

Optimality-Theoretic Syntax

Gereon Müller

Institut für Linguistik
Universität Leipzig

SoSe 2008

www.uni-leipzig.de/~muellerg

Hintergrund

Die Optimalitätstheorie (OT) ist Anfang der neunziger Jahre von Alan Prince, Paul Smolensky, John McCarthy und anderen entwickelt worden, und zwar zunächst hauptsächlich für die Phonologie. Die ausführlichste und bis heute im wesentlichen kanonische Darstellung der Theorie findet sich im Buch von Alan Prince & Paul Smolensky (1993; 2004): [Optimality Theory: Constraint Interaction in Generative Grammar](#); Oxford: Blackwell.

(1) *Zentrale Annahmen der Optimalitätstheorie:*

- a. *Universalität:*
Beschränkungen sind universell.
- b. *Verletzbarkeit:*
Beschränkungen können verletzt werden.
- c. *Geordnetheit:*
Beschränkungen sind geordnet.
- d. *Wettbewerb:*
Die Grammatikalität eines Kandidaten K ist nicht allein aufgrund interner Eigenschaften von K ermittelbar; vielmehr entscheiden externe Faktoren (der Wettbewerb von K mit anderen Kandidaten) über die Wohlgeformtheit von K.

Bemerkung:

K steht für eine grammatische Einheit, die Gegenstand des Optimierungsverfahrens ist, das über ihre Grammatikalität entscheidet. K ist jeweils die Grundeinheit von Phonologie, Morphologie, Syntax, Semantik (z.B.: Wort, Satz).

Basic Concepts

Note:

By definition, an optimality-theoretic syntax takes the general form in (2), with the grammar divided into a Gen part that creates the competing candidates, and a H-Eval part that selects the optimal candidate(s).

(2) Structure of a competition-based syntax:

- a. Gen creates the candidate set $\{C_1, C_2, \dots\}$.
- b. H-Eval determines the optimal candidate(s) C_i (C_j, \dots) in $\{C_1, C_2, \dots\}$.

Recall:

The notion of optimality in a minimalist syntax is a comparatively simple one: Optimality is determined by a small set of simple transderivational or translocal economy constraints, as in (3).

(3) Optimality in minimalist syntax:

A derivation D_i is optimal iff there is no derivation D_j in the same reference set that is preferred over D_i by a transderivational or translocal constraint.

Optimality

But:

In optimality-theoretic syntax, there is only one translocal or transderivational constraint that determines optimality: Optimal (and grammatical) is a candidate that has the best constraint profile – or, more precisely, a candidate for which there is no competitor that has a better constraint profile; cf. (4).

(Whether this optimality principle is transderivational or translocal depends on whether the competing candidates are assumed to be derivations or output representations.)

(4) **Optimality in optimality-theoretic syntax:**

A candidate C_i is optimal (= grammatical) iff there is no candidate C_j in the same candidate set that has a better constraint profile.

Constraint Profile

(5) Constraint Profile:

A candidate C_j has a better constraint profile than a candidate C_i iff there is a constraint Con_k such that (a) and (b) hold:


- a. C_j satisfies Con_k better than C_i .
- b. There is no constraint Con_l ranked higher than Con_k on which C_i and C_j differ.

Note:

C_j satisfies a constraint Con better than C_i iff C_j has fewer violations of Con . This implies the case that C_i violates Con once (or more often), and C_j does not violate Con at all.

Notation

Note:

This presupposes that in addition to the local (or global) constraints employed by the Gen component, which are inviolable and unranked, the H-Eval component relies on a system of local (or global) constraints that are **violable** and **ranked** (and, by assumption, **universal**) in order to determine the best constraint profile, hence, optimality. The ranking among the violable local constraints of the H-Eval component is indicated by the symbol \gg ; the H-Eval constraints themselves are typically written with small capitals. Optimality-theoretic competitions are often illustrated by tables (so-called **tableaux**); optimality of a candidate is indicated by the **pointing finger**: ; violation of a local constraint is shown by a star * in the appropriate column of the table; if this violation is fatal for a candidate (i.e., responsible for its suboptimality), an exclamation mark ! is added. In the abstract H-Eval competition in table T_1 , in which the candidate set consists of C_1 – C_5 , C_1 emerges as the optimal candidate: It avoids a violation of the high-ranked constraints A and B (unlike C_3 – C_5), and it minimizes a violation of the low-ranked constraint C (unlike C_2). Hence, there is no competing candidate with a better constraint profile than C_1 .

Determining Optimality and Parametrization

T₁: Determining optimality

Candidates	A	B	C
☞ C ₁			*
C ₂			**!
C ₃		*!	
C ₄	*!		
C ₅		*!	*

Constraint reranking = parametrization:

By reranking the constraints B and C in T₁, candidate C₃ would emerge as the optimal candidate. Reranking of constraints forms the basis of the concept of parametrization in optimality-theoretic syntax.

Non-Cumulativity

Non-cumulativity:

A further characteristic feature of this approach is that it is essentially non-cumulative; i.e., no number of violations of a low-ranked constraint can outweigh a single violation of a higher-ranked constraint. Thus, suppose that there were an additional, lowest-ranked constraint D in T_1 that C_1 violates, say, five times, and that C_2 – C_5 do not violate at all. This would not undermine C_1 's optimality.

Candidates and Violable Constraints

Candidates and violable constraints:

Something needs to be said about the nature of candidates and candidate sets.

Optimality-theoretic syntax is strongly influenced by work in optimality-theoretic phonology. Since the latter is characterized by an orientation that is predominantly representational (cf. Prince & Smolensky (1993) and McCarthy & Prince (1995)), it does not come as a surprise that many approaches in optimality-theoretic syntax postulate that the competing candidates created by Gen are output representations. This holds, e.g., for what can arguably be viewed as the three most influential analyses in optimality-theoretic syntax so far (outside the LFG work by Bresnan, Choi, Sells, and others), viz., Grimshaw (1997), Pesetsky (1998), and Legendre, Smolensky & Wilson (1998). However, there is no inherent reason why the candidates that are subject to optimization should not be syntactic objects of a more complex type, like $\langle D\text{-structure}, S\text{-structure}, LF \rangle$ tuples as in the Principles-and-Parameters approach, or, indeed, complete derivations, as in the minimalist program. The choice of candidate type goes hand in hand with the choice of local (or global) constraint type that shows up in the H-Eval part as a violable and ranked: If candidates are representations, constraints will be representational, if candidates are derivations, constraints will be derivational (or global), and if candidates are $\langle D\text{-structure}, S\text{-structure}, LF \rangle$ tuples, constraints can take any of the forms sketched in [2].

Candidate Sets

Candidate sets:

Similarly, candidate sets can be defined in various ways, which of course significantly influences the nature of the competition. Basically, all of the definitions of [reference sets in minimalist syntax](#) that have been proposed (see [5], and Sternefeld (1997) for an overview) are also potential definitions of candidate sets in optimality-theoretic syntax. A further influential definition of candidate sets comes from Grimshaw (1997). She postulates that two candidates (S-structure representations) compete iff they are realizations of the same predicate/argument structure and have non-distinct logical forms (or non-distinct interpretation).

H-Eval vs. Gen

Note:

By making optimality depend on an intricate system of violable and ranked constraints, H-Eval – and hence, the concept of competition – becomes even more important than in minimalist syntax and blocking syntax. As a matter of fact, much work in optimality-theoretic syntax tries to minimize the role of the Gen component, and maximize the role of the H-Eval component (but see Pesetsky (1997; 1998) for some cautionary remarks).

Evidence for Optimization

What is optimality theory good for in syntax?

An optimality-theoretic approach gains immediate support in all those contexts where postulating a competition of syntactic objects is initially plausible. This includes, but is by no means confined to, contexts where notions of economy seem to play a role. A prototypical case is one in which the wellformedness of a sentence S_i that exhibits an otherwise peculiar property seems to depend on the unavailability of another sentence S_j that exhibits the property one would normally expect. Here, S_i is often referred to as a “repair” or “last resort” form; a typical instance is the English **do**-support construction. Accordingly, **do**-support was among the first phenomena to be tackled in optimality-theoretic syntax (see Speas (1995) and Grimshaw (1997)). Most of the constructions discussed in [5] can also be viewed as suggesting an underlying competition; and indeed, they can fruitfully be addressed in optimality-theoretic syntax.

