Optimality theory in morphology and syntax

Dieter Wunderlich (Berlin)

Nov. 2004

to appear in: Encyclopedia of Language & Linguistics – 2nd edition, Oxford: Elsevier

abstract: The article gives a brief summary of work dealing with morpho-syntactic phenomena in OT, and also addresses more general problems. It highlights on questions of optimal ordering, optimal case, and optimal agreeement, including phenomena such as morphological gaps, repairs and case-split. It finally addresses the concepts of stochastic OT and bidirectional OT.

key words: bidirectional OT, candidate set, case-split, constraint-ranking, faithfulness, harmonic alignment, lexical marking, optimal agreement, optimal case, optimal word, optimality theory, parallel optimization, repair, serial optimization, stochastic OT

1. Some basic issues

Optimality Theory (OT) has originally been developed for dealing with phonological problems, abandoning the assumption that grammatical constraints are inviolable (Prince & Smolensky 1993/2004, McCarthy & Prince 1995). The fundamental new ideas have then soon been adopted in other grammatical domains, too, in syntax (Grimshaw 1997, Ackema & Neeleman 1998, Grimshaw & Samek-Lodovici 1998, Barbosa et al. 1998, Dekkers et al. 2000, Legendre et al. 2001) as well as in morphology (Legendre 2000, Wunderlich 2001).

The main innovations of OT can be summarized as follows: (i) Grammatical constraints can conflict with each other and are violable under certain conditions (not inviolable, as assumed before). (ii) Grammatical constraints are ordered according to their respective weight (not equally ranked). (iii) Different rankings of constraints are responsible for differences between individual grammars or languages. (What was formerly conceived as grammatical parameters, is now reconstructed as constraints being differently ranked.) (iv) A construction is grammatical (not only by virtue of its own properties, or by virtue of its generative history, but rather) if it wins the competition in a candidate set, because it satisfies best the higher-ranked constraints.

At first sight, many phenomena that had troubled linguists over years seemed to be good candidates for application of OT, such as variation in word order, selective blockings, morphological gaps, repairs, or the emergence of unmarked forms. However, as soon as one starts to analyze these phenomena more systematically, many important theoretical questions have to be answered: Must all the constraints belong to a universal set of constraints, or can there be language- or even construction-specific constraints, which are based on the instantiation of lexical categorization? Is it possible to represent cumulative constraints in situations where the effects of individual constraints seem to be added up? Which restrictions over constraints are necessary? Given the concept that each re-ranking of constraints defines a possible grammar, how can one avoid the problem of over-generation of grammars? And how complex does the OT architecture develop, if one assumes an infinite set of candidates, each with its own internal structure? Is it possible to restrict the set of possible candidates, that is, to define some relevant set of candidates? How is optionality to be dealt with (where more than one candidate seems to be optimal), and how ineffability (where no candidate seems to be optimal)? Can OT also cover the phenomenon of gradient grammaticality? And last but not least, how can an OT grammar be learned? (This last issue is dealt with in Tesar & Smolensky 2000.)

Many of these questions have been addressed in the phonological OT literature (Kager 1999, McCarthy 2004), as well as in the OT syntax. Lack of space prevents me from addressing these questions in detail. Moreover, note that OT syntax can be performed in the shape of several distinct morpho-syntactic theories, such as Government and Binding theory (GB) or Minimalist Program (MP) (Grimshaw 1997, Barbosa et al. 1998, Müller 2000, Legendre et al. 2001), Lexical Functional Grammar (LFG: Bresnan 2000 XXX), Lexical Decomposition Grammar (LDG: Wunderlich 1999, Stiebels 2000), or Role and Reference Grammar (RRG: Nakamura 1999).

If one steps from phonological to morphological and finally to syntactic optimization, an important problem arises, concerning the question of serial or parallel optimization. Under the parallel view, which is prevailing among researchers, even a rather complex construction is evaluated in one step, the different sub-aspects performed simultaneously, whereas under the serial view, several local optimizations (of the word, or of the clause) are performed cyclically. Kiparsky (2000) argues for the cyclic model as most realistic for the semitic verb morphology (distinguishing optimal stem and optimal word), see also Blevins (1997) and McCarthy (2000). On the basis of detailed analyses of various phenomena (including remnant movement, superiority, reflexives and binding, pied piping), Müller and co-workers argue for local, i.e. cyclic, optimization in the syntax (Heck & Müller 2000, Müller 2002, Fischer 2004).

The candidate set is generated on the basis of a given input (underlying form). In the classic OT (the so-called correspondence theory), each candidate (a potential output) is evaluated according to faithfulness: dependence constraints penalize the occurrence of elements (or features) in the output that do not have a correspondent in the input, maximality constraints penalize the presence of elements (or features) in the input that do not have a correspondent in the output, and identity constraints penalize a featural mismatch between input and output. Needless to say, further constraints are needed to check the linear ordering in the output, whether some morpheme or constituent is initial or final, or whether it is immediately to the left or to the right of the verbal root/the finite verb, and so on.

Grimshaw (1997) holds that the input for a syntactic evaluation must be some representation of meaning, for instance an LF representation; the same should hold for morphological evaluations. However, Heck et al. (2002), in discussing several examples of dependent as well as movement structures, come to the conclusion that input can be dispensed with in OT-syntax. They argue that syntax is information-preserving, while phonology is not. This conclusion, however, crucially depends on a specific conception of syntax, it does not hold for LFG or LDG.

In the remainder of this article, more illustrative OT examples are given that concern syntactic movement (section 2), the constitution of optimal words (section 3), the selection of optimal case (section 4), the effects of harmonic alignment (section5), and the selection of optimal agreement (section 6). Section 7 considers some further developments.

2. Syntactic movements and repairs

The pioneering work in OT syntax is Grimshaw (1997, Manuscript 1994), dealing with the sentence structure of English (including interrogatives, *wh*-movement, *do*-support and negative-induced inversion) within GB. Observe that only the last of the sentences in (1) is grammatical.

(1) a. *They will read what?
b. *Will they read what?
c. *What they will read?
d. What will they read?

