}}\ N_{[\text{GEND}:\textit{fem}]}\ t_2\ ]]$		*!		

would block it were not there at an earlier stage in the derivation).

4.4 Problems for Local Domains for Competition Resolution

Serial local optimization makes a number of interesting predictions, opens up new areas for research, and chimes in well with recent developments in the minimalist program. Nevertheless, it faces challenges in domains where it looks as though more information must be available for optimization procedures than would be permitted under local optimization. *Repair* or *last resort* phenomena that seem to involve long-distance dependencies, like resumptive pronouns in island contexts and instances of long-distance binding, are a case in point. Thus, recall from subsection 2.2 that resumptive pronouns that show up with movement dependencies across islands (and only there) can straightforwardly be analyzed in terms of OT, as involving a violation of a DEP constraint (like INCLUSIVENESS) that may become optimal only if all competing outputs violate higher-ranked constraints (i.e., those that trigger movement, and those that block movement across an island). However, there is a problem with a naive transfer of the standard (parallel, global) OT approach to resumptive pronouns to a serial local OT approach: On the one hand, the crucial decision (trace or resumptive pronoun) must be made *very early* in the derivation, when only little structure is present yet. On the other hand, an island that serves as the trigger for last-resort resumption may come *quite late* in the derivation; the island may be separated from the extraction site by a huge amount of intervening structure. Consequently, the island will typically be too far away from the extraction site to successfully trigger resumption. The problem is illustrated in (52).

(52) ?(the man) who(m) I don't believe $[_{\text{DP}_1}$ the claim that anyone thinks that Mary believes that Bill $[_{\text{CP}_2}$ said that John $[_{\text{vP}_3}$ t $[_{\text{VP}_4}$ saw him]]]]

Here, VP_4 is the domain where the decision must be made under extremely local optimization of derivational steps (or of phrases); vP_3 is the domain if phases are optimization domains; and the decision must be made in CP_2 if clauses are optimization domains. However, the domain in which the relevant information (viz., presence of an island) becomes available is DP_1, which is far beyond any of the local domains for optimization that have been proposed.

One can either consider problems of this type to be fatal, or one can take them to pose an interesting challenge to the local optimization enterprise. Assuming the latter, there are various ways to look for solutions. For instance, it has been proposed that morphological realization in (extended) chains permits exceptions to the Strict Cycle Condition (see Chomsky (1973)), so that the derivation may in fact backtrack, and selectively change earlier material (see particularly

Fischer's (2006) local OT approach to binding in terms of a 'wormhole theory').[33] Alternatively, one might argue that the relevant information (concerning islands) is in fact already present presyntactically (in the numeration); the decision can be made before the syntactic derivations starts.[34] Irrespective of these considerations, though, data like (52) suggest that potential problems with local optimization arise independently of the exact size of the optimization domain (given that it is not the entire sentence). Arguably, for conceptual reasons, this might then favour the choice of the smallest possible optimization domain, at least as a plausible research strategy.

5. Conclusion

What is the current status of OT syntax in the field of linguistics? One cannot help but notice that as a common research program, OT syntax is not well. Various considerations support this conclusion: First, at the time of writing, there do not seem to be regular workshops expressly devoted to OT syntax anymore.[35] Second, very few OT syntax papers have appeared in leading journals over the last few years. Third, the few papers that have appeared in leading journals 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 OT syntax. Fourth, whereas new edited volumes with a focus on OT syntax came out on a regular basis for some time (see, e.g., Archangeli & Langendoen (1997), Dekkers et al. (2001), Fanselow & Féry (2002b), Legendre, Grimshaw & Vikner (1998), Müller & Sternefeld (2001b), Sells (2001b)), this seems to have all but stopped.[36] Fifth, few influential dissertations on OT syntax have appeared in recent years (since, say, Zepter (2004), Engels (2004), Fischer (2004)), and virtually none (as far as I can tell) in the US. All this is very different from the situation in morphology, semantics, pragmatics (here see particularly the work on bidirectional OT going back to Blutner (2000) and Jäger & Blutner (2000)) and, especially, phonology, where OT thrives to this day.

