

The state of the art in OT syntax

By Ralf Vogel

Reviewing Géraldine Legendre, Jane Grimshaw and Sten Vikner's (eds)
Optimality Theoretic Syntax

A number of collections presenting work in optimality theoretic syntax have been published in recent years. *Optimality Theoretic Syntax* (OTS) is special among them in that it brings together most of the important researchers in the field in the late 1990s. It therefore offers a *representative* overview of the state of the art in Optimality Theoretic Syntax at the beginning of the new millenium. Since Prince & Smolensky's pioneering study from 1993, Optimality Theory has become extremely successful and influential within phonology. Generative syntacticians were less enthusiastic about OT. But the constantly growing body of work in OT syntax has, nevertheless, received a certain amount of attention and respect. In my discussion of *OTS*, I will focus on the question to what extent we can speak of OT syntax as a school of its own or a direction within generative grammar, in the same way as we do, for instance, in the case of minimalism, LFG or HPSG.

Apart from an introductory chapter, the other 15 articles in the book are ordered alphabetically by the names of the authors. This choice is a bit unfortunate. It would have been easier for the reader to recognise and compare alternative treatments of the same, or closely related, phenomena, if these had been grouped into the same chapter. Six contributions deal with the syntax, semantics and typology of pronouns. Five papers focus on questions of word order. The typology of case and voice systems is dealt with in three further contributions. Two papers mainly address questions of the general architecture of OT syntax.

Guiding ideas of optimality theory are the assumptions that grammars are constraint-based, that these constraints are violable, and that they are ranked. Grammatical expressions are those which perform better than all possible alternatives in an evaluation based on the constraint hierarchy. An essential part of OT models is the candidate set, a set of possible expressions generated by the generation function GEN. It is impossible to tell the grammaticality of a single expression in isolation, i.e., without reference to the whole candidate set. In this holistic perspective, OT differs fundamentally from other generative

approaches where such 'trans-derivationality' usually plays only a minor role, if any.

Géraldine Legendre remarks in her introductory chapter (chapter 1: "Introduction to Optimality Theory") that "... properties of the input and candidate sets are to a large extent determined by the underlying substantive theory of syntax" (page 20). But on the other hand, "... the question of the substantive nature of syntactic constraints and representations is largely independent of the claims made by OT" (*ibid.*). While, for example, LFG or minimalism can crucially be distinguished by their assumptions about syntactic representations, OT analyses can be formulated on the basis of any representational format. *OTS* reflects this flexibility. Joan Bresnan (chapter 5: "The Emergence of the Unmarked Pronoun"), Hye-Won Choi (chapter 6: "Binding and Discourse Prominence: Reconstruction in 'Focus' Scrambling") and Peter Sells (chapter 12: "Form and Function in the Typology of Grammatical Voice Systems") are well-known syntacticians working within LFG. Most of the other contributions are framed within the Chomskyan branch of generative syntax. But consequences that follow from this difference can hardly be found in the work presented in *OTS*. Choi's paper and the one by João Costa (chapter 7: "The Emergence of the Unmarked Word Order"), who works within the Chomskyan tradition, are remarkably close in their conception of OT syntax. The fact that Costa's candidates contain traces and Choi's do not is a negligible minor difference here.

But we also find a plurality of opinions on other central aspects. One example is what authors assume as the 'input' of a competition – the object that determines the selection/generation of the candidate set. Most of the articles assume a model where the input contains basically semantic information – lexical elements, argument structure, scope specification, information structure. On this basis, a candidate set of syntactic structures is generated. This model has been introduced by Grimshaw (1997) and is widely accepted. It follows a view on syntax familiar in the generative tradition: the OT system chooses the

Title: *Optimality Theoretic Syntax*. Editors: Géraldine Legendre, Jane Grimshaw and Sten Vikner. Publication date: 2001.
Publisher: MIT Press: Cambridge, Mass. ISBN 0-262-62138-X (paper). Price: cloth \$105, paper \$44. Pages: xviii+548 (incl. index).

Ralf Vogel, Universität Potsdam, Institut für Linguistik /
Allgemeine Sprachwissenschaft, Postfach 601553, D-14415
Potsdam, Germany, rvogel@ling.uni-potsdam.de

optimal surface structure for an underlying, 'deep', structure. In this sense, such OT models can still be considered derivational.