The set of candidates based on the input ‘read(they, which books)’ includes structural representations of all potential strings constituted by the given items. Among the constraints Grimshaw assumes are the following:

(2) STAY: Do not move.
OP-SPEC: Operators are realized in SPEC position.
OBL-HD: Heads must be lexically filled.

The evaluation in (3) shows that two movements (in violation of STAY) are needed in order to satisfy the higher-ranked constraints. (The constraint columns are ordered from left to right; ‘*’ indicates a violation, while ‘*!’ indicates a fatal violation. The optimal candidate is indicated by ‘☞’.)

(3) Interrogatives in English (Grimshaw 1997: 378)

candidates	OP-SPEC	OBL-HD	STAY
a. [$_{IP}$ they will [$_{VP}$ read what]]	*!		
b. [$_{CP}$ will$_k$ [$_{IP}$ they t_k [$_{VP}$ read what]]]	*!		*
c. [$_{CP}$ what$_i$ [$_{IP}$ they will [$_{VP}$ read t_i]]]		*!	*
☞ d. [$_{CP}$ what$_i$ will$_k$ [$_{IP}$ they t_k [$_{VP}$ read t_i]]]			**

The same set of constraints explains the negative-induced inversion in (4) (Grimshaw 1997: 401)

(4) a. *She said that she would work this hard under no circumstances again.
b. She said that under no circumstances would she work this hard again.

Besides movement, many kinds of repairs have been investigated in OT-syntax. They involve the introduction of empty elements such as expletives or *do* in English *do*-support. Another example of repair is the German ersatzinfinitive, which is the neutralized form of a past participle governing an infinitive (5b). This construction does not appear with durative verbs such as *bleiben* ‘remain’ (6c).

(5) German ersatzinfinitive
a. *Ich habe das Buch lesen gewollt.
I have the book read.INF want.PASTP
b. Ich habe das Buch lesen wollen.
I have the book read.INF want.INF
‘I wanted to read the book.’
c. Die Leute sind stehen geblieben.
the people are stand.INF remain.PASTP
‘The people remained standing.’

The solution proposed by Schmid (2002) makes use of the constraints in (6).

(6) a. MORPH: Morphological selectional properties of lexical items must be observed.
b. *PASTP/+INF.MOD: The past participle of a modal verb must not be a sister of a VP whose head is an infinitive.

c. *PASTP/+INF.DUR: The past participle of a durative verb must not be a sister of a VP whose head is an infinitive.

The avoid constraint is split into a family of constraints according to hierarchically ordered subclasses of verbs governing an infinitive; some are ranked higher than MORPH, while others are ranked lower. This is a standard solution for category-specific syntactic regularities.

(7) German ersatzinfinitive (Schmid 2002: 145)

candidates	*PASTP/+INF.MOD	MORPH	*PASTP/+INF.DUR
a. [[$_{VP}$ ... lesen$_{INF}$] gewollt$_{PASTP}$]	*!		
☞ b. [[$_{VP}$ ... lesen$_{INF}$] wollen$_{INF}$]		*	
☞ a. [[$_{VP}$ stehen$_{INF}$] geblieben$_{PASTP}$]			*
b. [[$_{VP}$ stehen$_{INF}$] bleiben$_{INF}$]		*!	

Lots of other syntactic problems have been discussed in the OT framework: long extraction, topicalization and focusing, word ordering in the German middle field, pied piping, quantifier scope, anaphoric binding, and resumptive pronouns.

3. Optimal words

An inflected word is constituted by various inflectional affixes attached to a stem. One point in question is how the ordering of affixes can be described best: is it possible to predict the ordering of affixes on the basis of a syntactic structure (according to the mirror hypothesis proposed by Baker), or are there language-specific templates that integrate all the idiosyncratic properties of individual languages? Many languages show patterns deviating from syntactic predictions, but on the other hand, the templatic account misses generalizations that one would like to make. The proponents of Minimalist Morphology (Wunderlich & Fabri 1995) claim that inflectional affixes are (bound) lexical items that can be freely attached to a stem, thus yielding elements of a candidate set, which are evaluated by means of a constraint ranking.

Yimas, a language of Papua New Guinea, is a good case for illustration because it shows some variation as to how pronominal affixes are realized on the verb (Foley 1991, Wunderlich 2001). Besides free pronouns (always unmarked for case), Yimas exhibits ergative (E), accusative (A) and nominative (N) prefixes in the 1st and 2nd person, as well as ergative and nominative prefixes and dative (D) suffixes in the 3rd person. With transitive verbs, only E or A affixes show up, the former in case of 3rd person objects (8a), and the latter in case of 1st or 2nd person objects (8b,c); the marked case is always closer to the stem. However, in the combination 1Ag/2Pat one finds either a pronominal gap with the free variant of the pronoun instead (8d), or a repair phenomenon, namely a fused morpheme (8e).

(8) Transitive verbs in Yimas

a. Ergative pattern:

na- mpu- tay	pu- ka- tay
3sgN- 3plE- saw	3plN- 1sgE- saw
'They saw him.'	'I saw them.'

b. Accusative pattern:

pu- ŋa- tay	ma- ŋa- tay
3plN- 1sgA- saw	2sgN- 1sgA- saw

‘They saw me.’ ‘You saw me.’

d. Gap in the combination 1Ag/2Pat:
   ipa kul- tay (*ipa-kul-tay, *ipwa-kay-tay)
   1pl 2plA-saw 1plN-2plA-saw 2plN-1plE-saw
   ‘We saw you(pl)’

e. Fused morpheme in the combination 1Ag/2sgPat:
   ipa kampan- tay (*ipa-nan-tay, *ma-kay-tay)
   1pl 1E/2sgA-saw 1plN-2sgA-saw 2sgN-1plE-saw
   ‘We saw you(sg)’

With ditransitive verbs, the dative is only possible in case of 3rd person recipients (9a), otherwise the recipient is marked accusative, so that ergative and accusative can be combined (9b). In the combination of 1st and 2nd person (Ag or Rec), one finds either a gap (9c) or the same fused morpheme as above (9d). (VI represents the gender class.)