Binding Principles A and B

(6) Consequences of Principle A:

- a. $[_{CP} C [_{TP} John_1 [_{T'} \emptyset [_{VP} t_1 \text{ likes himself}_1]]]]$
 b. $*[_{CP} C [_{TP} John_1 [_{T'} \emptyset [_{VP} t_1 \text{ thinks } [_{CP} \text{ that } [_{TP} Mary_2 [_{T'} \emptyset [_{VP} t_2 \text{ likes himself}_1]]]]]]]]$

(7) Consequences of Principle B:

- a. $*[_{CP} C [_{TP} John_1 [_{T'} \emptyset [_{VP} t_1 \text{ likes him}_1]]]]$
 b. $[_{CP} C [_{TP} John_1 [_{T'} \emptyset [_{VP} t_1 \text{ thinks } [_{CP} \text{ that } [_{TP} Mary_2 [_{T'} \emptyset [_{VP} t_2 \text{ likes him}_1]]]]]]]]$

Generalization:

By and large, pronouns are allowed to express binding relations in English in just those cases in which anaphors are not allowed to do so. This has been taken to indicate that there is a competition of the two strategies: We do not have to stipulate both Principle A and Principle B. Rather, one of the two constraints (usually Principle A) is adopted, and the other is derived as an elsewhere case by invoking syntactic competition (see Fanselow (1989; 1991), Burzio (1991; 1998), Reinhart (1991), and Richards (1997), among others). The following analysis essentially goes back to Wilson (2001).

Binding, OT Constraints

- (8) a. LOC-ANT^r (“Local Antecedent”):
 If a binding domain contains an anaphor, then it must also contain the anaphor’s antecedent.
- b. REF-ECON^r (“Referential Economy”):
 An argument must not have lexical ϕ -feature specification.
- (9) **Ranking in English:**
 LOC-ANT \gg REF-ECON.

Note:

- (i) LOC-ANT is a version of Principle A. Other things being equal, this constraint favours pronouns.
- (ii) REF-ECON inherently prefers anaphors to pronouns, assuming that anaphors do not have a lexical ϕ -feature specification, whereas pronouns do.

Binding Competitions

T₂: Reflexivization

Candidates	LOC-ANT	REF-ECON
☞ C ₁ : John ₁ likes himself ₁		
C ₂ : John ₁ likes him ₁		*!

T₃: Pronominalization

Candidates	LOC-ANT	REF-ECON
C ₁ : John ₁ thinks that Mary likes himself ₁	*!	
☞ C ₂ : John ₁ thinks that Mary likes him ₁		*

Complementizer-Trace Effects

Note:

Recall from [3] that the Principles-and-Parameters approach accounts for complementizer-trace effects (= that-trace effects) on a purely local basis, without postulating a competition with the complementizer-less variant from which only the latter would emerge as optimal.

(10) **The ECP account of that-trace effects:**

- a. *Who₁ do you think [_{CP} t'₁([+γ]) that [_{TP} t₁([-γ]) will leave]] ?
- b. Who₁ do you think [_{CP} t'₁([+γ]) ∅ [_{TP} t₁([+γ]) will leave]] ?

Background assumptions

This view is abandoned in Déprez (1991), which is the basis of the optimality-theoretic account advanced in Grimshaw (1997). As background, Grimshaw assumes that the size of clauses is variable. Clauses are extended projections of V; they are minimally VPs, but they can be TPs, CPs, or functional projections of an even bigger size, depending on the outcome of optimization. Bridge verbs in English permit both CP-embedding (with a complementizer – a declarative CP without a complementizer will typically fatally violate a high-ranked constraint that precludes empty head positions) and TP- or VP-embedding (without a complementizer). In the latter case, TP must be chosen if an auxiliary or **do** is present (i.e., if the need arises to accommodate an additional lexical head); VP otherwise. A co-occurrence of CP- and TP- (VP-) embedding is possible if the two candidates have an identical constraint profile. This implies that the presence of **that** does not violate any constraint, an assumption that Grimshaw (1997) makes even though it is not completely unproblematic (as she acknowledges). The main constraints that are needed in the account of complementizer-trace effects are listed in (11).

Grimshaw's OT Analysis: Constraints

- (11) a. OP-SPEC^r (“Operator in Specifier”):
 Wh-operators must occupy a specifier position from which they c-command all elements of the extended V projection over which they take scope.
- b. T-LEX-GOV^r (“Lexical Government of Traces”):
 A trace is lexically governed.
- c. STAY^r (“Economy of Movement”):
 Trace is not allowed.
- (12) **Ranking in English:**
 OP-SPEC \gg T-LEX-GOV \gg STAY.

STAY is a local representational version of the transderivational economy constraint Fewest Steps; OP-SPEC is a version of the Wh-Criterion that has often been postulated in the Principles-and-Parameters approach (see, e.g., Lasnik & Saito (1992)). (However, we have seen that the effects of such a constraint can be derived from postulating [*Q*] features on C_[+wh] items.)

Assumption:

Candidates with and without **that** compete.

Subject Movement

T₄: Subject wh-movement

Candidates	OP-SPEC	T-LEX-GOV	STAY
C ₁ : ... who ₁ you think [CP that [TP t ₁ will leave]]		*!	*
☞ C ₂ : ... who ₁ you think [TP t ₁ will leave]			*
C ₃ : ... you think [CP that [TP who ₁ will leave]]	*!		
C ₄ : ... you think [TP who ₁ will leave]	*!		

Analysis:

C₃ and C₄ fatally violate OP-SPEC. Both C₁ and C₂ violate STAY, but C₁ violates T-LEX-GOV in addition: t₁ in C₁ is not lexically governed (**that** being unable to do so), whereas t₁ in C₂ is lexically governed (by the matrix V). In contrast, an embedded V governs object traces throughout, irrespective of the presence or absence of a complementizer **that**; hence, T-LEX-GOV is satisfied equally well by C₁ and C₂ in table T₅. Given that C₁ and C₂ do not differ with respect to any other constraint either, optionality of a complementizer is correctly predicted in cases of object extraction, due to an identical constraint profile.

Object Movement

T₅: Object wh-movement

Candidates	OP-SPEC	T-LEX-GOV	STAY
☞ C ₁ : ... who ₁ you think [CP that [TP she will invite t ₁]]			*
☞ C ₂ : ... who ₁ you think [TP she will invite t ₁]			*
C ₃ : ... you think [CP that [TP she will invite who ₁]]	*!		
C ₄ : ... you think [TP she will invite who ₁]	*!		

Adjuncts?

Observation:

Thus far, there is no evidence for treating T-LEX-GOV as a violable constraint in the H-Eval part of the grammar (rather than as an inviolable constraint in the Gen part). Such evidence can be gained by considering adjunct extraction. In this case, T-LEX-GOV is violated by both candidates involving **wh**-movement (C_1 , C_2). However, given that there is no competing candidate that can avoid a violation of T-LEX-GOV without violating a higher-ranked constraint (e.g., a candidate that employs a resumptive pronoun; see below), C_1 and C_2 can emerge as optimal despite this violation.

Adjunct Movement

T₆: Adjunct wh-movement

Candidates	OP-SPEC	T-LEX-GOV	STAY
☞ C ₁ : ... why ₁ you think [CP that [TP she has left t ₁]]		*	*
☞ C ₂ : ... why ₁ you think [TP she has left t ₁]		*	*
C ₃ : ... you think [CP that [TP she has left why ₁]]	*!		
C ₄ : ... you think [TP she has left why ₁]	*!		

Subjacency and Resumptive Pronouns

(13) Resumptive pronouns in Complex NP Constraint contexts:

- a. the man [_{CP} who(m)₁ I saw t₁]
- b. *the man [_{CP} who(m)₁ I don't believe [_{DP} the claim [_{CP} t'₁ that anyone saw t₁]]]
- c. *the man [_{CP} who(m)₁ I saw him₁]
- d. ?the man [_{CP} who(m)₁ I don't believe [_{DP} the claim [_{CP} t'₁ that anyone saw him₁]]]

Resumptive pronouns often seem to be possible only as last resort strategies, in cases where traces are blocked. Competition-free approaches to syntax have no obvious means to relate one construction to the other; but the case is different in optimality-theoretic syntax. An optimality-theoretic account of resumptive pronoun strategies is developed in Legendre, Wilson & Smolensky (1998) (on the basis of evidence from Chinese) and Pesetsky (1998) (on the basis of English data comparable to those in (13), as well as evidence from Hebrew, Russian, and Polish). (A similar last resort analysis is given in Hornstein (2000) within the minimalist program.) The details of the analyses differ a lot, but the gist of the explanation is identical; it centers around two constraints like those in (14).