Thus, the immediate prospects for OT syntax as a self-sufficient, viable research programme can be viewed as bleak. However, there is a legacy of OT syntax: In a sense, it lives on in other theories. In particular, 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 important part of the minimalist program, even though this is typically not acknowledged. For reasons of space, I cannot possibly go through a substantial number of the cases of "hidden" optimization here, or provide detailed argumentation to support the claim that hidden optimization is often involved in work that purports to do without optimization; but it is clear that many of the relevant analyses are concerned with phenomena that suggest constraint conflict, repair (last resort), or defaults. Let me confine myself to listing a few examples where implicit optimization procedures (that must be construed with violable and ranked constraints if made explicit) show up in work in the Principles and Parameters tradition. In Müller (2000b), I argue for hidden optimization in Chomsky's (1981) analysis of pronoun vs. PRO in English gerunds based on the transderivational constraint Avoid Pronoun (reconstructed ranking: OBLCONTROL $\gg$ *PRON); in Haegeman's (1995) analysis of pro vs. overt pronoun in pro-drop languages based on Avoid Pronoun (reconstructed ranking: TOP/PRO $\gg$ *PRON); in von Stechow & Sternefeld's (1988) analysis of lexical vs. structural control in German (reconstructed ranking: FAITH(LEX) $\gg$ OBLCONTROL); in Kayne's (1994) analysis of complementizer-finality and the absence of overt *wh*-movement in Japanese (reconstructed ranking: IP-CRIT $\gg$ WH-CRIT); in Grewendorf's (2001) analysis of multiple *wh*-questions in German (reconstructed ranking: *COMPLEX-WH $\gg$ WH-REAL); and in Roberts's (1997) approach to phonological realization in head chains (reconstructed ranking: *COMPLEXHEAD $\gg$ HEAD-REAL). In Heck, Müller & Trommer (2008), we show that analysis of definiteness marking in Swedish DPs in Embick & Noyer (2001) relies on an implicit ranking of various constraints: N-DEF, D-DEF, HMC $\gg$ N-TO-D $\gg$ *DISSOCIATION, FULL-INT. Lahne (2009) observes that the analysis of Agree relations in Haegeman & Lohndal (2008) depends on a ranking MINIMALITY, FEATURE MATCHING $\gg$ AGREE. Samek-Lodovici (2006a) points out that the analysis of strong and weak pronouns in Cardinaletti & Starke (1999) is ultimately based on a ranking CHECK-F, PARSE $\gg$ *STRUC $\gg$ STAY. And so on.

As remarked above, optimization procedures play an important role in the minimalist program, independently of particular analyses of linguistic phenomena of the type just mentioned. First, earlier versions of the minimalist program regularly employed transderivational constraints like Fewest Steps and Shortest Paths, which involve optimization of a type that is very similar to that adopted in standard OT (see Müller & Sternefeld (2001a) for an overview, and Chomsky (1993; 1995), Collins (1994), and Bošković (1997) for some relevant cases). Second, at the heart of the minimalist program are elementary operations like Agree, Merge, Move, Delete, Transfer, and 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. This is what Heck & Müller (2007; 2010) argue for in the case of Agree vs. Move (or, more generally, Agree vs. Merge; see above). A far more widespread interaction of requirements for elementary operations concerns Merge vs. Move operations, as they have been discussed under the label of "Merge before Move" in Chomsky (1995; 2000; 2001; 2005), Frampton & Gutmann (1999), Hornstein (2001; 2009) and many other minimalist analyses, for a variety of phenomena including expletive sentences and adjunct control. In its original conception, Merge before Move is a transderivational constraint. Frampton & Gutmann (1999) suggest the formulation in (53), which brings the constraint closer to the perspective adopted in the previous section.

(53) *Merge before Move*:
Suppose that the derivation has reached stage Σ_n, and Σ_{n+1} is a legitimate instance of Merge, and Σ'_{n+1} is a legitimate instance of Move. Then, Σ_{n+1} is to be preferred over Σ'_{n+1}.

The optimality-theoretic reconstruction is straightforward: A MERGE CONDITION outranks a more specific MOVE CONDITION, as in (54), with the derivational step as the (extremely local) optimization domain.[37]

(54) a. MERGE CONDITION (MC):
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.

Does the ranking that is required to derive the effects of (53) have to be universal, or can it be reversed in principle (as suggested by Broekhuis & Klooster (2001))? If the latter is the case, can the ranking vary from one syntactic domain (or category) to another one? At present, these are open questions which, however, strike me as quite important, and which should definitely incite further interesting research.

Another example illustrating hidden optimization in core parts of minimalist syntax concerns the Inclusiveness condition adopted in Chomsky (2001) and much subsequent related work (see above). An INCLUSIVENESS constraint demands that nothing may enter the syntactic derivation which is not part of the original numeration; however, this DEP-type constraint 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. Arguably, the same conclusion can be drawn for the mechanism of *feature valuation* as part of Agree; the copy mechanism required here gives rise to a straightforward DEP violation. Similarly, the *copy theory of movement* (Chomsky (1993)) would seem to systematically require violability of INCLUSIVENESS.