(9) Ditransitive verbs in Yimas

a. Ergative pattern plus D:
   uraŋ k- ka- tkam-r- mpun
   coconut.VIsg VIsg- 1sgE- show-PERF-3plD
   ‘I showed them the coconut’

b. Three-partite pattern N-E-A:
   uraŋ k- mpu- ŋa- tkam-t
   coconut.VIsg VIsg- 3plE- 1sgA- show-PERF
   ‘They showed me the coconut’

c. Gap:
   ipa uraŋ k- kul- tkam-t (*k- nan- ŋa- tkam-t)
   1pl coconut.VIsg VIsg- 2plA- show-PERF VIsg-2plE-1sgA- show-PERF
   ‘We showed you(pl) the coconut’

d. Fused morpheme:
   uraŋ k- mpan - tkam-t (*k- ka- nan- tkam-t)
   coconut.VIsg VIsg- 1E/2sgA- show-PERF VIsg-1sgE-2sgA-show-PERF
   ‘I showed you(sg) the coconut’

In the context of a modal prefix, nominative turns to ergative, as shown in (10).

(10) Modal prefixes in Yimas

a. ama-tmuk-t
   1sgN-fall-PERF
   ‘I fell down’

b. ant-ka-tmuk-t
   POT-1sgE-fall-PERF
   ‘I almost fell down.’

c. pu-ŋa-tpul
   3plN-1sgA-hit
   ‘They hit me.’

d. ka-mpu-ŋa-tput-n
   LIKE-3plE-1sgA-hit-PRES
   ‘They are going to hit me.’

The constraints that can deal with the observed distribution include the following:

(11) Morphological constraints in Yimas (slightly simplified). (‘M’ indicates the morphological or word domain.)

a. MAX$(arg)_M$: All arguments are realized by pronominal affixes.
b. *$(arg)_M$/+FOCUS: Avoid pronominal affixes for arguments that bear the feature [+FOCUS]
c. PERSON: The higher person is realized closer to the stem (scale: 1 > 2 > 3).

d. DEFAULT: Every linking domain displays the default linker (nominative).
e. NOM-INITIAL: Nominative prefixes only appear word-initially.
f. CASE: The less marked case is realized closer to the stem (scale: ACC > ERG).
g. IDENT(person): All person features have identical values in the input and the output.

Most of these constraints follow from universal conditions. It is better to realize an argument pronominally than nominally, and it should be realized in the closest possible domain (which is the word in Yimas) (11a). However, an argument that introduces a new referent in the discourse, or is questioned or relativized (all three situations could be described by the feature [+FOCUS]), should not be realized by a pronominal affix, because pronouns generally are less suitable to bear focus (11b). *$(arg)_M$/+FOCUS thus automatically restricts MAX$(arg)_M$. The person scale 1 > 2 > 3 (or 1,2 > 3) is universal, but how it interacts with ordering problems is free (11c). (In Maung, a language of Australia, it is the lower person that is realized closer to the stem, Donohue 1998.) All case systems include nominative as the default case (11d). The nominative prefixes of Yimas are nearly identical to the free pronouns and probably originate from cliticisation; this explains why (11e) holds (called EDGEMOST in Legendre 2000). The markedness scale ACC > ERG is universal, but again, the way how it interacts with ordering problems is free (11f). Finally, IDENT(person) is universal, too.

The requirements of PERSON and CASE conflict with each other in case of the 1Ag/2Th (or 1Ag/2Rec) combination, and here the position of IDENT(person) in the constraint-ranking is decisive. In Yimas, a gap is easier tolerated than neutralization of person (while in Dalabon, a language of Australia, neutralization is preferred: a 1st person object can be realized by a 3rd person prefix, Evans et al. 2001). However, it is still open which of the two arguments then becomes gapped in the morphology. Only if one adds the ranking MAX(ACC) » MAX(ERG) ('Realize ACC/ERG'), one obtains the desired result, as shown in the tableau (12). We will turn to these constraints in the next section.

(12) A gap introduced by conflicting ordering constraints: Evaluation in the combination 1Ag/2plRec (or 1Ag/2plTh without the first 3N-prefix)

candidates	PERSON	CASE	ID(PERSON)	MAX$(arg)_M$	MAX(ACC)	MAX(ERG)
3N-1E-2A	*!					
3N-2A-1E		*!				
3N-3E-2A			*!			
3N-1E				*	*!	
☞3N-2A				*		*

Conflicting alignment constraints leading to gaps (or repairs) have also been considered by Legendre (2000) and Gerlach (2002), who studied the formation of clitic sequences in South Slavic and Romance, respectively.

4. Optimal case

Syntactic clauses are usually headed by a verb, which can have several arguments: one (intransitives), two (transitives), three (ditransitives) or even four or five (causativized ditransitives, and so on). One major issue is therefore the set of conditions that regulate case, realized by pronominal affixes or by morphological case; one may also regard syntactic positions of arguments as a realization of 'abstract case'. Most accounts consider

the semantic roles of arguments as the determining factor. However, a purely structural account is possible under the assumption that the arguments of a predicate are strictly ordered (normally, of course, controlled by some semantic conditions, but often also idiosyncratically). In many instances (such as causativization) the argument hierarchy can be read off from the decomposition of the complex predicate. Lexical Decomposition Grammar (LDG, Wunderlich 1997, Wunderlich & Lakämper 2001, Stiebels 2002) assumes that this hierarchy inheres 'abstract case', which is then realized by morphological or syntactic means.