Resumptive Pronouns: OT Analysis

- (14) a. $CNPC^r$ (“Complex NP Constraint”):
 $* \dots \alpha_1 \dots [DP \dots [CP \dots t_1 \dots]]$...
- b. RES^r (“Resumptive Pronoun Constraint”):
 Resumptive pronouns are prohibited.
- (15) **Ranking in English:**
 $CNPC \gg RES$.

Note:

The $CNPC$ prohibits traces in certain (non-local) environments; RES disfavors resumptive pronouns (i.e., pronouns that have c-commanding co-indexed antecedent in an A-bar position) in general. As with the **wh**-movement construction discussed in the last section, it must be ensured that overt movement of the relative operator takes place in examples like those in (13). Suppose that this is independently taken care of – e.g., by a high-ranked (or inviolable Gen-) constraint demanding deletion of [$*rel^*$].

Resumptive Pronouns: Transparent Contexts

T₇: Trace vs. resumptive pronoun in transparent contexts

Candidates	CNPC	RES
☞ C ₁ : the man [CP who(m) ₁ I saw t ₁]		
C ₂ : the man [CP who(m) ₁ I saw him ₁]		*!

Analysis:

Both candidates respect CNPC. Consequently, the RES violation incurred by the resumptive pronoun in C₂ becomes fatal, and C₁ is optimal. However, in the competition illustrated in table T₈, C₁ violates the CNPC. In this case, C₂'s RES violation is tolerable, and the resumptive pronoun strategy emerges as optimal.

Resumptive Pronouns: Opaque Contexts

T₈: Trace vs. resumptive pronoun in CNPC contexts

Candidates	CNPC	RES
C ₁ : the man [CP who(m) ₁ [TP I don't believe [DP the claim [CP t' ₁ that anyone saw t ₁]]]]	*!	
☞ C ₂ : the man [CP who(m) ₁ [TP I don't believe [DP the claim [CP that anyone saw him ₁]]]]		*

Avoid Pronoun

Recall:

As noted in [3], PRO may block overt pronouns in English gerunds. PRO and a lexical pronoun can both occur in principle; however, PRO must be used instead of a lexical pronoun if it can fulfill the Constraint on Control. This was derived by adopting a transderivational/translocal Avoid Pronoun constraint in Chomsky (1981).

(16) PRO in English gerunds:

- a. John₁ would much prefer [PRO₁ going to the movie]
- b. *John₁ would much prefer [PRO_{2/arb} going to the movie]

(17) Pronouns in English gerunds:

- a. *John₁ would much prefer [his₁ going to the movie]
- b. John₁ would much prefer [his₂ going to the movie]

Constraints

- (18) CONTROL^r (“Constraint on Control”, Manzini (1983)):
If PRO is minimally dominated by a declarative clausal complement α , then it must be controlled by an antecedent within the minimal CP that dominates α .
- (19) **Avoid Pronoun**^{td/tl}:
If two derivations D_1 and D_2 are in the same reference set and D_1 uses a lexical pronoun where D_2 uses an empty pronoun, then D_1 is to be preferred over D_2 .

OT Analysis

Note:

Chomsky's (1981) approach can straightforwardly be translated into optimality theory:

- 1 The Constraint on Control in (18) can directly be viewed as an optimality-theoretic constraint.
- 2 The Avoid Pronoun constraint in (19) can be simplified by turning this transderivational/translocal constraint into a local (though violable) one; cf. (20).

(20) *PRON^r ("Avoid Pronoun"):
Pronouns are prohibited.

(21) **Ranking in English:**
CONTROL \gg *PRON.

Competition: Co-Indexing

Candidate sets:

Suppose that candidate sets are defined in such a way that candidates with PRO and candidates with a lexical pronoun can compete, but, crucially, that sentences with different indexings (hence, different logical forms) do not compete. Then, the facts fall into place. The blocking of a lexical pronoun by PRO in cases where CONTROL can be satisfied is illustrated in table T_9 .

T_9 : *PRO vs. pronoun under co-indexing*

Candidates	CONTROL	*PRON
C ₁ : John ₁ would much prefer [his ₁ going to the movie]		*!
☞ C ₂ : John ₁ would much prefer [PRO ₁ going to the movie]		

Competition: Contra-Indexing

But:

Table T₁₀ illustrates the case where PRO is not co-indexed with the matrix antecedent, thereby violating CONTROL. Here, the *PRON violation incurred by all pronouns is non-fatal, and the pronoun strategy is optimal.

Competition: Contra-Indexing

But:

Table T_{10} illustrates the case where PRO is not co-indexed with the matrix antecedent, thereby violating CONTROL. Here, the *PRON violation incurred by all pronouns is non-fatal, and the pronoun strategy is optimal.

T_{10} : *PRO vs. pronoun under contra-indexing*

Candidates	CONTROL	*PRON
☞ C_1 : John ₁ would much prefer [his ₂ going to the movie]		*
C_2 : John ₁ would much prefer [PRO ₂ going to the movie]	*!	

Complexity

Note:

Optimality-theoretic syntax inherits the **complexity** problem from minimalist syntax. Candidate sets are typically large (as in Pesetsky (1998)), often infinite (as in Grimshaw (1997)). In addition, there are several open issues that are specific to the optimality-theoretic approach. The focus is on two of these in what follows: inputs and absolute ungrammaticality.

The Input and Its Functions

An important optimality-theoretic concept that has played no role so far is the notion of **input**. In optimality theory (cf. Prince & Smolensky (1993)), Gen does not create competing candidates (the **outputs**) freely; rather, it does so on the basis of a given input. In phonology, inputs are underlying representations stored in the lexicon; here, inputs qualify as roughly the same types of objects as outputs. In syntax, it is much less clear what the input might look like (see Archangeli & Langendoen (1996)). The null hypothesis – that the input is a completely articulated potential sentence of the same type as the output candidates – is not unproblematic because it would seem to imply the assumption that all possible sentences are “stored,” which cannot possibly be true. **To find out what the input in syntax is, it is instructive to consider its theory-internal functions.** By and large, there are two.

- First, the input is standardly taken to define the competition, i.e., **candidate sets**.
- Second, the input serves as a basis for **faithfulness** constraints that demand input/output identity and thereby minimize deviations from the input in the optimal output.

Let us consider the second function first. Faithfulness constraints can be viewed as global constraints because they cannot be checked by considering solely an output representation or a derivational step.

Faithfulness Constraints 1

Faithfulness constraints:

Faithfulness constraints play an important role in phonology. Constraints of the PARSE (or MAX) family prohibit deletion of input material in the output; constraints of the FILL (or DEP) family prohibit insertion of output material that is not part of the input; constraints of the IDENT family prohibit modifying input material. Faithfulness constraints are also adopted in much recent work in optimality-theoretic syntax. The following two constraints are taken from Legendre, Smolensky & Wilson (1998) and Baković & Keer (2001), respectively.

(22) Syntactic faithfulness constraints:

- a. PARSE[SCOPE]^g:
Scope assignment in the input must be realized by chain formation in the output.
- b. FAITH[COMP]^g:
The output value of [\pm COMP] is the same as the input value.

Faithfulness Constraints 2

Note:

(22-a) implies that the input is a more complex object than just a lexical array or a predicate/argument structure; it must be a highly structured representation that exhibits the relative scope of operators. (22-b) presupposes an abstract feature $[\pm\text{COMP}]$ that for present purposes we can assume to be located on a V that selects a proposition. Let us consider candidates that violate these constraints. Suppose that (23-a) is the input for output candidate (23-b), and (23-c) is the input for output candidate (23-d). Legendre, Smolensky & Wilson (1998) assume that (23-b) violates $\text{PARSE}[\text{SCOPE}]$ because matrix scope for **how**₁ in the input (23-a) (indicated by $[\text{+wh}]_1$) is reduced to embedded scope in the output (again indicated by $[\text{+wh}]_1$). Similarly, Baković & Keer (2001) assume that (23-d) violates $\text{FAITH}[\text{COMP}]$ because a $[-\text{COMP}]$ specification in the input contrasts with a $[\text{+COMP}]$ specification (hence, a complementizer) in the output.