Finally, it is worth pointing out that implicit optimization in the minimalist program is not confined to conflicting demands imposed by basic operations. For instance, an idea that has been widely pursued in recent years is that attempts at carrying out an Agree operation may in principle fail without necessarily giving rise to ungrammaticality. Rather, a second, different attempt can be made to establish an Agree operation; see Béjar & Řezáč (2009), Bošković (2009), Patel (2010), and d'Alessandro (2012), among others, on such "second-cycle Agree" (Georgi (2010) even argues for "third-cycle Agree" effects). On the simplest interpretation, this clearly presupposes violability of the constraint that triggers Agree in a well-formed output. In the same way, Chomsky's idea (presented at various talks in recent years) that movement is *triggered* by the necessity to break an otherwise existing symmetry in syntactic structure (which then may preclude a labelling of constituents; see Moro (2007), Boeckx (2008), Ott (2012) for some applications of this or a similar idea) would seem to strongly suggest an optimality-theoretic mechanism at its very core.

Thus, OT syntax may be endangered as a research programme sui generis, but based on the preceding remarks, I would like to contend that minimalist syntax is inherently optimality-theoretic at its very core. Independently of this, OT syntax is, in my view, well worth pursuing, and not just for the more obvious reasons having to do with the existence of repair phenomena, constraint conflict, and default forms in natural languages: OT syntax permits a radically new perspective on various kinds of phenomena, one that would not be available in approaches that do not envisage constraint violability and constraint ranking. To see this, consider, finally, the gist of the account of *wh*-island effects developed in Legendre, Smolensky & Wilson (1998); unlike most other accounts, this analysis does not rely on a concept of intervention (as in Rizzi (1990; 2004)). In this alternative account, *all* movement from an embedded clause significantly violates locality constraints. Such a violation is fatal if the ultimate target position of the *wh*-phrase that is supposed to undergo long-distance movement can be changed from the matrix clause to the embedded clause without triggering a violation of selection requirements. This is possible, hence obligatory, with embedded *wh*-clauses, which are objects of [+wh]-selecting verbs. However, such a locality violation with movement from a clause is permissible as a last resort if the ultimate target position of the *wh*-phrase that is supposed to undergo long-distance movement cannot be relocated to the embedded clause without violating selection requirements. This is the case with embedded declarative clauses, which are objects of [–wh]-selecting verbs. So, surprisingly, what *rules out wh*-island constructions is the fact that a violation of locality *can be avoided* by relocating the *wh*-scope to the embedded clause; and what *permits* extraction from declarative complements is the fact that a violation of locality *cannot be avoided* here. Evidently, there is no room for elegant and highly innovative reasonings of this type in non-optimality-theoretic approaches. For reasons like this, a renaissance of OT syntax, however unlikely, might do the field good.

Notes

[*]I am grateful to an anonymous reviewer for very helpful comments on a previous version of this chapter.

[1]Here and henceforth, 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).

[2]It has sometimes been argued that there is a difference between, e.g., "weak" and "strong" violations of constraints on movement, such as the Subjacency Condition vs. the Empty Category Principle (ECP) in Chomsky (1986). However, this is just stipulated on top of the grammatical decision procedure (yes/no), and does not reflect a genuine interaction of constraints (let along a ranking).

[3]Similar questions, and similar kinds of variation, can be found in (mostly early) versions of the minimalist program (Chomsky (1995; 2001)) that rely on *transderivational constraints* which choose among a set of competing derivations in a candidate set; see Sternefeld (1996) and references cited there. In the minimalist tradition, candidate sets are usualy referred to as *reference sets*.

[4]It should be emphasized that this interaction of components, including the order (generation → optimization) instantiates an abstraction as it is standard in most grammatical theories. It should not be understood as making claims about OT as a model of linguistic cognition in the narrow sense, such that, e.g., first a (possibly infinite; see below) set of candidates is generated in actual language processing that is then subjected to optimization, with the best candidate chosen. Here, techniques have been suggested (at least for certain empirical phenomena) that make it possible to determine the optimal candidate without building all competitors first; see Tesar (1995) and Riggle (2004).

[5]In Heck et al. (2002) it is argued that the two basic motivations for inputs in phonology – viz., (i) defining candidate sets and (ii) providing information for faithfulness constraints (see below) – are either unavailable or irrelevant in syntax. More specifically, (i) is unavailable because candidate sets cannot adequately be defined by resorting to input information only, and (ii) is irrelevant because syntax, unlike, phonology, is an information-preserving system, with, e.g., subcategorization information present on a verb throughout the derivation.

[6]Under the input-free conception of OT syntax mentioned above, DEP, MAX and IDENT constraints all have to be reformulated as constraints that are purely output-oriented. In the case of DEP constraints, this will involve markedness constraints banning items with property P, where P is the property that kept the item from appearing in the input in the first place. For instance, rather than violating faithfulness qua not appearing in the input, expletives, under the input-free view, would violate markedness qua being semantically empty, which would be just the property responsible for their non-occurrence in inputs in the standard OT model; see Grimshaw (1997) for such an approach.

[7]See Prince & Smolensky (2004, ch. 10) and Smolensky & Legendre (2006, part I) for detailed elaboration of the differences between these two kinds of approaches.