Eric Baković & Edward Keer (chapter 4: "Optionality and Ineffability"), henceforth B&K, convincingly argue that this standard conception is unable to account for true syntactic optionality and ineffability in a straightforward way (one could use one of the workarounds that have been developed for such situations, but these do not offer a principled solution). One of the cases they discuss is the complementiser optionality in English subordinate clauses like 'I think (that) the coat doesn't fit him'. The two versions of this clause (with and without 'that') do not differ in any of the semantic properties which are assumed to constitute the input in standard OT syntax. Therefore, one form should block the other in the OT competition. Optionality is unexpected. Non-optionality is in fact one prediction of this architecture: if the input contains only semantic information, then no two different winning, i.e. grammatical, output forms can have the same meaning. Many cases of word order variation are successfully dealt with on this basis, especially those which reflect information structural distinctions. In *OTS*, such accounts are presented by Choi, Costa and Vieri Samek-Lodovici (chapter 11: "Crosslinguistic Typologies in Optimality Theory").

The solution that B&K propose for the problem of semantically identical optional forms makes crucial use of faithfulness constraints. Each of the two optional forms is the winner of its own competition, i.e., they are optimal for different inputs. For this to work, the syntactic difference between the two forms has to be present in the input already: "... in addition to the lexical features, argument structure, tense and aspect of the Grimshavian input, we posit that there are functional features such as [\pm COMP] and [\pm WH]. ..." (p. 99). They note that their input comes close to the numeration in minimalism. In fact, a minimalist analysis solves the optionality problem in exactly the same way: the two optional structures have different numerations.

The function of the faithfulness constraints is that features of the input are preserved in the output, if these constraints are highly ranked. An interesting difference to a minimalist account is that the two optional structures are included in the candidate sets of both competitions. This opens the possibility to solve another notorious problem for OT analyses, ineffability. For instance, optionality of the complementiser breaks down with embedded topicalisation: "I think *(that) on him, no coat looks good". Faithfulness still favours a complementiser-less structure with the appropriate input. But there is a higher ranked markedness constraint, in this case PUREEP ('no adjunction to the highest node of an extended projection'), which prefers the structure with complementiser in both competitions.

Peter Sells argues against such a conception of the input in his article: "... I take it that the input does

not contain a piece of 'passive' information for a Passive output, and an 'antipassive' for Antipassive, and so on. This would seem to rob the OT account of most of its predictive power, by building properties of the intended output into the input." (p. 357). This position appears to me a bit exaggerated. B&K's analysis successfully reconstructs the intuition that the complementiser is optional under 'normal' or unmarked conditions. Embedded topicalisation, for example, creates a marked environment which enforces the unmarked version of the subordinate clause. I do not see a 'loss of predictive power' in such a strategy.

On the contrary, the problem of optionality is not resolved in an account as offered by Sells and, likewise, Judith Aissen (chapter 3: "Markedness and Subject Choice in Optimality Theory"). Among other things, the choice of active and passive in Lummi which depends on the person features of subject and object is modeled by Aissen on the basis of universal prominence and markedness scales. She uses an input specification that is neutral about voice and is related to a candidate set which includes the active and passive version of the clause. For many competitions, the prediction of Aissen's model is that one form blocks the other. This might be desirable for the languages she discusses. But the system presumably predicts ineffability of passive and active in more cases than we actually find. At least in German there is clear optionality of active and passive in the case of topicalisation of a lower argument. Compare "Den Hans hat Maria besucht" ('the-acc H. has M. visited', active, topicalisation) and "Der Hans wurde von Maria besucht" ('the-nom H. was by M. visited', passive). I wonder, how such optionality can be accounted for, if the input is neutral about the morpho-syntactic differences between the two options.

Margaret Speas uses another alternative to the standard model in her contribution (chapter 13: "Constraints on Null Pronouns"). Her input is a complete syntactic structure derived by a minimalist grammar, with copies as traces, features etc. OT regulates the mapping from syntax to the interface levels LF and PF in this model. Speas argues that an OT model of syntax is incomplete without an explanation of the origin of the input structures. This argument can also be turned against B&K's paper, insofar as their input is not a complete syntactic structure, but it is also a problem in minimalism: where does the numeration come from? Speas' solution makes such unanswerable questions superfluous. But restricting OT to interface mapping seems to make a treatment of the issues that Baković & Keer raise impossible. An interesting option would be a combination of her approach with the one by B&K: a faithfulness OT system with complete syntactic structures in both the input and the candidate set.