(23) Violations of syntactic faithfulness constraints:

- a. $[\text{+wh}]_1 \dots \text{wonder}_{[\text{+wh}]} [[\text{+wh}]_2 \dots \text{what}_2 \dots \text{how}_1 \dots]$ (input)
- b. You wonder_[\text{+wh}] [_{CP} $[\text{+wh}]_1$ $[\text{+wh}]_2$ how₁ John did what₂] (output)
- c. $\dots V_{[-\text{Comp}]} [\dots]$ (input)
- d. I think [_{CP} that [_{PP} on him]₁ no coat looks good t₁] (output)

Faithfulness Constraints 3

A reinterpretation of faithfulness constraints:

At this point, we need not go into the actual analyses in which these constraints play a role (as it happens, both faithfulness violations turn out to be non-fatal, i.e., (23-bd) are optimal). The crucial question is: Is it really necessary to refer to the concept of input here, or is it possible to read the respective violations off the output forms, without any reference to inputs? At least for the cases at hand, the answer is straightforward: By enriching output representations in ways that have independently been proposed, a reference to inputs becomes unnecessary. (23-ab) is a case where the intended matrix scope is not reached by chain formation in the candidate. Employing **abstract scope markers** (Σ) in S-structure representations (cf., e.g., Williams (1986)), we can equivalently encode this input information in the output, as in (24-a). As for the case in (23-cd), the only assumption that we have to make (which is completely standard) is that selectional properties of lexical heads are accessible in syntax; cf. (24-b).

Faithfulness Constraints 4

(24) Violations of syntactic markedness constraints:

- a. Σ_1 you wonder_[+wh] [_{CP} [+wh]₁ [+wh]₂ how₁ John did what₂] (output)
- b. I think_[-Comp] [_{CP} that [_{PP} on him]₁ no coat looks good t₁] (output)

PARSE[SCOPE] and FAITH[COMP] can now be modified in obvious ways, without reference to inputs, as regular local (so-called *markedness*) constraints.

(25) Faithfulness constraints as markedness constraints:

- a. PARSE[SCOPE]^r (revised):
Scope markers must be reached by chain formation.
- b. FAITH[COMP]^r (revised):
Lexical [\pm COMP] selection requirements must be respected.

Faithfulness Constraints 5

Note:

If this result can be generalized, and all syntactic faithfulness constraints can be reanalyzed in this way, we can conclude that these constraints do not support the concept of input anymore. Why should it be that the notion of input is relevant for phonological faithfulness constraints, but not for their syntactic counterparts? The answer follows from what appears to be a fundamental difference between syntax and phonology: Syntax is an **information-preserving system** with richly structured output candidates, whereas phonology is a system that is standardly taken to lose information in the course of a derivation (e.g., from input to output), so that reference to an underlying input is necessary in constraints.

Candidate Sets

Candidate Sets:

With this in mind, consider the other input function noted above, that of defining candidate sets. Since syntactic output candidates are richly structured, all the relevant information that they must share in order to compete can be read off them, independently of what notion of candidate set is adopted; again, this is in sharp distinction to phonology. Hence, reference to inputs seems unnecessary for this purpose in syntax.

Conclusion:

From all this, it seems that one can conclude that it may eventually be possible to dispense with the notion of input in syntax; but further research is needed in this domain (see Heck et al. (2002)).

Absolute Ungrammaticality

Observation:

Another important open question in optimality-theoretic syntax is how to account for the phenomenon of **absolute ungrammaticality** or **ineffability**, i.e., cases where there does not seem to be a candidate in a candidate set that is grammatical. As an example, consider the following ungrammatical example involving **wh**-extraction across an adjunct island in German:

(26) A consequence of the Adjunct Condition:

*Was₁ ist Fritz eingeschlafen [CP nachdem er t₁ gelesen hat] ?
 what is Fritz fallen asleep after he read has

(27) Adjunct Condition^d:

Movement must not take place from an XP that has been merged without a deletion of selectional features.

(28) ADJ-CON^r (“Adjunct Condition”, representational version):

A trace must not be separated by a non-selected XP from its antecedent.

Absolute Ungrammaticality: Gen

Approach no. 1: Gen:

Pesetsky (1997; 1998) emphasizes that certain sentences may be ungrammatical not because they are classified as suboptimal in the H-Eval part of the grammar, but because they cannot be generated by Gen in the first place. Thus, a constraint like (28) might be part of Gen.

Absolute Ungrammaticality: Semantic Illformedness

Approach no. 2: semantic illformedness:

It is suggested in Grimshaw (1994) and Müller (1997) that certain optimal candidates may have properties that make them inaccessible for other domains of the language faculty like, e.g., semantic interpretation. ADJ-CON might be part of H-Eval, but ranked higher than OP-SPEC. On this view, (29) could block (28) as suboptimal; but this optimal candidate would be uninterpretable (indicated by #) and, hence, unusable.

(29) Semantic illformedness:

#Fritz ist eingeschlafen [CP nachdem er was₁ gelesen hat] ?
 Fritz is fallen asleep after he what read has

Note:

These two approaches have in common that they allow the possibility that absolute ungrammaticality is not located in the H-Eval component of grammar; but in a component that precedes (Gen) or follows (interpretation) optimization. If, however, H-Eval is to be held responsible for the ungrammaticality of (26), there must be a competing candidate with a better constraint profile that blocks it.

Absolute Ungrammaticality: Last Resort

Approach no. 3: last resort (fails):

A priori, this might be a candidate that employs a resumptive pronoun strategy which is only legitimate in this context, as a last resort. If this were so, the ineffability problem would be spurious in the case at hand. However, (30) shows that the resumptive pronoun strategy is not an option in German (a constraint like RES must outrank ADJ-CON and other locality constraints in German):

(30) Last resort:

*Was₁ ist Fritz eingeschlafen [_{CP} nachdem er es₁ gelesen hat] ?
 what is Fritz fallen asleep after he it read has

Absolute Ungrammaticality: Null Parse

Approach no. 4: null parse:

What, then, could the optimal candidate blocking (26) look like? Following Prince & Smolensky (1993), Ackema & Neeleman (1998) propose that the empty candidate \emptyset (the “null parse”) is part of every candidate set. This candidate violates the constraint in (31), which is typically ranked high.

- (31) $*\emptyset^r$ (“Avoid Null Parse”):
 \emptyset is prohibited.

Absolute Ungrammaticality: Null Parse cont'd

Note:

Constraints that are ranked higher than $*\emptyset$ in effect become inviolable (given that there is no other constraint except $*\emptyset$ that \emptyset can violate). In this sense, $*\emptyset$ introduces a dividing line into rankings. Thus, if both ADJ-CON and the constraint that triggers **wh**-movement (e.g., OP-SPEC) outrank $*\emptyset$, adjunct islands become inviolable. This is shown in table T₁₁.

T₁₁: Adjunct islands and the null parse

Candidates	ADJ-CON	OP-SPEC	$*\emptyset$
C ₁ : was ₁ ... [CP nachdem er t ₁ V]	*!		
C ₂ : – ... [CP nachdem er was ₁ V]		*!	
☞ C ₃ : \emptyset			*

Absolute Ungrammaticality: Neutralization 1

Approach no. 5: neutralization:

A final possibility to be discussed here is the **neutralization** approach to absolute ungrammaticality in syntax. Such an approach has been adopted by Legendre, Smolensky & Wilson (1998), Schmid (1998), Baković & Keer (2001), and Wilson (2001), among others. For the present case, a neutralization analysis might posit that the optimal candidate blocking (26) is (32).

(32) Neutralization:

Fritz ist eingeschlafen	[_{CP}	nachdem er	was ₁	gelesen hat]
Fritz is fallen asleep		after	he something	read	has

Note:

The crucial difference to (29) is that **was**₁ is turned into an indefinite pronoun, and the matrix $C_{[+wh]}$ is turned into a $C_{[-wh]}$. Thus, there is a feature change from [+wh] in (26) to [-wh] in (32), and the sentence is interpreted as declarative, rather than a question. If (32) is to block (26) as suboptimal, this presupposes that candidates that differ in their **wh**-feature specification can compete. But then, the problem arises that we would also wrongly expect one of the sentences in (33) to block the other.

Absolute Ungrammaticality: Neutralization 2

(33) A problem?

- a. Was₁ hat er t₁ gelesen ?
 what has he read
- b. Er hat was₁ gelesen
 he has something read

The solution:

The neutralization approach solves this problem as follows. The $[\pm wh]$ -specification is unambiguously specified in the input; an input with a $[+wh]$ specification on some item and a minimally different input with a $[-wh]$ specification count as different, and define different candidate sets. The important assumption is that there is a faithfulness constraint that demands preservation of the $[\pm wh]$ feature specification in the output:

(34) FAITH[WH]^g:

The output value of $[\pm wh]$ is the same as the input value.