The argument roles of a hierarchy can be encoded by means of two relational 'abstract case' features: [+hr] 'there is a higher role', and [+lr] 'there is a lower role' (similar to an earlier proposal by Kiparsky). The highest argument is [–hr] ('there is no higher role'), and the lowest argument is [–lr] ('there is no lower role'); in addition, depending on the presence of further arguments, the positive features are encoded. Thus, canonical encoding yields the following theta-roles (lambda abstractors over arguments):

(13) Abstract case features:

a. intransitives: λx verb(x)
[–hr,–lr]

b. transitives: λz λx verb(x,z)
[+hr,–lr] [–hr,+lr]

c. ditransitives: λz λy λx verb(x,y,z)
[+hr,–lr] [+hr,+lr] [–hr,+lr]

However, single features can be overridden by specific lexical assignments, which often reflect semantic conditions. For instance, dative-subject verbs such as German *gefallen* or Icelandic *likar* (both 'like') can be characterized by (14) and thus suggest the experiencer-reading 'The highest argument behaves as affected'.

(14) Lexical assignment in the presence of abstract case:

λz λx verb(x,z)
[+hr,–lr] [**+hr**,+lr]

The same features (but only those with positive values) are used to characterize overt case:

(15) DAT: [+hr,+lr]
ACC: [+hr]
ERG: [+lr]
NOM: [ ]$_{case}$

Observe that dative is only compatible with the medial argument of ditransitive verbs or the lexically marked argument in (14), while accusative is compatible with more arguments (including all those that could in principle be realized by dative). Finally, nominative as the maximally underspecified case is compatible with all arguments. Therefore further constraints are necessary; this approach to case thus calls for OT. In this framework, lexical items with the format of (13) or (14) constitute the input, and the relevant output candidates have to show some pattern of overt case.

The relation between input and output is minimally determined by faithfulness (MAX); in addition, markedness controls that not too many features are realized.

(16) a. MAX(+F): Every feature [+F] in the input has a correspondent in the output.
b. *[+F]: Avoid [+F] in the output.

The rankings MAX(+hr) ≫ *[+hr] and *[+lr] ≫ MAX(+lr) characterize a pure accusative system, while the reverse rankings characterize a pure ergative system. However, dative can appear in both types of system, which has to be guaranteed by a further constraint (MAX(+hr,+lr)), which ranks above both MAX(+hr) and MAX(+lr). Moreover, lexically assigned features should have precedence (17a) (see also Woolford 2001), nominative should appear for reasons of economy (17b), and the morphological cases of a local domain should be distinct (17c).

(17) General case constraints
- a. MAX(lexF): Every lexically assigned feature in the input has a correspondent in the output.
- b. DEFAULT: Every linking domain displays the default linker (nominative).
- c. UNIQUENESS: Each linker applies only once in a local domain.
- d. MAX(+hr,+lr): Every feature combination [+hr,+lr] in the input has a correspondent in the output.

Wunderlich (2003) shows that this inventory is also sufficient to characterize all case patterns occurring in German and Icelandic, provided that all non-canonical case patterns are characterized by additional lexical features. Moreover, these two languages share the same constraint-ranking (although Icelandic exhibits by far more marked case patterns). How this constraint-ranking works is illustrated here for lexically marked ditransitive verbs and their passives.

(18) Lexically marked ditransitive verbs:

a. weil der Arzt den Patienten einem Test unterzog
because the.NOM doctor the.ACC patient a.DAT test exposed
'because the doctor exposed the patient to a test'

	λz	λy	λx	EXPOSE(x,y,z)
lexical	**+lr**			
default	+hr	+hr	-hr	
		+lr	+lr	

b. weil der Patient einem Test unterzogen wurde
because the.NOM patient a.DAT test exposed was
'because the patient was exposed to a test'

	λz	λy	∃x	EXPOSE(x,y,z)
lexical	**+lr**			
default	+hr	+hr		
		+lr		

(19) Lexically marked ditransitive verbs

	z y x	MAX (lexF)	DEFAULT	UNIQUE-NESS	MAX (+hr,+lr)	MAX (+hr)	*[+lr]	MAX (+lr)	*[+hr]
☞	DAT ACC NOM				*		*	**	**
	ACC DAT NOM	*!			*		*	**	**
	DAT DAT NOM			*!			**	*	**
	ACC ACC NOM	*!		*	**			***	**

(20) Passive of lexically marked ditransitive verbs

	z y	MAX (lexF)	DEFAULT	UNIQUE-NESS	MAX (+hr,+lr)	MAX (+hr)	*[+lr]	MAX (+lr)	*[+hr]
☞	DAT NOM				*	*	*	*	*
	NOM DAT	*!			*	*	*	*	*
	DAT ACC		*!		*		*	*	**
	DAT DAT		*	*			**		**

Stiebels (2000, 2002) bases her typology of linking systems on the inventory given in (16) and (17). Further OT-work on optimal case includes Nakamura (1997, 1999) in RRG, and Woolford (2001) in GB.

5. Harmonic alignment

Harmonic alignment is an OT-concept that captures implicational generalizations (Prince & Smolensky 1993). Suppose a scale X > Y, next to a scale a > b > ... > z. Then harmonic alignment of the two dimensions is the pair of the following scales:

(21) a. X/a ⊃ X/b ⊃ ... ⊃ X/z
b. Y/z ⊃ ... ⊃ Y/b ⊃ Y/a

A possible interpretation is that X and Y are linguistic entities, where X is more frequent or more expressive or less marked than Y. The elements of the other scale are possible contexts, ordered according to salience. Then X/a ('X in the context of a') is less marked than X/b, whereas Y/a is more marked than Y/b. What is more marked is likely to be more avoided. Therefore, the constraint alignment is a pair of constraint hierarchies:

(22) a. *X/z » ... » *X/b » *X/a
b. *Y/a » *Y/b » ... » *Y/z

Aissen (1999) was the first who applied this concept to morpho-syntactic phenomena. She assumed subject > object as the first scale, and the animacy or definiteness scale as the contextual scale. Silverstein (1976) observed that ergative is often restricted to 3rd person as opposed to 1st or 2nd person, to inanimates as opposed to humans, or to indefinites as opposed to definites, thus characterizing various kinds of ergative-split. A similar, but reversed split is found with objects: accusative is often restricted to 1st or 2nd person, to animates or to definites; this is analyzed in OT by Aissen (2003).