[8]For the sake of clarity, a specification of the input is provided in the form of a numeration; this is of no further importance in the present context.

[9]Arguably, (8) can (and, if so, should) be derived from yet more general constraints and their interaction; Grimshaw (2001; 2006) has come up with promising attempts to achieve this (also see Steddy & Samek-Lodovici (2009) for an application of the underlying logic to universal constraints on DP-internal order of D, Number, A and N): On the one hand, it can be observed that all syntactic constituents violate so-called alignment constraints that dictate the left-peripheral or right-peripheral placement of items. Given dichotomies like HEAD-LEFT/HEAD-RIGHT and COMPLEMENT-LEFT/COMPLEMENT-RIGHT, with (due to the universality of constraints) both inherently conflicting constraints of a pair active in every language even if only one of the two actually determines a given order, it is clear that more structure will invariably imply more violations of alignment constraints (viz., the ones which are violated in any given structure). Movement is structure-building; therefore, any ECONOMY violation will also trigger a violation of alignment (see Grimshaw (2001)). On the other hand, as remarked above, all movement chains in outputs are trivial (i.e., single-membered) in the input. Movement gives rise to non-trivial (i.e., multi-membered) chains. This implies a violation of faithfulness (IDENT/UNIQUENESS; see Grimshaw (2006)).

[10]English may ultimately not be the best language to illustrate the phenomenon because (9-d) is somewhat marked. Still, it may do in the present context because we are only interested in the general pattern, not in a comprehensive and empirically impeccable analysis of a single language. That said, data from other languages instantiate exactly the same pattern, and in a clearer way. For instance, resumptive pronouns in Swedish are also confined to island contexts, where traces are blocked (with some well-defined exceptional cases where there is optionality); see Engdahl (1982; 1985), Sells (1984). Thus, resumptive pronoun insertion is a repair/last resort operation amenable to the same type of analysis given in the text.

[11]There may eventually be much more general constraints, or sets of constraints, replacing both REL-CRIT and CNPC, but this is immaterial for the logic of the argument.

[12]The ranking of REL-CRIT and CNPC is actually not crucial here but will be assumed to be strict to simplify matters.

[13]As a matter of fact, this presupposes that optimization can be serial and local. See section 4 below.

[14]Whenever it seems to be unproblematic, as in approaches to syntax that envisage blocking (see Williams (1997), Fanselow (1991)), or in Distributed Morphology (Halle & Marantz (1993)), this is due to the fact that the approach in fact shares crucial features with OT – in the case at hand, that it is based on competition and candidate sets, too.

[15](17-a) may plausibly be viewed as a subcase of a more general constraint demanding faithfulness to lexical case specifications. At this point, it can be noted that despite initial appearances to the contrary, OT is arguably not perfectly well designed to capture lexical exceptions via faithfulness to lexical specifications. Here is why: Suppose that a lexical item α is lexically specified as demanding property P in the output (e.g., a verb governs genitive case on its internal argument DP). If a faithfulness constraint demanding preservation of this information in the output is sufficiently highly ranked, P shows up in the output, as desired (e.g., the DP is marked genitive). However, there is no intrinsic requirement for faithfulness for lexical specifications to outrank conflicting constraints in a language, and this means that the situation may well occur that exceptional lexical specifications may be present on lexical items without ever showing up in optimal outputs. To take a far-fetched example: All transitive verbs in German might be lexically specified as governing ergative case for their subject, or as governing instrumental case on a direct object, but with high-ranked case-government constraints outranking the respective faithfulness constraints demanding ergative (or instrumental), this information can never make it to the surface. On this view, peculiar ambiguities may arise: A grammar of German with ergative specifications on transitive verbs and another grammar without will yield the same output; this would give rise to principled redundancies. Potential problems of this type show up systematically in OT. They have been addressed by invoking a meta-optimization procedure ('input optimization') that is related to learnability considerations in Prince & Smolensky (2004). Also see Prince (2006, 52-58) for pertinent discussion.

[16]To name just one further phenomenon: Movement often seems to obey an order preservation constraint. However, to permit permutation at all (as it arises, e.g., when an object *wh*-DP moves to a local SpecC position across a subject DP), such a constraint must clearly be violable (see Williams (2003)). This suggests an OT approach where the constraint demanding order preservation is ranked low but springs into action when all pertinent higher-ranked constraints do not distinguish the candidates; see Müller (2001) and Engels & Vikner (2006).

[17]Why, then, do transitive clauses without lexical case specification on the verb not always just involve nominative marking on all arguments? One possible answer might be that there is a high-ranked (or inviolable) case distinctness requirement counter-acting such a state of affairs; crucially, though, such a requirement would always be fulfilled in quirky subject contexts like the ones currently under consideration.