Absolute Ungrammaticality: Neutralization 3

Analysis:

Suppose now that *ADJ-CON* and *OP-SPEC* are ranked higher than *FAITH[WH]*. Then, (32) will have a better constraint profile than (26) both in the competition that has a *[-wh]* specification in the input, and in the competition that has a *[+wh]* specification in the input. Thus, there is a “neutralization” of different input specifications in the output. This is shown in tables *T₁₂* and *T₁₃*.

Absolute Ungrammaticality: Neutralization 4

T_{12} : Adjunct islands and neutralization; [-wh] in the input

Candidates	ADJ-CON	OP-SPEC	FAITH[WH]
C_1 : was _{1[+w]} ... [CP nachdem er t ₁ V]	*!		*
C_2 : - ... [CP nachdem er was _{1[+w]} V]		*!	*
☞ C_3 : - ... [CP nachdem er was _{1[-w]} V]			

T_{13} : Adjunct islands and neutralization; [+wh] in the input

Candidates	ADJ-CON	OP-SPEC	FAITH[WH]
C_1 : was _{1[+w]} ... [CP nachdem er t ₁ V]	*!		
C_2 : - ... [CP nachdem er was _{1[+w]} V]		*!	
☞ C_3 : - ... [CP nachdem er was _{1[-w]} V]			*

Absolute Ungrammaticality: Neutralization 5

Note:

In transparent contexts, where movement may occur without a violation of a high-ranked locality constraint like ADJ-CON (cf. (33)), FAITH[WH] violations become fatal, and the candidate that maintains the $[\pm wh]$ specification of the input emerges as optimal.

Absolute Ungrammaticality: Conclusion

Of the four working approaches to absolute ungrammaticality discussed here (Gen, semantic illformedness, null parse, neutralization), the neutralization approach is arguably the most elegant one. Still, it is not without problems. One conspicuous peculiarity is that neutralization creates **massive derivational ambiguity**. A well-formed sentence like (32) can have different “histories,” being an optimal candidate in two candidate sets with different inputs. This vacuous ambiguity may be considered problematic from the point of view of language acquisition and parsing; and it can only be avoided by additional meta-optimization procedures that compare the competitions in T_{12} and T_{13} ; cf. the notion of **input optimization** in Prince & Smolensky (1993) (called **lexicon optimization** in phonology).

General Conclusion

Note:

This does not exhaust the list of open issues that are currently under debate in optimality-theoretic syntax. Here are some others:

- 1 Optionality
- 2 Degrees of Grammaticality
- 3 Cumulativity
- 4 Parametrization
- 5 Serial vs. parallel optimization

Outlook:

Do optimality-theoretic and minimalist approaches differ? Yes, they do, but not as much as one might think at first sight, and in comparing these two approaches, it is important to identify issues that are orthogonal (e.g., derivational vs. representational constraints/candidates, and the issue of transderivational/translocal constraints).

Optionalität

Beobachtung:

Eigentlich sollte doch immer nur ein Kandidat der beste sein (auch wenn die Definition von Optimalität in (4) im Prinzip mehrere Gewinner in einem einzigen Wettbewerb zulässt). Was macht man dann aber mit Situationen, wo es so aussieht, als könnten mehrere Kandidaten optimal nebeneinander bestehen?

Optionalität

Beobachtung:

Eigentlich sollte doch immer nur ein Kandidat der beste sein (auch wenn die Definition von Optimalität in (4) im Prinzip mehrere Gewinner in einem einzigen Wettbewerb zulässt). Was macht man dann aber mit Situationen, wo es so aussieht, als könnten mehrere Kandidaten optimal nebeneinander bestehen?

(35) Komplementierer-Tilgung im Englischen:

- a. I think – John will leave
- b. I think that John will leave

Optionalität

Beobachtung:

Eigentlich sollte doch immer nur ein Kandidat der beste sein (auch wenn die Definition von Optimalität in (4) im Prinzip mehrere Gewinner in einem einzigen Wettbewerb zulässt). Was macht man dann aber mit Situationen, wo es so aussieht, als könnten mehrere Kandidaten optimal nebeneinander bestehen?

(35) Komplementierer-Tilgung im Englischen:

- a. I think – John will leave
- b. I think that John will leave

(36) Partielle W-Bewegung im Deutschen:

- a. Wen glaubst du [dass man t einladen sollte] ?
- b. Was glaubst du [wen man t einladen sollte] ?

Weitere Beispiele für Optionalität

(37) **W-Bewegung im Französischen:**

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

Weitere Beispiele für Optionalität

(37) **W-Bewegung im Französischen:**

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

(38) **Extraposition im Deutschen:**

- a. dass eine Frau [die ich mag] zur Tür reingekommen ist
- b. dass eine Frau t zur Tür reingekommen ist [die ich mag]

Weitere Beispiele für Optionalität

(37) **W-Bewegung im Französischen:**

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

(38) **Extraposition im Deutschen:**

- a. dass eine Frau [die ich mag] zur Tür reingekommen ist
- b. dass eine Frau t zur Tür reingekommen ist [die ich mag]

(39) **Freie Wortstellung im Deutschen:**

- a. dass keiner den Fritz gesehen hat
- b. dass den Fritz keiner gesehen hat

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) **Optionalität zweier Kandidaten K_i , K_j :**

a. **Pseudo-Optionalität:**

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) **Optionalität zweier Kandidaten K_i , K_j :**

a. **Pseudo-Optionalität:**

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. **Echte Optionalität:**

K_i , K_j haben ein identisches Beschränkungsprofil.

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

(ii) Lokal hierarchische Kopplung

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

(ii) Lokal hierarchische Kopplung

(iii) Lokal konjunktive Kopplung

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

(ii) Lokal hierarchische Kopplung

(iii) Lokal konjunktive Kopplung

(iv) Lokal disjunktive Kopplung

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

(ii) Lokal hierarchische Kopplung

(iii) Lokal konjunktive Kopplung

(iv) Lokal disjunktive Kopplung

(v) Global disjunktive Kopplung

Verfahren zur Ableitung von Optionalität in der Optimalitätstheorie

(40) Optionalität zweier Kandidaten K_i , K_j :

a. Pseudo-Optionalität:

K_i , K_j gehören zu verschiedenen Kandidatenmengen und interagieren nicht.

b. Echte Optionalität:

K_i , K_j haben ein identisches Beschränkungsprofil.

c. Kopplung ('tie'):

K_i , K_j unterscheiden sich bei zwei Beschränkungen A, B; diese Beschränkungen sind gekoppelt ("tied").

(i) Global hierarchische Kopplung

(ii) Lokal hierarchische Kopplung

(iii) Lokal konjunktive Kopplung

(iv) Lokal disjunktive Kopplung

(v) Global disjunktive Kopplung

d. Neutralisierung:

K_i , K_j gehören zu verschiedenen Kandidatenmengen, interagieren aber.

e. Stochastische Optimalitätstheorie

Pseudo-Optionalität

Annahme:

Kandidatenmengen werden so definiert, dass weniger Wettbewerb besteht.

Problem:

Zu wenig Wettbewerb schwächt die Theorie und verstärkt die Unaussprechbarkeitsproblematik.

(41) Wettbewerb von partieller und langer W-Bewegung im Deutschen:

- a. ?Wen glaubst du nicht [dass man t einladen sollte] ?
- b. *Was glaubst du nicht [wen man t einladen sollte] ?

(42) Wettbewerb von W-Bewegung und W-in-situ im Französischen:

- a. Je me demande [qui C tu as vu t]
ich frage mich wen du hast gesehen
- b. *Je me demande [– (que) tu as vu qui]
ich frage mich dass du hast gesehen wen

Echte Optionalität

Annahme:

Zwei (oder mehr) Kandidaten können tatsächlich dasselbe (optimale) Beschränkungsprofil haben.

Echte Optionalität

Annahme:

Zwei (oder mehr) Kandidaten können tatsächlich dasselbe (optimale) Beschränkungsprofil haben.

Problem:

Dies ist praktisch unmöglich, z.B. aufgrund der Existenz von Treuebeschränkungen.

Kopplungen

Annahme:

Zwei (oder mehr) Beschränkungen sind gleich wichtig, d.h., gekoppelt ("tied"). Kandidaten, die sich nur bei diesen Beschränkungen unterscheiden, können alle optimal sein.