Peter Ackema & Ad Neeleman (chapter 2: "Competition between Syntax and Morphology"), hence-

forth A&N, analyse compound structures in various languages, in particular, Dutch particle verbs. These verbs are traditionally classified as 'separable compounds'. They behave like single lexical units morphologically and can, for instance, be productively combined with derivational affixes. Thus, the verb 'opmerken' ('up-notice') can be turned into the adjective 'opmerkelijk' ('up-notice-able'). But on the other hand, verb and particle have to be separated in syntax, if the verb is finite and undergoes verb-second movement: 'De onderzoeker (*op)merkte dit feit niet *(op)' ('the researcher noticed this fact not up'). A&N focus on the question whether the compounding of such verbs takes place in syntax or morphology. Their answer is that this decision is made casewise. Syntactic and morphological compounding compete in an OT competition. Syntactic compounding is the default option. Morphological compounding has to be enforced, for example by the presence of derivational affixes like '-lijk'. The input of the OT competition contains the items to be combined. A&N argue that it is the advantage of OT that it offers the tools to handle such flexibility to yield an empirically more adequate account of the phenomenon.

An important difference between this paper and others in OTS is that we here have competition between two *components* of the grammar, syntax and morphology. Usually, candidates that compete are generated by the same rule system. A&N argue that such a competition must proceed locally. It does not evaluate a clause in toto, but only the particular compound. Thus, that competition is separate from the syntax-only competition. The picture of the grammar that emerges is a sequence of OT competitions. This proposal is another example of the usage of OT in handling interface issues. A&N state that a definition of the candidate set in terms of shared lexical material and shared semantics "is impossible in the case of competition between components, because at the moment of competition the overall semantics of the structure is unknown (the morphological output may give rise to a different semantics than the syntactic one)". If I understand it correctly, then the idea of the authors is that a compound either enters syntax as already merged (in the morphology), or the parts of the compound enter syntax separately and will be merged syntactically. A&N claim that this decision has to be taken pre-syntactically. But a single competition OT system that contains all the competing structures and constraints is, to my mind, possible. It only requires a bit more variability in input and candidate set than A&N seem to be willing to allow.

Jane Grimshaw (chapter 8: "Optimal Clitic Positions and the Lexicon in Romance Clitic Systems") draws a non-serial picture of the syntax-morphology interaction. She deals with the clitic systems and the serialisation of clitics in Romance languages. Grimshaw proposes a system of faithfulness and markedness constraints on morphological features that derives clitic inventories, and another system of

alignment constraints that derive the linearisation patterns. Grimshaw then shows how these two systems interact in a single OT system to account for an interesting substitution phenomenon in Spanish where the sequence "le lo" ('to him it') comes out as "se lo" (lit. "SELF it"). This particular case shows an interaction of morphology and syntax that appears to me to be impossible in the model that Ackema & Neeleman propose.

Géraldine Legendre's contribution (chapter 9: "Masked Second-Position Effects and the Linearization of Functional Features") is an interesting discussion of the syntax-phonology interface. She demonstrates that second position effects with finite verbs in Macedonian and Breton can successfully be analysed as being "... a fairly superficial process involving a lot of jostling at PF" (p. 269). The core of the analysis are phonological alignment constraints like NON-INITIAL(x) and EDGEST(x). Languages differ in the degree to which these phonological constraints crucially interact with syntactic and other constraints. This in turn is related to the morphological status of the verb. In Macedonian, finite auxiliaries are clitics, and their relative position is mainly determined by the interaction of phonological alignment constraints, while in Breton finite verbs are independent lexical items and thus are also subject to syntactic constraints. Here we find richer interaction between syntactic and phonological constraints. Legendre's contribution shows that the OT conception of the interfaces allows for much more interdependence among modules than traditional theories of grammar. Constraint interaction and free variation in constraint rankings also predict that this interdependence could be stronger in one language than in others.

The paper by Colin Wilson (chapter 15: "Bidirectional Optimization and the Theory of Anaphora") deals with the syntax-semantics interface. A theory of anaphora binding has to account for two complementary questions: which is the best form (personal or reflexive pronoun) for a given binding relation, and which is the binding relation for a given form? The first question takes the semantics as fixed and seeks the optimal form, the second question takes the form as fixed and seeks the optimal interpretation. Wilson argues for a *bidirectional* model of OT that combines these two perspectives. His candidates and inputs have the same structure, they are form-meaning pairs consisting of a syntactic and a semantic representation. *Interpretive* optimisation defines a competition where the candidates share the form given in the input and vary the meaning, i.e., the binders for the anaphoric expression. *Expressive* optimisation keeps the meaning fixed and varies the form of the bound pronoun.

