OPTIMALITY AND FUNCTIONALITY:
SOME CRITICAL REMARKS ON OT SYNTAX

Frederick J. Newmeyer
University of Washington

fjn@u.washington.edu

July 2000

OPTIMALITY AND FUNCTIONALITY:
SOME CRITICAL REMARKS ON OT SYNTAX

ABSTRACT

This paper presents a critique of optimality-theoretic syntax. First, it questions the frequently encountered claim that an OT approach to a particular problem is 'restrictive', in the sense that it limits the class of possible languages. Such claims, it will be argued, are in principle difficult if not impossible to substantiate. Since constraints are, by hypothesis, universal, there will almost surely be some additional constraint whose interaction with the constraints of the proposed analysis would yield the undesired grammars. Secondly, the paper criticizes the attempt to provide substance to the idea that constraints are universal by requiring for each a functional motivation. While identifying a function for each OT constraint is trivially accomplishable, it is an undesirable 'accomplishment'. Functions and functionally-motivated hierarchies are neither part of grammar nor linked directly to grammatical constructs. Rather, their effect is manifest at the level of language use and language acquisition and (therefore) language change. Finally, OT falters in its treatment of optionality, whether discourse-independent or discourse-dependent. In both cases, it needs to posit otherwise unneeded competition sets and in the latter case it requires the blurring of the distinction between purely syntactic generalizations and those relating syntactic structure to information structure.

OPTIMALITY AND FUNCTIONALITY:
SOME CRITICAL REMARKS ON OT SYNTAX[1]

1. Introduction

The paper begins with some commentary on the fact that optimality theory (OT) is a rare example of a conceptual framework that was designed to handle phonology and has been adopted by syntacticians (§1.1). Section 1 goes on to discuss the appeal of OT approaches to syntax (§1.2) and to outline the limitations of OT that will be treated in detail in the following sections (§1.3).

1.1 Syntax, phonology, and optimality theory

Syntacticians tend to gaze upon the world of phonological theorizing with wonder and envy. Phonologists appear to have their act together in ways that have long been unknown to syntacticians. For one thing, new approach after new approach in phonology seems to rally most of the field behind it. Lexical phonology, metrical phonology, and autosegmental phonology all had the effect of sparking a research program that was endorsed — if not always practiced — by a great majority of the world's phonologists. And now for almost a decade, OT has followed suit. While there are holdouts, of course, and an increasing number of individuals with reservations about the enterprise, it seems accurate to use the expression 'swept the field' to describe the impact of OT in phonology.

Secondly, and even more impressively, these different approaches are not at root incompatible — they simply focus on different aspects of phonological patterning. So there is no intellectual inconsistency, say, in writing a metrical paper one year, an autosegmental the next, and having as one's current research program the task of recasting some of the generalizations in these works in OT terms.

The world of syntactic theorizing could not be more different. Every new approach has tended to reject fundamental aspects of its predecessor. In syntax, rival approaches are just that — for the most part at root incompatible. The move from government-binding to the minimalist program or from GPSG to HPSG entailed abandoning central hypotheses of the earlier models. Only an intellectual schizophrenic or anarchist would write an LFG paper one year and a categorial grammar paper the next, without having had severe qualms about the adequacy of the former and having been intrigued by the possibilities of the

latter. And it goes without saying that there exists a plethora of 'functionalist' and 'cognitive' approaches to syntax that do not share even the questions that formal approaches ask, much less the answers. As a consequence, syntax is in a depressingly fragmented state.

Given the relative degree of harmony among phonologists, it might seem surprising that models of syntax have not borrowed more conceptual apparatus from phonology. Until recently, autolexical syntax was perhaps the only modern syntactic framework whose architecture had phonological roots. If phonologists know something that we (syntacticians) don't know, then why not adapt this knowledge to our own work? The answer to this question is not a mystery — for the most part syntax is simply too different from phonology for any attempt to calque the former on the latter to meet with a great deal of success. Also, given the predominant view that the phonological component takes as its input the output of the syntactic component, phonologists *have to* pay attention to syntax, while syntacticians can — and often do — feel free to ignore the latest developments in phonology.