T_{14} : Kopplung von Beschränkungen: $B \circ C$

	A	B	C	D
☞ O ₁			*	
☞ O ₂		*		
O ₃		*(!)	*(!)	
O ₄	*!			

Typen von Kopplungen

Bemerkung:

Es existieren viele unterschiedliche (und inkompatible) Kopplungskonzepte in der Literatur. Eine Grundunterscheidung ist die zwischen **globalen** und **lokalen** Kopplungskonzepten. Globale Kopplungen sind Abkürzungen für die simultane Präsenz verschiedener Beschränkungsordnungen in einer Sprache; lokale Kopplungen sind spezielle Beschränkungstypen. Am weitesten verbreitet sind wohl die global hierarchische Kopplung und die lokal konjunktive Kopplung.

Beobachtung:

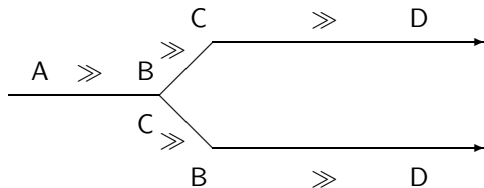
Bei globalen Kopplungen können sich die optimalen Kandidaten unterhalb der Kopplung unterschiedlich verhalten; bei lokalen Kopplungen führt unterschiedliches Verhalten bei tieferen Beschränkungen zum Zusammenbruch der Optionalität.

Global hierarchische Kopplungen

(43) Global hierarchische Kopplung:

Es sei $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine partielle Beschränkungsordnung der Sprache L und B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$. Dann gilt: Für jede mögliche Subordnung O der Beschränkungen in B_i ist auch Γ_O eine Beschränkungsordnung der Sprache L, wobei Γ_O sich von Γ nur dadurch unterscheidet, dass B_i durch O ersetzt ist.

(44) Diagramm der global hierarchischen Kopplung $B \circ C$

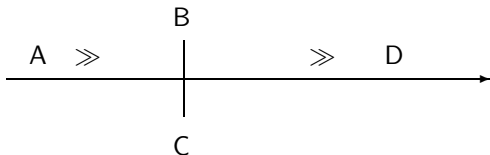


Lokal konjunktive Kopplungen

(45) Lokal konjunktive Kopplung:

Es sei $\Gamma = \langle B_1 \gg \dots \gg B_n \rangle$ eine vollständige Beschränkungsordnung der Sprache L und B_i ($1 \leq i \leq n$) $\in \Gamma$ eine Kopplung $B_{i_1} \circ \dots \circ B_{i_n}$. Dann gilt: B_i ist von einem Kandidaten K verletzt gdw. es eine Beschränkung B_{i_j} in der Kopplung B_i gibt, die K verletzt.

(46) Diagramm der lokal konjunktiven Kopplung $B \circ C$



Stochastische Optimalitätstheorie

Beobachtung:

Häufig sind die an der Optionalität beteiligten Konstruktionen nicht gleich frequent bzw. unmarkiert bzw. wohlgeformt → Grade der Akzeptabilität.

(47) Präferenzen bei Optionalität der Position englischer Possessoren:

- 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

Terminologie:

SpecN = Position links von N; CompN = Position rechts von N.

Skalen

Beobachtung:

Es sind unabhängig funktional Belebtheits- und Definitheitsskalen motiviert. Diese spielen eine Rolle in Beschränkungen.

(48) Beschränkungssubhierarchien qua Belebtheits- und Definitheitsskalen:

- a. *SpecN/unbelebt \gg *SpecN/belebt \gg *SpecN/menschlich
- b. *CompN/menschlich \gg *CompN/belebt \gg CompN/unbelebt
- c. *SpecN/indefinit \gg *SpecN/definit \gg *SpecN/Name \gg *SpecN/Pron
- d. *CompN/Pron \gg *CompN/Name \gg *CompN/definit \gg
*CompN/indefinit

Überlappung der Bereiche von Beschränkungen

Idee:

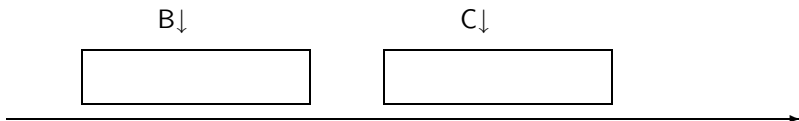
Beschränkungen sind nicht notwendigerweise fest, kategorial zueinander geordnet. Vielmehr können ihre Geltungsbereiche überlappen. Bei Überlappung entsteht Optionalität.

Überlappung der Bereiche von Beschränkungen

Idee:

Beschränkungen sind nicht notwendigerweise fest, kategorial zueinander geordnet. Vielmehr können ihre Geltungsbereiche überlappen. Bei Überlappung entsteht Optionalität.

(49) Kategoriale Ordnung von Geltungsbereichen von Beschränkungen:

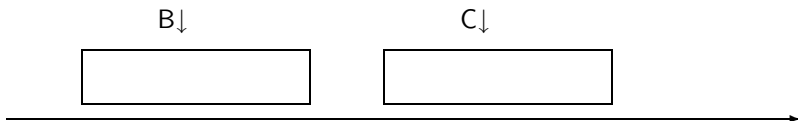


Überlappung der Bereiche von Beschränkungen

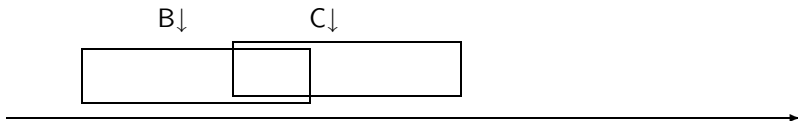
Idee:

Beschränkungen sind nicht notwendigerweise fest, kategorial zueinander geordnet. Vielmehr können ihre Geltungsbereiche überlappen. Bei Überlappung entsteht Optionalität.

(49) Kategoriale Ordnung von Geltungsbereichen von Beschränkungen:



(50) Überlappende Ordnung von Geltungsbereichen von Beschränkungen:



Mechanismus

Annahme:

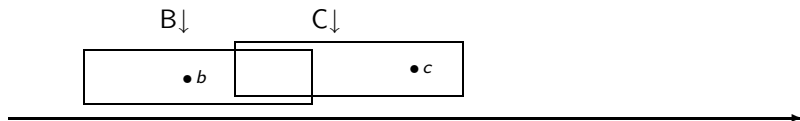
Ein Kandidat wird zu einer **Evaluationszeit** bewertet; er ist grammatisch, wenn er dabei optimal ist. Für eine Evaluation wird im Geltungsbereich einer Beschränkung ein beliebiger Punkt gewählt. Eine Beschränkung B ist zu einer Evaluationszeit höher geordnet als eine andere Beschränkung C, wenn der für B gewählte Punkt über dem für C gewählten liegt. Sind die Bereiche von B und C kategorisch geordnet, liegt der Punkt von B immer über dem von C: keine Optionalität. Sind die Bereiche von B und C dagegen überlappend, entsteht Optionalität, je nachdem, ob der Punkt von B über dem von C liegt oder umgekehrt. (Dies entspricht im Prinzip der globalen Kopplung.)

Erfassung von Präferenzen:

Punktwahl zum Evaluationszeitpunkt ist an sich frei; aber je kleiner der gemeinsame Bereich von B und C ist, desto wahrscheinlicher ist es, dass der für die höhere Beschränkung, z.B. B, gewählte Punkt über dem für C gewählten liegt. Je wahrscheinlicher die höhere Position der B-Punkte über den C-Punkten zu einer gegebenen Evaluationszeit ist, desto stärker ist dann die von B favorisierte Konstruktion in einem Optionalitätspaar gegenüber der von C favorisierten Konstruktion präferiert (und desto frequenter ist sie in **Korpora**).

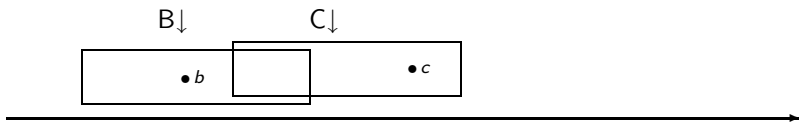
Typische und untypische Ergebnisse

(51) **Typisches Ergebnis:** $B \gg C$

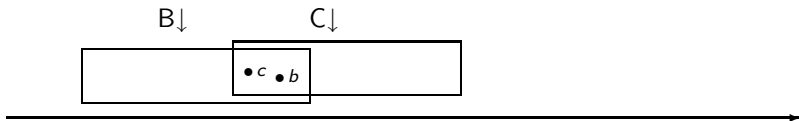


Typische und untypische Ergebnisse

(51) **Typisches Ergebnis:** $B \gg C$



(52) **Seltenes Ergebnis:** $C \gg B$



Die Lokalitätstheorie von Legendre, Smolensky & Wilson (1998)

Lit.:

Legendre, Géraldine, Paul Smolensky & Colin Wilson. 1998. When is Less More? Faithfulness and Minimal Links in Wh-Chains. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 249-289. Cambridge, Mass.: MIT Press & MITWPL.