[18]Note that issues of input optimization of the type discussed in footnote 15 become relevant here again.

[19]The reason is that the two constraints are in a special-to-general relation. However, one must be careful here. As observed by Prince & Smolensky (2004) two cases must be distinguished with constraints that are in a special-to-general relation: On the one hand, the two constraints may impose conflicting requirements on candidates. In that case, they form a *Paninian relation*, and ranking the more specific one lower than the conflicting, more general one will invariably imply that the former one becomes inactive. On the other hand, the two constraints may actually push candidates in the same direction, as in the case currently under consideration. The constraints can then be said to form a *stringency relation* (see Baković (1995)); here, the more specific constraint can in principle be the lower-ranked one and still carry out some work if a more complex system of constraints is considered.

[20]A further objection to OT analyses that one can hear every now and then is that the theory is inherently unconstrained in the sense that "anything goes"; e.g., potentially problematic predictions of an existing analysis can be avoided by adding another high-ranked ad hoc constraint. While technically correct, such a criticism misses a fundamental point: Criteria of elegance and simplicity hold for OT syntax in the same way that they hold for other syntactic theories. Consequently, adding stipulative, highly specific constraints is ultimately not an option in OT for the same reason that, say, adding stipulative, highly specific exceptions to constraints is rightly frowned upon in non-OT work. This is tacitly presupposed in all good work in OT syntax; it is explicitly stated in (inter alia) Grimshaw (1998) and Smolensky & Legendre (2006).

[21]Also see Prince & Samek-Lodovici (1999) for the extended concept of *collective harmonic bounding*, i.e., cases where it is not a single candidate but a set of candidates that harmonically bounds a candiate, which can therefore never become optimal in any language.

[22]Ultimately, a bit more will have to be said in this kind of neutralization approach, concerning, e.g., whether C is also marked [+wh] vs. [–wh] in the two inputs, and this feature value must then also be altered. However, this issue does not affect the logic of the neutralization approach as such.

[23]For more on ineffability, see Müller (2000b), Fanselow & Féry (2002a), Legendre (2009), Vogel (2009b).

[24]However, recall that the definition of optimality in (1) is in principle compatible with there being more than one winner.

[25]Terminology is not uniform is this domain. Legendre (2001) reserves the term "tie" for what I call "local tie," and refers to what I call "global tie" as "floating constraints" or "partial orderings." Prince & Smolensky (1993) label what I refer to as a local tie "crucial nonranking".

[26]However, also see Kiparsky (2000), Itô & Mester (2002), and Bermúdez-Otero (2008) for stratal OT, where this reasoning does not hold.

[27]In fact, given the more recent concept of phase-based spell-out, there is *no final representation of the full sentence* in minimalist syntax. Under this assumption, Legendre et al.'s first option is not available on principled grounds.
At this point, one may ask whether it would be possible to postulate local serial optimization in a classical transformational setting, or in a strictly representational approach. The answer to both questions would seem to be affirmative (Janes Grimshaw, p.c.): Even in a representational approach, one might first optimize a subpart (e.g., the most deeply embedded clause); and then take the subpart to be invariant, and let it influence properties of the next domain (i.e., bring about changes in the existing structure there). However, to the best of my knowledge, such serial, local optimization of complete LEs has not yet been pursued in OT syntax.

[28]Note that these numbers are given solely for the purpose of illustration; they may or may not come close to the actual state of affairs, depending on a variety of further decisions about clause structure that are orthogonal to present concerns.

[29]Also see Pesetsky (1998), Broekhuis & Dekkers (2000), Broekhuis (2000; 2006; 2008), Fanselow & Ćavar (2001), and Heck & Müller (2000), among others.

[30]Also see Epstein & Seely (2002, 77), who argue that "each transformational rule application constitutes a 'phase'."

[31]Taking (50-b) to be a serious competitor of (50-a) is presumably not an artefact of the theory. Young children as well as second language learners of German have well-documented problems with getting the two types of gender agreement (i.e., root vs. inflection) with third-person possessive pronouns right; see, e.g., Ruff (2000). Furthermore, gender mistakes with possessive pronouns regularly occur even in adult speech, and are then frowned upon by language mavens; see, e.g., Sick (2006, 108).

[32]Some remarks on notation. [$*$GEND$_r$:$\square *$], [$*$GEND$_i$:$\square *$] are unvalued gender features of the root and inflectional parts of the possessive pronoun, respectively, that require valuation by Agree with a gender-bearing nominal; [●EPP●] is a property of the possessive pronoun that triggers movement of the possessor DP_2 to SpecD, yielding the eventual surface order of constitutents.

[33]Also compare multidominance approaches to movement, which require a similar assumption; see Gärtner (2002), Frampton (2004), among others.