OT, however, promises to reverse this trend. A significant number of syntacticians have looked to this model — one designed originally to handle phonology — for the purpose of capturing syntactic generalizations. The reason that OT has the potential for winning over syntacticians is that it is, at root, not a theory of phonology per se. Rather, it is a theory governing the interaction of constraints. Where the constraints come from is another question — a question that is independent of whether or not one has chosen the OT approach to constraint-interaction.

Despite its success in phonology, however, OT has not swept the field of syntactic theory. Predictably, perhaps, it has simply increased the fragmentation of an already-fragmented area of investigation. Part of the explanation for that fact derives from the nature of OT itself. It provides no new theory of UG per se, but merely organizes the principles provided by existing theories of UG. So in other words, assessing the adequacy of an OT syntactic analysis involves not just assessing the results of constraint-interaction in that analysis, but also the constraints themselves and, therefore, the theory that provides the constraints. In the past, one could work, say, in the MP, in LFG, or in functionalist syntax. Now, one can work in the MP, with or without an OT orientation, in LFG, with or without an OT orientation, and in functionalist syntax, with or without an OT orientation. The possibilities open to syntacticians are therefore double what they were before the advent of OT.

This paper presents a critical, mainly negative, appraisal of OT syntax. It is intended to make no claim, positive or negative, about OT approaches to phonology. Some of the criticism will clearly transfer to phonological work within OT, and some will not. But it is left to others to argue for that transfer. A basic knowledge of the mechanics of OT will be assumed here; the following section will give a short overview of how OT might work for syntax and discuss

briefly what there is about this approach that some syntacticians have found appealing.

1.2 The appeal of OT syntax

As in OT phonology, constraints in OT syntax are universal and violable. What is language-specific is the ranking of the constraints. The first question to ask, then, is what might compose a candidate set, whose members 'compete'. The answer is simple enough in phonology. The input is the underlying form of the word and the output an annotated phonetic representation. The competition set is an infinite set of forms, though, because of faithfulness constraints, only those bearing some plausible phonological connection to the input have any chance of surfacing as the output. As far as syntax is concerned, it is not totally obvious what competes with what. Following the most cited paper in OT syntax, Grimshaw (1997), the input generally taken to be an extended projection — that is, a lexical head and associated arguments and functional elements belonging to that extended projection. So for a verbal extended projection, the input is a lexical head and its associated argument structure, along with the associated tense and aspect elements. The input is passed to GEN, which generates all extended projections conforming to X-bar principles. The nature of GEN, of course, depends on the framework being assumed — Grimshaw, for example, assumes some version of principles-and-parameters syntax in her work, while Bresnan (1997; 2000a; b) presents an LFG-based approach to OT.[2] So at least in the simplest cases a competition set is made up of phrases with the same heads and dependents (and hence the same LF), but which differ in their arrangement (Legendre, Smolensky, and Wilson 1998, however, argue that competing structures need not have the same LF). In fact, the literature reveals a world of variety with respect to what can constitute a competition set. We find in competition whole sentences, sentence fragments that are not even constituents, different forms of personal pronouns, and so on. Expository convenience accounts for some of this variety, but certainly not for all of it.

Let us examine a simple (and simplified) tableau which demonstrates how OT might account for the presence or absence of overt *Wh*-Movement in a particular language. Two crucial constraints are posited, STAY and OP-SPEC:

(1)   STAY: Trace is forbidden

(2)   OP-SPEC: Syntactic operators must be in specifier position

---

[2] Bresnan (2000a: 21) writes that 'GEN must be universal. That is, the input and the candidate set are the same for all languages. Systematic differences between languages arise from different rankings…' This remark is puzzling, since in all OT work, including Bresnan's own, candidate sets contain language-particular functional morphemes (such as auxiliary *do*), as well as language-particular expletive elements.

Tableau I represents the ordering of STAY and OP-SPEC in a language with overt *Wh*-Movement, Tableau II their ordering in a language without:

	OP-SPEC	STAY
☞ Who$_i$ did you see t$_i$		*
You saw who	*!	