The account of Aissen is rather complicated, and it fails to give a conceptually and empirically satisfying solution. The main reason is that she appeals to grammatical roles (subject, object) rather than case features. Stiebels (2000, 2002) proposes to make use of the scale [+hr] > [+lr] instead ('It is better to mark a lower argument (an object) than a higher argument (the subject)'). Given the salience scales 1, 2 > 3, animate > inanimate, definite > indefinite, pronoun > noun, imperfect > perfect, harmonic alignment then yields the constraint hierarchies in (23) without further complications. (For convenience, +hr is abbreviated as ACC, and +lr is abbreviated as ERG.)

(23) a. *ERG/1,2 » *ERG/3 *ACC/3 » *ACC/1,2
b. *ERG/+anim » *ERG/–anim *ACC/–anim » *ACC/+anim
c. *ERG/+def » *ERG/–def *ACC/–def » *ACC/+def
d. *ERG/+pro » *ERG/–pro *ACC/–pro » *ACC/+pro

e. *ERG/–perf » *ERG/+perf *ACC/+perf » *ACC/–perf

Effects that are based on these universal tendencies are found in the inventory of pronouns or pronominal affixes ('emergence of the unmarked'), as well as in all sorts of case-splits. Yimas (see section 3) has accusative affixes in the 1st and 2nd person, but lacks them in the 3rd person (according to *ACC/3 » *ACC/1,2). A similar case is Dyirbal, a Pama-Nyungan language of Australia (Dixon 1994): Pronouns of the 1st and 2nd person can only be marked for accusative (*-na*), while demonstratives as well as nouns can only be marked for ergative (*-ŋgu*). Such a language allows for four different case patterns with transitive verbs: ERG-ACC, ERG-NOM, NOM-ACC, and NOM-NOM; each is signalling a certain salience distribution of subjects and objects.

Hindi displays these four case patterns, too: Ergative is restricted to perfect, whereas accusative is only found with human, specific animate or definite inanimate objects.

(24) The four-way case-split of Hindi (Mohanan 1994)

a. niinaa-ne baalikaa-ko ut^{h}aa-y-aa.
Nina.F-ERG girl-ACC lift-PERF-M
'Nina lifted up a/the girl.'

b. niinaa baalikaa-ko ut^{h}aa-eg-ii.
Nina.F.NOM girl-ACC lift-FUT-F
'Nina will lift up a/the girl.'

c. niinaa-ne roTii khaa-y-ii.
Nina.F-ERG bread.F.NOM eat-PERF-F
'Nina ate bread.'

d. niinaa kelaa khaa-eg-ii.
Nina.F.NOM banana.M.NOM eat-FUT-F
'Nina will eat a banana.'

A split can also occur between two coexisting case-linking devices. Warlpiri, another Pama-Nyungan language of Australia, shows ergative case on nouns and pronouns, but accusative-based agreement: *-rna* '1sg' agrees with the subject of both intransitive and transitive verbs (25a-b), while *-ju* '1sg' agrees with the object (25c); consequently, *-ju* has to be classified as accusative (A), and *-rna* as nominative (N).

(25) Ergative case and accusative agreement in Warlpiri (Andrews 1985: 106f.)

a. Ngaju ka-rna parnka-mi.
I.NOM PRES-1sg.N run-NONPAST
'I am running.'

b. Ngajulu-rlu ka-rna-ngku nyuntu nya-nyi.
1-ERG PRES-1sg.N-2sg.A you.NOM see-NONPAST
'I see you.'

c. Ngaju ka-npa-ju nyuntulu-rlu nya-nyi.
1.NOM PRES-2sg.N-1sg.A you-ERG see-NONPAST
'You see me.'

Such a mixed system in which head-marking exhibits the accusative type and case-marking the ergative type is found in several languages, whereas the reverse constellation does not exist. This fact again follows from harmonic alignment, if we consider the agreement morpheme as a bound pronoun (pron$_M$): The two scales ACC > ERG and pron$_M$ > pron ('Morphological realization is preferred') induce the constraint hierarchies *ACC/pron »

*ACC/pron$_M$ and *ERG/pron$_M$ » *ERG/pron, so that a split with ACC only on pronouns and ERG only on affixes is ruled out.

Zeevat & Jäger (2002) propose a functional explanation of the effects of harmonic alignment. They argue that the contextualized constraints (such as those in (23)) emerge from statistical generalizations about the perceived linguistic input rather than being part of grammar. However, an unequal distribution of ACC and ERG in the data must have been produced by speakers prior to learners' perception. The assumption of a cognitive rationale of the scales to be combined therefore seems more plausible.

6. Optimal agreement

Grammatical agreement is a phenomenon in which word forms that co-occur in a clause exhibit covariation. Inflected forms often agree with respect to their values of number, gender, or person, as can be seen from the contrasting examples in (26). All items in (26a) refer to a single (feminine) entity, whereas in (26b) they refer to a plural (masculine) entity.

(26) French
a. La^{fsg} fille^{fsg} est^{3sg} belle^{fsg}.
'The girl is beautiful.'
b. Les^{pl} garcons^{mpl} sont^{3pl} beaux^{mpl}.
'The boys are handsome.'

Languages that use rich agreement often allow more freedom in word order, as can be seen from the following example from Latin.

(27) Grandia per multos tenuantur flumina rivos. (Ov. *Rem*. 445)
big.npl to numerous.mpl.ACC dissolve.3pl river.npl stream.mpl.ACC
'Big rivers dissolve into numerous streams'

However, the extent to which agreement is observed differs from language to language. Besides languages that show multiple agreement, there are also languages that avoid agreement under specific conditions or in certain domains, as well as languages that lack agreement altogether.

Some languages exhibit an agreement split with respect to number, similar to the various types of case splits discussed in the preceding section. Georgian shows plural agreement only with animate subjects (28b), whereas Classical Arabic does not allow plural agreement for non-pronominal postverbal subjects, where instead the neutralized singular form is found.

(28) Georgian (Harris 1981: 21)
a. Burt-eb-i gorav-s / *gorav-en.
ball-pl-NOM roll-3sg / roll-3pl
'The balls are rolling.'
b. Knut-eb-i *gorav-s / gorav-en.
kitten-pl-NOM roll-3sg / roll-3pl
'The kittens are rolling.'