[34]In this context, see Boeckx's (2010) objection to minimalist approaches where most syntactic operations are determined by the properties of lexical items (e.g., their structure-building and probe features), and where syntax "blindly follows lexical instructions". If he is right, solving the locality problem with resumptive pronouns should be straightforward in such an approach.

[35]There were such workshops on OT 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.

[36]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 "OT syntax". Also, 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. However, it can be noted that a volume (mainly) on OT syntax (even though it is to some extent based on the earlier two working paper volumes) is advertised for 2012: *Linguistic Derivations and Filtering*, edited by Hans Broekhuis and Ralf Vogel. Another possible exception is the colletion of papers in Samek-Lodovici (2006b) that goes back to an LSA workshop in 2005.

[37]The MC put to use in subsection 4.3 would have to be adapted accordingly.

References

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

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

Aissen, Judith (2003a): Differential Object Marking: Iconicity vs. Economy, *Natural Language and Linguistic Theory* 21, 435–483.

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

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

Aoun, Joseph, Norbert Hornstein & Dominique Sportiche (1981): Aspects of Wide Scope Interpretation, *Journal of Linguistic Research* 1, 69–95.

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

Baković, Eric (1995): A Markedness Subhierarchy in Syntax. Ms., Rutgers University.

Baković, Eric (1998): Optimality and Inversion in Spanish. In: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds., *Is the Best Good Enough?*. MIT Press and MITWPL, Cambridge, Mass., pp. 35–58.

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

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

Bermúdez-Otero, Ricardo (2008): *Stratal Optimality Theory*. Book Ms., University of Manchester. To appear: Oxford University Press.

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

Boeckx, Cedric (2008): *Bare Syntax*. Oxford University Press, Oxford.

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

Boeckx, Cedric (2010): Defeating Lexicocentrism, Part A. Ms., ICREA/UAB.

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

Bošković, Željko (1997): *The Syntax of Nonfinite Complementation. An Economy Approach*. MIT Press, Cambridge, Mass.

Bošković, Željko (2009): Unifying First and Last Conjunct Agreement, *Natural Language and Linguistic Theory* 27, 455–496.

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.

Broekhuis, Hans & Wim Klooster (2001): On Merge and Move/Attract. In: *Progress in Grammar: Articles at the 20th Anniversary of the Comparison of Grammatical Models Group in Tilburg*. Tilburg University, pp. 1–20.

Broihier, Kevin (1995): Optimality Theoretic Rankings with Tied Constraints: Slavic Relatives, Resumptive Pronouns and Learnability. Ms., MIT, Cambridge, Mass.

Büring, Daniel (2001): Let's Phrase It. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 69–105.

Butt, Miriam & Tracy Holloway King (2012): Lexical Functional Grammar. this volume.

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.