Ziel:

Entwicklung einer optimalitätstheoretischen Lokalitätstheorie für Bewegung, die gleichzeitig flexibel und restriktiv ist.

Wichtige theoretische Konzepte:

- Neutralisierung
- lokale Konjunktion von Beschränkungen

Grundannahmen

- (53) *Kandidatenmenge* (Legendre, Smolensky & Wilson (1998, 257)):
 Zwei Kandidaten K_i , K_j sind in derselben Kandidatenmenge gdw. gilt:
- a. K_i und K_j sind Realisierungen von identischen Prädikat-/Argument-Strukturen.
 K_i und K_j zielen auf identische LFs.

Bemerkung:

Der Wettbewerb kann rein über Input-Identität festgelegt werden: Im Input liegt eine Prädikat-/Argument-Struktur vor, für die eindeutig eine LF vorgegeben ist. Die (von den Basispositionen abweichenden) Zielpositionen für LF-Interpretation sind dabei durch abstrakte Skopusmarker festgelegt. Legendre et al. nennen ihren komplexen syntaktischen Input [Index](#).

Fallstudie: W-Extraktion im Chinesischen

Annahme:

W-in-situ-Sprachen haben S-strukturelle W-Bewegung. Der einzig relevante Unterschied zu W-Bewegungs-Sprachen ist, dass nicht die erste Spur (= Kopie) ausgesprochen wird, sondern die letzte.

(54) *Lange Adjunktbewegung aus deklarativen Sätzen:*

- a. Ni renwei [_{CP} Lisi yinggai zenmeyang chuli zhe-jian
 du denkst Lisi sollte wie behandeln diese-CL
 shi] ?
 Angelegenheit
- b. "Wie₁ denkst du, dass Lisi diese Angelegenheit t₁ behandeln sollte?"

(55) *Lange Adjunktbewegung aus eingebetteten Fragesätzen:*

- a. Ni xiang-zhidao [_{CP} shei zenmeyang chuli zhe-jian shi] ?
 du fragst dich wer wie behandeln diese-CL Angelegenheit
- b. “*Wie₁ fragst du dich, wer diese Angelegenheit t₁ behandelt hat?”
- c. “Du fragst dich, wer diese Angelegenheit wie behandelt hat.”

- (56) a. **SEL** (“Selektion”, “Select”):
Lexikalisch markierte Selektion muß respektiert werden.
- b. **BAR^{2[-ref]}** (“Zwei-Barrieren-Bedingung für Adjunktketten”):
Ein Kettenglied einer nicht-referentiellen (Adjunkt-) Kette darf nicht zwei Barrieren überkreuzen.
- c. **SKOP-TREUE** (“Skopus-Treue”, “Parse-Scope”):
Ein im Index markierter Skopus muß durch S-strukturelle Kettenbildung in diese Position hinein realisiert werden.
- (57) **Barriere** (Chomsky (1986)):
Eine XP ist eine Barriere gdw. sie nicht L-markiert ist.
- (58) **L-Markierung** (Chomsky (1986)):
 α L-markiert β gdw. gilt:
- α ist eine lexikalische X^0 -Kategorie.
 - α θ -markiert β .
 - β ist Schwester von α .

Konsequenz:

VP und IP sind Barrieren.

(Dies ist zunächst einmal auch in Chomsky (1986) so. Diese Konsequenz ist dort aber unerwünscht und wird durch Zusatzstipulationen zunichte gemacht.)

(59) *Ordnung im Chinesischen:*

SEL \gg BAR^{2[-ref]} \gg SKOP-TREUE

T₁₅: W-Inseln für weiten Adjunktskopus im Chinesischen

Kandidaten	SEL	BAR ² _[-ref]	SKOP-TREUE
K ₁ : [S Q ₁] ... V _[+w] [CP t' ₁ ... wie ₁ ...]		*!	
☞ K ₂ : [S -] ... V _[+w] [CP Q ₁ ... wie ₁ ...]			*

T₁₆: Kurzer Adjunktskopus im Chinesischen: Neutralisierung

Kandidaten	SEL	BAR ² _[-ref]	SKOP-TREUE
☞ K ₁ : ... V _[+w] [CP [S Q ₁] ... wie ₁ ...]			
K ₂ : Q ₁ ... V _[+w] [CP [S -] ... wie ₁ ...]		*!	*

(60) *Input-Optimierung:*

Man nehme an, dass verschiedene unterschiedliche Inputs I_1, I_2, \dots, I_n in einer Grammatik zu korrespondierenden Outputs O_1, O_2, \dots, O_n führen, die alle als dieselbe Form Φ realisiert werden. Es muß einer dieser Outputs der harmonischste sein, weil er die am wenigsten signifikanten Verletzungen hervorruft; dieser Output sei O_k . Dann sollte der Lerner als zugrundeliegende Form von Φ den Input I_k wählen.

Frage:

Diese Erklärung des W-Insel-Effektes beruht nicht auf der Präsenz einer W-Phrase als Intervenierer in SpecC (anders als praktisch sämtliche existierenden Erklärungen). Wieso kann dann lange Bewegung aus einem Deklarativsatz einen analogen Neutralisierungseffekt vermeiden?

Lösung:

Genial!

T₁₇: Deklarativsätze und weiter Adjunktskopus im Chinesischen

Kandidaten	SEL	BAR ² _[-ref]	SKOP-TREUE
☞ K ₁ : [S Q ₁] ... V _[-w] [CP t' ₁ ... wie ₁ ...]		*	
K ₂ : [S -] ... V _[-w] [CP Q ₁ ... wie ₁ ...]	*!		*

Extraktion und lokale Konjunktion

Treuebeschränkungen

Schon bekannt:

SKOP-TREUE.

Dazu kommen:

(61) *Treuebeschränkungen:*

a. **W-TREUE** (“Parse(wh)”):

Ein im Index auf XP auftretendes W-Merkmal muß im Output durch eine Operator-Variablenkette realisiert werden.

b. **TOP-TREUE** (“Topik-Treue”, “Parse(top)”):

Ein im Index auf XP auftretendes Top-Merkmal muß im Output durch eine Operator-Variablenkette realisiert werden.

Lokalitätsbeschränkungen

- (62) a. **BAR**^{2[-ref]} (“Zwei-Barrieren-Bedingung für Adjunktketten”):
Ein Kettenglied einer nicht-referentiellen Kette darf nicht zwei Barrieren überkreuzen.
- b. **BAR** (“Barrierenbedingung”):
Ein Kettenglied darf keine Barriere überkreuzen.

Bemerkung:

(62-a) geht auf (62-b) zurück.

T₁₈: Eine falsche Prognose unter BAR

Kandidaten	BAR
☞ $K_1: \alpha_1 \dots \beta \dots \beta \dots t'_1 \dots \beta \dots t_1$	** *
☞ $K_2: \alpha_1 \dots \beta \dots \beta \dots \beta \dots t_1$	***

(63) *Lokale Konjunktion:*

- a. Die lokale Konjunktion zweier Beschränkungen B_1 , B_2 bzgl. eines Bereichstyps D ergibt eine neue Beschränkung $B_1 \&_l B_2$, die verletzt wird, falls zwei unterschiedliche Verletzungen von B_1 und B_2 in einem einzigen Bereich vom Typ D auftreten.
- b. Universell gilt: $B_1 \&_l B_2 \gg \{B_1, B_2\}$

(64) *Subhierarchie der minimalen Kettenglieder*

('minimal link subhierarchy'):

- a. $BAR \&_l BAR = BAR^2$:
Ein Kettenglied darf nicht zwei Barrieren überkreuzen.
- b. $BAR^2 \&_l BAR = BAR^3$:
Ein Kettenglied darf nicht drei Barrieren überkreuzen.
- c. BAR^n :
Ein Kettenglied darf nicht n Barrieren überkreuzen.
- d. *Fixe Ordnung*:
 $\dots \gg BAR^3 \gg BAR^2 \gg BAR^1$

T_{19} : Eine korrekte Prognose unter der BAR-Subhierarchie

Kandidaten	BAR ³	BAR ²	BAR ¹
☞ $K_1: \alpha_1 \dots \beta \dots \beta \dots t'_1 \dots \beta \dots t_1$		*	*
$K_2: \alpha_1 \dots \beta \dots \beta \dots \beta \dots t_1$	*!		