Wilson proposes that the two optimisations operate in sequence, where interpretive feeds expressive optimisation. This is necessary to avoid the otherwise unescapable prediction of grammars without long binding, in fact, without bound pronouns. This is a consequence of the constraint set that Wilson

assumes. One constraint, REFECON ('referential economy'), prefers reflexives over pronouns, very much like the 'Avoid Pronoun Principle' from Government and Binding theory. If this constraint is undominated and applies in an unrestricted fashion, then the language in question has no pronouns. A second constraint, in fact a family of constraints, LOCAL-ANT[δ] ('local antecedent'), requires anaphors to be bound in domain δ , where δ ranges over syntactic domains like NP and infinitival, subjunctive and finite versions of IP and/or CP. If both of these constraints are active, i.e., ranked highly, then a pronoun in the input will always be mapped onto an anaphor in the output, which will then be interpreted as locally bound. Wilson's serial model avoids this by first computing an *intermediate inventory* by interpretive optimisation. For a given form-meaning pair it is checked whether the form can express the intended meaning. In the following expressive optimisation only those form-meaning pairs compete that won in interpretive optimisation and have the same meaning. This has the effect that in case of local binding anaphor and pronoun compete, while a non-locally bound anaphor is neutralised to local binding, unless faithfulness prohibits this. The effects of REFECON are restricted to those cases where anaphor and pronoun can have the same meaning, ideally cases of local binding.

Though the bidirectional perspective for the syntax-semantics interface itself is reasonable, I wonder whether the particular direction of the serial model could be justified on independent grounds or whether it is only an artefact of the analysis. For instance, the crucial locality constraint LOCALANT is defined in terms of interpretive optimisation: "if X is an anaphor then X is locally bound". An expressive mirror image constraint formulated in the following way is missing, and I wonder why: "if X is locally bound, then X is an anaphor" (LOCALBIND). What LOCALBIND formulates is in fact derived by the interaction of Wilson's constraints. But with LOCALBIND instead of LOCALANT, and a constraint favouring pronouns instead of REFECON we would yield results equivalent to those of Wilson's system if the order of the optimisations is reversed. I wonder how a decision between these two options could be made on independent grounds. Another conceptual worry is that if the serial bidirectional architecture is needed to avoid global application of REFECON, then perhaps only this constraint needs redefinition, and there is no need to blow up the architecture of the model.

Four papers focus mainly on word order phenomena. Sten Vikner (chapter 14: "V-to-I movement and do-Insertion in Optimality Theory") presents an analysis of V-to-I movement. He contrasts French and Icelandic which have V-to-I movement with Danish and English which do not have it. Most of the constraints that he uses are well established within the Principles and Parameters/minimalist framework, like the head movement constraint, the case filter, and economy of movement. The constraints

responsible for the contrast are 'Proper binding' (PRBD), a constraint that requires traces to be properly bound and here essentially bans affix hopping, and 'No Lexical Movement' (LxMv), a special version of economy of movement which punishes movement of lexical categories. Higher rank of LxMv triggers affix hopping from I to V, higher rank of PRBD triggers V-to-I movement. A number of related phenomena are discussed and predicted by Vikner's model. The paper shows that OT syntactic theorising can be closely related to ideas and assumptions that are common within generative syntax, also for the reason that violability and prioritisation of constraints is already implicitly present in much of the earlier generative work.

Joao Costa uses a small set of five constraints to derive the basic and marked word order patterns of transitive clauses in Spanish, Portuguese, Greek, Italian, and a couple of other languages. These constraints are two information structural constraints: ALIGNFOCUS (the rightmost constituent bears focus), TOPFIRST (the topic is clause-initial); two constraints on case, SUBJCASE and OBJCASE, requiring subject and object to occur in their case positions, SpecCP and SpecAgrOP; and the economy constraint on movement, STAY, from Grimshaw (1997). Costa uses rightwards branching structures. All the relevant candidates have V moved to I. The relative rank of SUBJCASE, OBJCASE and STAY derives languages with underlying SVO (SUBJCASE >> STAY), VSO (STAY >> SUBJCASE, OBJCASE) and VOS (OBJCASE >> STAY >> SUBJCASE) orders. High rank of the information structure constraints yields additional orders in special contexts, like, e.g., VOS order with focused subjects in Italian (OBJCASE >> ALIGNFOCUS >> SUBJCASE >> STAY). Small systems of constraints like the one used by Costa can easily be explored in a systematic and transparent way for their typological predictions. This transparency is certainly an advantage of OT.