Choi, Hye-Won (1999): *Optimizing Structure in Context. Scrambling and Information Structure*. CSLI Publications, Stanford.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.
Chomsky, Noam (1970): Remarks on Nominalization. In: R. Jacobs & P. Rosenbaum, eds., *Readings in English Transformational Grammar*. Ginn and Company, Waltham, Mass., pp. 184–221.
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 (1975): *The Logical Structure of Linguistic Theory*. Plenum Press, New York.
Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.
Chomsky, Noam (1986): *Barriers*. MIT Press, Cambridge, Mass.
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.
Chomsky, Noam & Howard Lasnik (1993): Principles and Parameters Theory. In: J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann, eds., *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. Vol. 1, de Gruyter, Berlin, pp. 506–569.
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.
Costa, João (1998): Word Order Variation. PhD thesis, Universiteit Leiden.
d'Alessandro, Roberta (2012): Agreement, Ergativity, and the Parametrization of Probes. Ms., Leiden University.
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.
Engdahl, Elisabet (1982): Restrictions on Unbounded Dependencies in Swedish, *Readings on Unbounded Dependencies in Scandinavian Languages* pp. 151–174.
Engdahl, Elisabet (1985): Parasitic Gaps, Resumptive Pronouns, and Subject Extractions, *Linguistics* 23(1), 3–44.
Engdahl, Elisabet, Maia Andréasson & Kersti Börjars (2004): Word Order in the Swedish Midfield – An OT Approach. In: F. Karlsson, ed., *Proceedings of the 20th Scandinavian Conference of Linguistics*. Vol. 36, Department of General Linguistics, University of Helsinki, pp. 1–12.
Engels, Eva (2004): Adverb Placement. An Optimality-Theoretic Approach. PhD thesis, Universität Potsdam.
Engels, Eva & Sten Vikner (2006): An Optimality-Theoretic Analysis of Scandinavian Object Shift and Remnant VP-Topicalisation, *Linguistics in Potsdam* 25, 195–231.
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 (2004): Copies, Traces, Occurrences, and All That. Ms., Northeastern University.
Frampton, John & Sam Gutmann (1999): Cyclic Computation, *Syntax* 2, 1–27.
Gabriel, Christoph (2010): On Focus, Prosody, and Word Order in Argentinean Spanish: A Minimalist Ot Account, *Revista Virtual de Estudos da Linguagem* 4, 183–222.
Gallego, Ángel (2007): Phase Theory and Parametric Variation. PhD thesis, Universitat Autónoma de Barcelona, Barcelona.
Gärtner, Hans-Martin (2002): *Generalized Transformations and Beyond*. Akademie-Verlag, Berlin.
Georgi, Doreen (2009): Local Modelling of Global Case Splits. Master's thesis, Universität Leipzig.
Georgi, Doreen (2010): Third Cycle Agree Effects in Mordvin. In: S. Bank, D. Georgi & J. Trommer, eds., *2 in Agreement*. Vol. 88 of *Linguistische Arbeitsberichte*, Universität Leipzig, pp. 125–161.
Gigerenzer, Gerd, P.M. Todd & the ABC Group (1999): *Simple Heuristics that Make us Smart*. Oxford University Press, New York.
Grewendorf, Günther (2001): Multiple Wh-Fronting, *Linguistic Inquiry* 32, 87–122.
Grimshaw, Jane (1994): Heads and Optimality. Ms., Rutgers University. (Handout of talk, Universität Stuttgart).
Grimshaw, Jane (1997): Projection, Heads, and Optimality, *Linguistic Inquiry* 28, 373–422.
Grimshaw, Jane (1998): Constraints on Constraints in Optimality Theoretic Syntax. Ms., Rutgers University, New Brunswick, New Jersey.
Grimshaw, Jane (1999): Heads and Clauses. Ms., Rutgers University.
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.
Grimshaw, Jane (2010): Do-Support and Last Resorts. Ms., Rutgers University.
Gutiérrez-Bravo, Rodrigo (2007): Prominence Scales and Unmarked Word Order in Spanish, *Natural Language and Linguistic Theory* 25, 235–271.
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 (2010): Extremely Local Optimization. Ms., Universität Leipzig. To appear in Hans Broekhuis and Ralf Vogel (eds.), *Derivations and Filtering*. Equinox Publishing.

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.

Hornstein, Norbert (2009): *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge University Press, Cambridge.

Itô, Junko & Armin Mester (2002): Lexical and Postlexical Phonology in Optimality Theory. In: G. Fanselow & C. Féry, eds., *Resolving Conflicts in Grammars: Optimality Theory in Syntax, Morphology, and Phonology*. Buske, Hamburg, pp. 183–207. Special issue of *Linguistische Berichte*.

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.

Keer, Ed & Eric Baković (2001): Optionality and Ineffability. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 97–112.

Keer, Ed & Eric Baković (2004): Have Faith in Syntax. In: E. Curtis, J. Lyle & G. Webster, eds., *Proceedings of the 16th WCCFL*. CSLI Publications, Stanford, California, pp. 255–269.

Keine, Stefan (2010): *Case and Agreement from Fringe to Core. Impoverishment Effects on Agree*. Linguistische Arbeiten, Mouton de Gruyter, Berlin.

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 (1973): 'Elsewhere' in Phonology. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 93–106.

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 (2000): Opacity and Cyclicity, *The Linguistic Review* 17, 351–367.

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.

Kroch, Anthony (2001): Syntactic change, *The Handbook of Contemporary Syntactic Theory* pp. 699–729.

Kuhn, Jonas (2001): Formal and Computational Aspects of Optimality-Theoretic Syntax. PhD thesis, Universität Stuttgart.

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, Colin Wilson, Paul Smolensky, Kristin Homer & William Raymond (1995): Optimality and Wh-Extraction. In: J. Beckman, L. Walsh-Dickie & S. Urbanczyk, eds., *Papers in Optimality Theory*. Number 18 *in* 'Occasional Papers in Linguistics', UMass, Amherst, Massachussetts, pp. 607–635.

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

McCarthy, John (2002): *A Thematic Guide to Optimality Theory*. Cambridge University Press, Cambridge.

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

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

McCarthy, John (2010): An Introduction to Harmonic Serialism. Ms., University of Massachusetts, Amherst.

Merchant, Jason (2001): *The Syntax of Silence - Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press, Oxford.

Moro, Andrea (2007): Some Notes on Unstable Structures. Ms., Universitá Vita–Salute San Raffaele.

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

Müller, Gereon (1999): Optimality, Markedness, and Word Order in German, *Linguistics* 37, 777–818.

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 (2001): Order Preservation, Parallel Movement, and the Emergence of the Unmarked. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 279–313.

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 (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 (2001a): The Rise of Competition in Syntax: A Synopsis. In:

G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 1–68.