Übereinkunft:

Selbstverständlich verletzt auch K_2 BAR² und BAR¹, denn es folgt unmittelbar aus dem Begriff der lokalen Konjunktion, dass BAR ^{n} zu BAR ^{$n-1$} als speziell zu generell in einer Stringenzrelation steht. Dies kann jedoch nie eine Rolle spielen, und daher muß man die entsprechenden Verletzungen nicht unbedingt mehr in Tabellen berücksichtigen.

Bemerkung:

Jede Lokalitätstheorie braucht ein Verfahren, um das unterschiedliche Verhalten von Argument- und Adjunktextraktion zu erfassen.

(65) **REF** (“Referentialität”):

Ketten sind referentiell.

(66) *Argumente für “Referentialität”:*

- a. *How many kilos do you wonder whether he weighs?
- b. ?Where do you wonder whether to go?
- c. *Who left why?
- d. Who lives where?

Annahme:

Es gibt lokale Konjunktion von REF mit BAR-Beschränkungen.

Frage:

Wie kommt man von (67-a) und (67-b) zu (67-c)?

- (67)
- a. Ein Kettenglied darf keine Barriere überschreiten.
 - b. Eine Kette ist referentiell.
 - c. Ein Kettenglied einer nicht-referentiellen Kette darf keine Barriere überschreiten.

(68) *Kompositionelle Interpretation von* $\text{BAR}^1[-ref]$:

- a. [Ein Kettenglied darf keine Barriere überschreiten] \wedge [Eine Kette ist referentiell].
- b. [Ein Kettenglied darf keine Barriere überschreiten] \vee [Eine Kette ist referentiell].

Situation:

Ein Satz der Art $A \vee B$ ist wahr genau dann, wenn wenigstens einer der beiden Teilsätze wahr ist, und falsch folglich nur dann, wenn beide Teilsätze falsch sind.

Das ist die intendierte Interpretation für $\text{BAR}^1[-ref]$: Diese Beschränkung ist verletzt nur dann, wenn sowohl BAR^1 , als auch REF verletzt ist. Als Ergebnis können wir somit festhalten, dass zumindest im vorliegenden Fall (de facto bei den meisten syntaktischen Anwendungen, die mir bekannt sind) die lokale **Konjunktion** logisch als **Disjunktion** zu interpretieren ist.

(Darüber hinaus ist eine Spezialannahme für reflexive lokale Konjunktion zu machen.)

(69) *Subhierarchie der minimalen nicht-referentiellen Kettenglieder:*

a. $\text{BAR}^n \&_{\text{REF}} = \text{BAR}^{n[-\text{ref}]}$:

Ein Kettenglied einer nicht-referentiellen (Adjunkt-) Kette darf nicht n Barrieren überkreuzen.

b. *Fixe Ordnung:*

$\dots \gg \text{BAR}^{3[-\text{ref}]} \gg \text{BAR}^{2[-\text{ref}]} \gg \text{BAR}^{1[-\text{ref}]}$

(70) *Universal: Adjunktbewegung ist restringierter als Argumentbewegung:*

$\text{BAR}^{n[-\text{ref}]} \gg \text{BAR}^n$

W-Bewegung im Englischen

(71) *Ordnung im Englischen:*

SEL \gg

W-TREUE \gg

SKOP-TREUE \gg

$$\text{BAR}^{4[-ref]} \gg \text{BAR}^{3[-ref]} \gg \text{BAR}^4 \gg \text{BAR}^3 \gg \text{BAR}^{2[-ref]} \gg \text{BAR}^2 \gg \text{BAR}^{1[-ref]} \gg \text{BAR}^1$$

(72) $\text{BAR}^2 \gg \text{BAR}^{1[-ref]}$

a. How₁ did [IP he [VP fix what₂] t₁] ?

b. *What₂ did [IP he [VP fix t₂] how₁] ?

(73) *Kurze Bewegung:*

a. How₁ did [IP she [VP do it] t₁] ?

b. What₁ did [IP she [VP do t₁]] ?

(19) *Ordnung im Englischen:*

SEL \gg BAR⁴^[-ref] \gg BAR³^[-ref] \gg BAR⁴ \gg BAR³ \gg
 W-TREUE \gg BAR²^[-ref] \gg
 SKOP-TREUE \gg BAR² \gg BAR¹^[-ref] \gg BAR¹

(74) *Bewegung aus einem Deklarativsatz:*

- a. How₁ do [IP you [VP think [CP t'₁ that [IP she [VP did it] t₁]]]] ?
 b. What₁ do [IP you [VP think [CP t'₁ that [IP she [VP did t₁]]]]] ?

(19) *Ordnung im Englischen:*

SEL \gg BAR⁴^[-ref] \gg BAR³^[-ref] \gg BAR⁴ \gg BAR³ \gg
 W-TREUE \gg BAR²^[-ref] \gg
 SKOP-TREUE \gg BAR² \gg BAR¹^[-ref] \gg BAR¹

(75) *Bewegung aus einer W-Insel:*

- a. *How₁ do [IP you [VP wonder [CP t'₁ what [IP PRO to fix t t₁]]]] ?
 b. What₁ do [IP you [VP wonder [CP t'₁ when [IP PRO [VP to fix t₁]]]]] ?

(76) *Neutralisierung:*

- You wonder [CP how₁ [IP PRO [VP to fix what] t₁]]
 b. *You wonder [CP what [IP PRO [VP to fix t] in some way₁]]

(19) *Ordnung im Englischen:*SEL \gg W-TREUE \gg SKOP-TREUE \gg

$$\text{BAR}^{4[-ref]} \gg \text{BAR}^{3[-ref]} \gg \text{BAR}^4 \gg \text{BAR}^3 \gg$$

$$\text{BAR}^{2[-ref]} \gg$$

$$\text{BAR}^2 \gg \text{BAR}^{1[-ref]} \gg \text{BAR}^1$$
(77) *Adjunktinseln:*a. *How₁ was [IP he [VP fired [CP after behaving t₁]]] ?b. *What₁ was [IP he [VP fired [CP after reading t₁]]] ?(78) *Neutralisierung:*a. He was [VP fired [CP after behaving in some way₁]]b. He was [VP fired [CP after reading something₁]]

(19) *Ordnung im Englischen:*

SEL \gg BAR^{4[-ref]} \gg BAR^{3[-ref]} \gg BAR⁴ \gg BAR³ \gg
 W-TREUE \gg BAR^{2[-ref]} \gg
 SKOP-TREUE \gg BAR² \gg BAR^{1[-ref]} \gg BAR¹

(79) *Subjektinseln:*

- a. *How₁ would [IP [CP t'₁ PRO to behave t₁] be inappropriate] ?
 b. *Who₁ would [IP [CP t'₁ PRO to kiss t₁] be inappropriate] ?

(80) *Intendiertes Ergebnis: Neutralisierung:*

- a. [IP [CP PRO to behave in some way₁] would be inappropriate]
 b. [IP [CP PRO to kiss someone₁] would be inappropriate]

Problem:

Es gibt nur zwei intervenierende Barrieren; das reicht nicht. (Mögliche Lösung: mehr Struktur in Subjektsätzen, wie z.B. eine leere NP-Schale.)

Einordnung

Potentielle Probleme:

(i) *Lokale Konjunktion:*

Der Mechanismus der lokalen Konjunktion von Beschränkungen ist aufgrund seiner Rekursivität zwar elegant, aber auch sehr mächtig. Ohne weitere Restriktionen gibt man damit eine elementare Annahme über Grammatiken auf, nämlich die, dass die Menge der Beschränkungen in einer Grammatik finit ist.

(ii) *Neutralisierung:*

Dasselbe gilt für den Mechanismus der Neutralisierung.

Konsequenz:

Die beiden Kernstücke des Ansatzes von Legendre, Smolensky & Wilson (1998) – lokale Konjunktion und Neutralisierung – führen zwar zu einer überzeugenden Theorie der Lokalität von Ketten, dies aber nicht umsonst: Der Preis ist ein Komplexitätsanstieg.