The paper by Samek-Lodovici is closely related to Costa's, and even more concentrated on the systematic exploration of the typological predictions of a small set of constraints. He uses four constraints: STAY, SUBJECT, which here does more or less the same as Costa's SUBJCASE, ALIGNFOCUSRIGHT, which is similar to Costa's ALIGNFOCUS, but restricted to alignment within VP, and its counterpart, ALIGNFOCUSLEFT, which can be shown to be crucial in some languages. Samek-Lodovici explores the predictions of this constraint system in detail. The article is an excellent demonstration of how typology is done within OT. Especially, Samek-Lodovici spends much effort on showing which kinds of languages are predicted not to exist. The paper can be recommended to anyone who is sceptical about the predictive capabilities of OT models.

Gereon Müller's paper (chapter 10: "Order Preservation, Parallel Movement, and the Emergence of the Unmarked") is an OT treatment of the problem of unmarked and marked orders. His aim is an account

of phenomena of shape conservation. The constraint PARMOV requires c-command relations between elements to be preserved at each level of representation. Müller returns, in this paper, to the GB-model with D-structure, S-structure and LF. PARMOV establishes subject-before-object as a default order, for instance. If objects precede subjects, then in order to fulfil a higher constraint. Such constraints are the Wh-criterion, a pronoun criterion which triggers pronoun fronting in German, and others. An important effect of PARMOV is that movement of multiple elements *for the same reason* is order preserving. One conceptual problem with PARMOV, as Müller himself mentions, is that it is a 'global rule'. It does not regulate individual movement steps. All movement that is not string vacuous necessarily violates PARMOV. It is therefore clear that such an account cannot be translated into the standard minimalist model which has no levels that a constraint like PARMOV could compare.

Though it is not explicitly mentioned, Müller's proposal is related to recent work of Edwin Williams (Williams, 1998, 2003). More sharply than Müller does, Williams formulates his account as an alternative to the minimalist conception of syntax. Shape conservation, the requirement of isomorphic mappings between representations, replaces economy of movement in his model. Interestingly, the constraint STAY appears in Müller's tableaux, but it is ranked lowest in all languages he discusses, and it never rules out a candidate. Instead of the GB-model, Williams has case structure, theta structure, focus structure etc. as representations with potentially conflicting requirements on surface structure. The conflicts are resolved by preferences, like in OT. Williams's 'Representation Theory' can in fact be viewed as another conception of OT syntax. It would be interesting to compare his account to Müller's in more detail. But such a discussion goes beyond the scope of this review.

The conception of markedness plays an important role in many of the papers. Limitations of space do not permit a careful discussion of the papers by Aissen, Bresnan, Sells, and Ellen Woolford (chapter 16: "Case Patterns"). Their common merit is that they show how some typological generalisations which are known from functionalist literature and confirmed by research on many languages of the world can be integrated into

a generative OT model, by using universal and independently motivated prominence and markedness scales as the base for the constraint system.

So what is the state of the art in OT syntax, according to *OTS*? The book demonstrates that OT offers new and promising perspectives on many syntactic phenomena. This is welcome especially in areas where traditional models have notorious weaknesses. Typical phenomena of this kind are those where a multiplicity of factors from diverse sources seem to interact. It is no coincidence that many of the papers in *OTS* deal with interface matters. The conception of constraint interaction offers a wide array of possibilities for the modeling of quite complex interactions between syntax and semantics, morphology, or phonology. It seems that this is an area where syntacticians are most willing to accept an OT perspective.

The concentration on the "syntactic periphery" goes hand in hand with an interesting pluralism on the side of core syntax. Is it not surprising that within OT functionalists, minimalists, LFG- and other grammarians are able to cooperate fruitfully, take up each other's ideas and learn from each other? The diversity of "underlying substantive theories of syntax" that we find in OT syntax might very well turn out to be its major advantage at some point in the future, if this discourse leads to a broad consensus about the ideal way of doing OT syntax. The two to my mind most innovative papers in *OTS* with respect to the architecture of OT syntax are written by PhD students at the OT "headquarters" (Baković & Keer at Rutgers, and Wilson at Johns Hopkins University). Perhaps the development of a genuine OT syntax requires a generation of linguists whose perspective on syntax has not already been pre-shaped by already existing 'schools'.

References

- GRIMSHAW, J. (1997) Projection, heads and optimality. *Linguistic Inquiry*, 28/3, 373–421.
- PRINCE, A. and SMOLENSKY, P. (1993) *Optimality Theory. Constraint Interaction in Generative Grammar*. Rutgers Optimality Archive, ROA-537, <http://roa.rutgers.edu/view.php3?roa=537>.
- WILLIAMS, E. (2003) *Representation Theory*, Cambridge, MA: MIT Press.