Müller, Gereon & Wolfgang Sternefeld, eds. (2001b): *Competition in Syntax*. Studies in Generative Grammar, Mouton de Gruyter, Berlin.

Ott, Dennis (2012): *Local Instability*. Vol. 544 of *Linguistische Arbeiten*, Mouton de Gruyter.

Patel, Pritty (2010): Disagree to Agree. Ms., MIT, Cambridge, Mass.

Pesetsky, David (1997): Optimality Theory and Syntax: Movement and Pronunciation. In: D. Archangeli & T. Langendoen, eds., *Optimality Theory. An Overview*. Blackwell, Oxford, pp. 134–170.

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.

Pollard, Carl J. & Ivan A. Sag (1994): *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Prince, Alan (2006): Implication and Impossibility in Grammatical Systems. Ms., Rutgers University. Roa.rutgers.edu/files/880-1006/880-PRINCE-0-0.PDF.

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.

Prince, Alan & Vieri Samek-Lodovici (1999): Optima. Ms., ROA 363.

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

Riggle, Jason (2004): Generation, Recognition, and Learning in Finite State Optimality Theory. PhD thesis, University of California, Los Angeles.

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.

Ross, John (1967): Constraints on Variables in Syntax. PhD thesis, MIT, Cambridge, Mass.

Ruff, Claudia (2000): Besitz und Besitzer in der Sprachentwicklung von deutschen und italienischen Kindern. PhD thesis, Universität Braunschweig.

Sailer, Manfred (2002): The German Incredulity Response Construction and the Hierarchial Organization of Constructions. Ms., Universität Göttingen. Talk, 2nd International Conference on Construction Grammar.

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 (1996): Constraints on Subjects. An Optimality-Theoretic Analysis. PhD thesis, Rutgers University, New Brunswick, New Jersey.

Samek-Lodovici, Vieri (2005): Prosody-Syntax Interaction in the Expression of Focus, *Natural Language and Linguistic Theory* 23, 687–755.

Samek-Lodovici, Vieri (2006a): 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.

Samek-Lodovici, Vieri (2006b): *Studies in OT Syntax and Semantics*. Elsevier, Amsterdam. *Lingua* Special Issue 9, vol. 117.

Schmid, Tanja (2001): OT Accounts of Optionality: A Comparison of Global Ties and Neutralization. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 283–319.

Schmid, Tanja (2005): *Infinitival Syntax. Infinitivus Pro Participio as a Repair Strategy*. Benjamins, Amsterdam.

Sells, Peter (1984): Syntax and Semantics of Resumptive Pronouns. PhD thesis, University of Massachusetts, Amherst.

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.
Sells, Peter, John Rickford & Thomas Wasow (1996): An Optimality Theoretic Approach to Variation in Negative Inversion in AAVE, *Natural Language and Linguistic Theory* 14, 591–627.
Shlonsky, Ur (1992): Resumptive Pronouns as a Last Resort, *Linguistic Inquiry* 23, 443–468.
Sick, Bastian (2006): *Der Dativ ist dem Genitiv sein Tod. Noch mehr Neues aus dem Irrgarten der deutschen Sprache. Folge 3*. 5 edn, KiWi Paperback. (2007).
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 (1991): *Syntaktische Grenzen. Chomskys Barrierentheorie und ihre Weiterentwicklungen*. Westdeutscher Verlag, Opladen.
Sternefeld, Wolfgang (1996): Comparing Reference Sets. In: C. Wilder, H.-M. Gärtner & M. Bierwisch, eds., *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 81–114.
Stiebels, Barbara (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.
Tesar, Bruce (1995): Computational Optimality Theory. PhD thesis, University of Colorado.
Tesar, Bruce (1998): Error-Driven Learing in Optimality Theory via the Efficient Computation of Optimal Forms. In: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds., *Is the Best Good Enough?*. MIT Press and MITWPL, Cambridge, Mass., pp. 421–435.
Vikner, Sten (2001a): V-to-I Movement and Do-Insertion in Optimality Theory. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 424–464.
Vikner, Sten (2001b): Verb Movement Variation in Germanic and Optimality Theory. Habilitation thesis, Universität Tübingen.
Vogel, Ralf (2001): Case Conflict in German Free Relative Constructions. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 341–375.
Vogel, Ralf (2009a): Skandal im Verbkomplex. Betrachtungen zur scheinbar inkorrekten Morphologie in infiniten Verbkomplexen im Deutschen, *Zeitschrift für Sprachwissenschaft* 28, 307–346.
Vogel, Ralf (2009b): 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.
Williams, Edwin (2003): *Representation Theory*. MIT Press, Cambridge, Mass.

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

Woolford, Ellen (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.

Zepter, Alexandra (2004): Phrase Structure Directionality: Having a few Choices. PhD thesis, Rutgers University, New Brunswick, New Jersey.

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