(29) Classical Arabic (Aoun et al. 1994: 197, 205)
a. Naam-a /*naam-uu l-ʔawlaad-u.
slept-3sg.m / slept-3pl.m DEF-children-NOM
'The children slept.'
b. *Naam-a /naam-uu hum.
slept-3sg.m / slept-3pl.m PRON.pl.m

'They slept.'

Ortmann (2002) provides an OT-analysis of these phenomena. He assumes the expressivity scale +pl > [ ] ('Plural is more informative than underspecification') and the two salience scales animate > inanimate, pronoun > noun. Harmonic alignment yields the two markedness scales in (30).

(30) a. *(+pl)/inan » *(+pl)/anim
b. *(+pl)/noun » *(+pl)/pron

The interpretation of these avoid constraints is that it is more marked to realize plural for referents of low salience than for those of high salience. Interpolating these constraint hierarchies with the relevant MAX constraint ($MAX(+pl)_{S\text{-}AGR}$: 'Realize plural in the subject agreement') yields the possible rankings in (31).

(31) a. $MAX(+pl)_{S\text{-}AGR}$ » $*(+pl)_{S\text{-}AGR}$/inan » $*(+pl)_{S\text{-}AGR}$ /anim
(unrestricted agreement: German)
b. $*(+pl)_{S\text{-}AGR}$ /inan » $MAX(+pl)_{S\text{-}AGR}$ » $*(+pl)_{S\text{-}AGR}$ /anim
(restricted agreement: Georgian)
c. $*(+pl)_{S\text{-}AGR}$ /inan » $*(+pl)_{S\text{-}AGR}$/anim » $MAX(+pl)_{S\text{-}AGR}$
(no agreement: Chinese)

Languages thus differ with respect to different cut-off points regarding the various possible salience scales, and, of course, they can also make different choices with respect to subject agreement, object agreement, and possessor agreement.

Ortmann (2000, 2002) also deals with another phenomenon. Hungarian (unlike English or German) avoids NP-internal plural agreement. Attributive adjectives do not agree with the noun (32a), and in the context of a numeral, which inherently contains the concept of plurality, the noun must not be marked for plural (32b). However, demonstratives (outside of DP) agree with the noun (32c).

(32) Hungarian plural agreement (Ortmann 2000, 2002)

a. gyors hajó-k
fast ship-pl
'fast ships'

b. öt hajó
five ship.sg
'five ships'

c. Ez-ek a hajó-k gyors-ak.
this-pl DEF ship-pl fast-pl
'These ships are fast.'

d. Az öt hajó gyors.
DEF five ship.sg fast.sg
'The five ships are fast.'

e. Az öt nagynéni sör-t isz-ik / *isz-nak.
DEF five aunt.sg beer-ACC drink-3sg/ *drink-3pl
'The five aunts drink beer'

Moreover, subject agreement strictly respects the value of number instantiated on the head element of the subject DP or DemP (32c-e). This latter fact suggests cyclic optimization: first, the DP is checked (with the nominal head being singular in (32d,e) despite of the cardinality > 1), and the resulting information is taken as input for subject agreement.

Ortmann proposes the constraint *MULTIPLE-PLUR(DP) ('DP-referents receive only one realization of 'plurality' within the DP'), which penalizes any uneconomic plural-assignment in the DP. This constraint also interacts with a plurality-split in the Hungarian possessor agreement, which results in so-called anti-agreement: Actually, the DP-internal possessor never agrees with the agreement information on the head noun, as shown in (33).

(33) Hungarian possessor anti-agreement (Ortmann 2002: 170)

a. az ö ház-uk
DEF PRON.3sg house-POSS.3pl
'their house'

b. a nagynéni-k ház-a
DEF aunt-pl house-POSS.3sg
'the aunts' house'

In harmony with the salience split (of Classical Arabic) as discussed above it is better for the head noun to agree with a pronoun than with a full noun, and because of *MULTIPLE-PLUR(DP) dominating MAX(+pl) only one plural marking should appear in the DP. (MAX(+pl) summarizes the requirements for morphological and syntactic realization.) The tableaus in (34) illustrate how the effects of anti-agreement come about by the constraint ranking.

(34) Evaluation of anti-agreement in Hungarian

a. Pronominal possessor	*(+pl)$_{P\text{-}AGR}$/ noun	MAX(+pl)$_{P\text{-}AGR}$	*MULT-PLUR	MAX(+pl)	*(+pl)$_{P\text{-}AGR}$/ pron
PRON.3sg house-P.3sg		*!		**	
☞ PRON.3sg house-P.3pl				*	*
PRON.3pl house-P.3sg		*!		*	
PRON.3pl house-P.3pl			*!		*

b. Nominal possessor	*(+pl)$_{P\text{-}AGR}$/ noun	MAX(+pl)$_{P\text{-}AGR}$	*MULT-PLUR	MAX(+pl)	*(+pl)$_{P\text{-}AGR}$/ pron
aunt.3sg house-P.3sg		*		**!	
aunt.3sg house-P.3pl	*!			*	
☞ aunt.3pl house-P.3sg		*		*	
aunt.3pl house-P.3pl	*!		*		

The morphology and syntax of agreement is certainly one of the aspects in which languages differ widely, this topic thus being one of the most interesting parts of linguistic typology.

7. Further developments

In contrast to classic OT, **Stochastic OT** (Boersma 1998, Boersma & Hayes 2001) constraints are ranked on a continuous scale: every constraint is assigned a real number which determines the ranking among each other and is a measure for the distance between them. For each evaluation, the placement of a constraint is modified by adding a noise with normal distribution, so that the actual placement of the constraints after adding this noise value can vary if the constraints are close to each other. However, if the constraints are more distant, a categorical ranking results. This account can incorporate linguistic variation that was outside the scope of classic OT. Bresnan at al. (2001) study the role of person in active-passive variation. In Lummi (Straits Salish, British Columbia), for example, the person of the subject argument cannot be lower than the person of a nonsubject argument (the star in (35) represents ungrammaticality). The same constraints that result in a categorical effect in Lummi are then shown to give rise to statistical preferences in English, which are modeled in stochastic OT.

(35) Active-passive variation in Lummi and English

	Lummi		English	
	active	passive	active	passive
1,2 subject, 3 object	obligatory	*	frequent	rare
3 subject, 1,2 object	*	obligatory	rare	frequent

Besides OT-syntax, also OT-semantics has been developed which is concerned, among others things, with the interpretation of pronouns (Hendriks & de Hoop 2001). Integrating the two perspectives (speaker-oriented: What is an optimal form? and hearer-oriented: What is an optimal interpretation?), Blutner (2001) proposed **Bidirectional OT**, which has been formalized by Jäger (2002, 2004). All candidates are <form, meaning> pairs, and the optimal candidate wins the competition of independent variation of both meaning and form. (In the weak version of bidirectional OT the evaluation is recursive.) This account offers an adequate solution of blocking phenomena, such as **more cheap* in the presence of *cheaper,* and it singles out <*cause to die*, cause death in indirect way> vs. <*kill*, cause death in direct way> as optimal pairs. Lee (2001) explains the 'emergence of the unmarked' effect on word order in scrambling languages (free word order freezes into a fixed, canonical order under certain circumstances) using bidirectional optimization. Bidirectional OT is promising in various domains because it allows for a unified account of optionality and ineffability on the one hand, which are problematic for production OT, and of ambiguity and uninterpretability on the other hand, which are problematic for comprehension OT (Beaver & Lee 2004).

≈ 5200 words (5500 words including the tableaus)

References:

Ackema, Peter & Ad Neeleman (1998). 'Optimal questions' *Natural Language and Linguistic Theory 16*, 443-490. (Ms. 1997)

Aissen, Judith (1999). 'Markedness and subject choice in optimality theory' *Natural Language and Linguistic Theory 17*, 673-711.

Aissen, Judith (2003). 'Differential object marking: Iconicity vs. economy' *Natural Language and Linguistic Theory 21*, 453-483.

Andrews, Avery (1985) 'The major functions of the noun phrase' In Shopen, Timothy (ed.) *Language typology and syntactic description*, vol. I: *Clause structure*. Cambridge University Press, 62-154.

Aoun, Joseph, Benmamoun, Elabbas & Sportiche, Dominique (1994). 'Agreement, word-order, and conjunction in some varieties of Arabic' *Linguistic Inquiry 25*, 195-220.

Barbosa, Pilar et al. (eds.) (1998). *Is the best good enough? Optimality and competition in syntax*. Cambridge, Mass.: MIT Press.

Beaver, David I. & Lee, Hanjung (2004) 'Input-output mismatches in OT' In Zeevat & Blutner (eds.)

Blevins, Juliette (1997). 'Rules in optimality theory: Two case studies' In Roca, Iggy (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon.

Blutner, Reinhard (2001). 'Some aspects of optimality in natural language interpretation', *Journal of Semantics 17*, 189-216.

Boersma, Paul (1998). *Functional phonology: Formalizing the interaction between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul & Hayes, Bruce (2001).'Empirical tests of the gradual learning algorithm' *Linguistic Inquiry* 32, 45-86.

Bresnan, Joan, Dingare, Shipra & Manning, Christopher D. (2001). 'Soft constraints mirror hard constraints: Voice and person in English and Lummi' In Butt, Miriam & King, Tracy H. (eds.) *Proceeedings of the LFG 01 conference*. Stanford: CSLI publications, 13-32.

Bresnan, Joan (2000). 'Optimal syntax' In Dekkers et al. (eds.), 343-385.

Bresnan, Joan (2001). *Lexical functional syntax*. Oxford & Malden, Mass.: Blackwell. **weg**

Dekkers, Joost, van der Leeuw, Frank & van de Weijer, Jeroen (eds.) (2000). *Optimality theory: phonology, syntax, and acquisition*. Oxford University Press.

Dixon, R. M. W. (1994). *Ergativity*. Cambridge University Press.

Donohue, Mark (1998). 'A note on verbal agreement in Maung' *Australian Journal of Linguistics 18*, 73-89.

Evans, Nicholas, Brown, Dunstan & Corbett, Greville (2001). 'Polysemy in the Dalabon pronominal prefix paradigm: a referral-based treatment' In Booij, Geert & van der Marle, Jaap (eds.) *Yearbook of Morphology 2000*, xxx

Fischer, Silke (2004). 'Optimal binding' *Natural Language and Linguistic Theory 22*, 481-526.

Foley, William A. (1991). *The Yimas Language of New Guinea*. Stanford University Press.

Gerlach, Birgit (2002). *Clitics between syntax and lexicon*. Amsterdam: Benjamins.

Grimshaw, Jane (1997). 'Projection, heads, and optimality' *Linguistic Inquiry 28*, 373-422. (Ms. 1994)

Grimshaw, Jane & Samek-Lodovici, Vieri (1998) 'Optimal subjects and subject universals' In Barbos et al (eds.), 193-219.

Harris, Alice C. (1981). *Georgian syntax. A study in relational grammar*. Cambridge University Press.

Heck, Fabian & Müller, Gereon (2000). 'Successive cyclicity, long distance superiority, and local optimization' In Billerey, Roger & Lillehaugen, Brook D. (eds.) *Proceedings of WCCFL 19*. Somerville, Mass.: Cascadilla Press, 218-231.

Heck, Fabian et al. (2002). 'On the nature of the input in optimality theory' *The Linguistic Review* 19, 345-376.

Hendriks, Petra & de Hoop, Helen (2001). 'Optimality Theoretic Semantics' *Linguistics and Philosophy 24*, 1-32.

Jäger, Gerhard (2002). 'Some notes on the formal properties of bidirectional optimality theory' *Journal of Logic, Language and Information* 11, 427-451.

Jäger, Gerhard (2004). 'Learning constraint sub-hierarchies. The bidirectional gradual learning algorithm' In Zeevat & Blutner (eds.), 251-287.

Kager, René (1999). *Optimality theory*. Cambridge University Press.

Kiparsky, Paul (2000) 'Opacity and cyclicity' *The Linguistic Review 17*, 351-367.

Lee, Hanjung (2001). *Optimization in argument expression and interpretation: A unified approach.* Ph.D. dissertation, Stanford University.

Legendre, Géraldine, Grimshaw, Jane & Vikner, Sten (eds.) (2001). *Optimality-theoretic syntax*. Cambridge, Mass.: MIT Press.

Legendre, Géraldine, Smolensky, Paul & Wilson, Colin (1998). 'When is less more? Faithfulness and minimal links in wh-chains' In Barbosa et al. 1998, 249-289.

Legendre, Géraldine (2000) 'Morphological and prosodic alignment of Bulgarian clitics' In Dekkers et al. (eds.), 423-462.

McCarthy, John & Alan Prince (1995). 'Faithfulness and reduplicative identity'. In Beckman, Jill, Wals-Dickie, Laura & Urbanczyk, Suzanne (eds.) *Papers in optimality theory*. Amherst, Mass.: UMass Occasional Papers in Linguistics 18, 249-384.

McCarthy, John (2000). 'Harmonic serialism and parallelism' In Hirotani, Masako et al. (eds.) *Proceedings of NELS 30*. Amherst: Mass.: GSLA, 249-384.

McCarthy, John (ed.) (2004). *Optimality theory in phonology – a reader*. Oxford: Blackwell.

Mohanan, Tara (1994). *Argument structure in Hindi*. Stanford: CSLI Publications.

Müller, Gereon & Sternefeld, Wolfgang (eds.) (2001) *Competition in syntax*. Berlin: Mouton de Gruyter.

Müller, Gereon (1997). 'Partial wh-movement and optimality theory' *The Linguistic Review 14*, 249-306.

Müller, Gereon (2000). *Elemente der optimalitätstheoretischen Syntax*. Tübingen: Stauffenburg.

Müller, Gereon (2002). 'Two types of remnant movement' In Alexiadou, Artemis et al. (eds.) *Dimensions of movement*. Amsterdam: Benjamins, 209-241.

Nakamura, Wataru (1997). *A constraint-based typology of case systems*. PhD Dissertation, State University of New York at Buffalo.

Nakamura, Wataru (1999). 'Optimality-theoretic account of the Japanese case system' *Studies in Language 23*, 597-

Ortmann, Albert (2000). 'Where plural refuses to agree: feature unification and morphological economy' *Acta Linguistica Hungarica 47*, 249-288.

Ortmann, Albert (2002). 'Economy-based splits, constraints, and lexical representations' In Kaufmann, Ingrid & Stiebels, Barbara (eds.) *More than words*. Berlin: Akademie Verlag, 147-177.

Prince, Alan & Paul Smolensky (1993). 'Optimality theory. Constraint interaction in generative grammar' Ms., Rutgers University and University of Colorado at Boulder.

Prince, Alan & Paul Smolensky (2004). 'Optimality theory. Constraint interaction in generative grammar' In McCarthy 2004, 3-71.

Schmidt, Tanja (2002). *West-Germanic IPP-constructions. An optimality theoretic approach*. Ph.D. Dissertation, University of Stuttgart.

Silverstein, Michael (1976). 'Hierarchy of features and ergativity' In Dixon, R.M.W. (ed.) *Grammatical categories in Australian languages*. Canberra: Australian Institute of Aboriginal Studies, 112-171.

Stiebels, Barbara (2000). 'Linker inventories, linking splits and lexical economy' In Stiebels, Barbara & Wunderlich, Dieter (eds.) *Lexicon in Focus*. Berlin: Akademie Verlag, 211-245.

Stiebels, Barbara (2002). *Typologie des Argumentlinkings: Ökonomie und Expressivität*. Berlin: Akademie Verlag.

Tesar, Bruce & Smolensky, Paul (2000). *Learnability in optimality theory*. Cambridge, Mass.; MIT Press.

Woolford, Ellen (2001) 'Case Patterns' In Legendre et al. (ed.), XXX.

Wunderlich, Dieter & Fabri, Ray (1995). 'Minimalist Morphology: An approach to inflection' *Zeitschrift für Sprachwissenschaft 14*, 236-294.

Wunderlich, Dieter & Lakämper, Renate (2001). 'On the interaction of structural and semantic case' *Lingua 111*, 377-418.

Wunderlich, Dieter (1997). 'Cause and the structure of verbs' *Linguistic Inquiry 28*, 27-68.

Wunderlich, Dieter (1999). 'The force of lexical case: German and Icelandic compared' In Wunderlich, Dieter (ed.) *Papers on Argument Linking*. Düsseldorf: Arbeiten des

Sonderforschungsbereichs 282, No.112, 81-102. (Includes, among others, preliminary versions of 'On the interaction of structural and semantic case' 35-61, 63-79,.)

Wunderlich, Dieter (2001). 'How gaps and substitutions can become optimal: An OT account of argument linking in Yimas' *Transactions of the Philological Society 99*, 315-366.

Wunderlich, Dieter (2003). 'Optimal case patterns: German and Icelandic compared' In Brandner, Ellen & Zinsmeister, Heike (eds.) *New Perspectives on Case Theory*. Stanford: CSLI Publications, 331-367.

Zeevat, Henk & Jäger, Gerhard (2002). 'A reinterpretation of syntactic alignment' In de Jongh, D., Zeevat, H. & Nilsenova, M. (eds.), *Proceedings of the 3rd and 4th International Symposium on Language, Logic and Computation*. Amsterdam: ILLC.

Zeevat, Henk & Blutner, Reinhard (eds.) (2004). *Optimality theory and pragmatics*. New York: Palgrave MacMillan.

author's address:

Dieter Wunderlich
Lückhoffstr.9
14129 Berlin
(professor emeritus of Heinrich Heine University Düsseldorf)
email: dieterwdl@aol.com

Postbank Berlin, bank code number 10010010, account no. 194272106.