Tableau I
A language with overt *Wh*-Movement

	STAY	OP-SPEC
Who$_i$ did you see t$_i$	*!	
☞ You saw who		*

<div align="center">

Tableau II
A language without overt *Wh*-Movement

</div>

If OP-SPEC » STAY, the candidate with overt *Wh*-Movement is the winner; if STAY » OP-SPEC, it is the candidate without overt movement.

It is easy to see why an OT approach to syntax has won a number of adherents. The idea of violable universal constraints as opposed to parameters rigidly set for a particular language is very seductive. Such an approach allows for flexibility, in that constraints need not be an all-or-nothing affair, and demands precision, in that the exercise of ranking constraints demands more formalization than is typically found in P&P approaches to syntax. Also the possibility of violable universal constraints as opposed to rigid parameters allows for interesting empirical predictions. By hypothesis, all constraints are operative in all languages — the only difference is in their ranking. What this means is that even high-ranked constraints can be violated in a language if a higher one exists and that the effects of low-ranked constraints can become evident, given the right requirements imposed by higher-ranked constraints. Let us examine briefly each possibility.

The first example is of a high-ranked constraint being violated. Costa (1998: 22-24) posits the following two constraints:

(3)   CASE: Move NP to a case-licensing position (subjects to [Spec, IP]; objects to [Spec, AgrOP])

(4)   WEIGHT: Postpose heavy NPs

Now CASE is obviously a high-ranked constraint: its GB ancestor, the Case Filter, was presented as an inviolable universal property of grammars. But notice that it *can* (seemingly) be violated. A sufficiently 'heavy' NP can (or must) move to a position to which Case is not assigned. In other words, WEIGHT appears to outrank CASE in English:

(5)   a. Mary regards John fondly.
        b. ?Mary regards fondly John.
        c. ?Mary regards all the students in her evening class in optimality theory fondly.
        d. Mary regards fondly all the students in her evening class in optimality theory.

Tableau III illustrates:

	WEIGHT	CASE
☞ [[S V [VP Adv t] [Heavy NP]		*
[S V [AgrOP Heavy NP [VP adv t t]]]	*!	

Tableau III
Weight effects dominating the Case requirement in English (Costa 1998)

The second example, that of a low-ranked constraint surfacing, is provided in Grimshaw (1997). STAY is a very low ranked constraint in English, a language which is quite promiscuous in its movement possibilities. Yet Subject-Auxiliary Inversion is blocked in subordinate clauses. Why? Because higher-ranked constraints conspire to prevent movement. The relevant constraints are the following:

(6)   PURITY OF EXTENDED PROJECTION (PURE-EP): No adjunction takes place to the highest node in a subordinate extended projection; and no movement takes place into the highest head of a subordinate extended projection.

(7)   CASE MARKING (CASE): DPs must be Case-marked

(8)   OBLIGATORY HEADS (OB-HD): A projection has a head

(9)   ECONOMY OF MOVEMENT (STAY): Trace is not allowed

Observe in Tableau IV how this ranking conspires to block inversion in subordinate clauses. The first entry violates the constraint against adjunction/movement in subordinate clauses and in the third entry, CASE is violated, since the DP *we* is never in a specifier-head relation with a Case-marking head:

Candidates (subordinate interrogatives)	PURE-EP	CASE	OB-HEAD	STAY
[CP who should_i [IP we e_i [VP see t]]]	*!			**
☞ [CP who e [IP we should [VP see t]]]			*	*
[IP who should [VP we see t]]		*!		*

Tableau IV
The surfacing of a low-ranked constraint (Grimshaw 1997)

1.3   A look ahead

What then might be fundamentally wrong with OT syntax? First of all, it should be stressed that one has every reason to be cautious in criticizing any new approach to any topic. Even the most ardent partisans of OT syntax will acknowledge that there are fundamental issues to be resolved. It would be uncollegial, to say the least, to tear into some novel approach just because there are foundational questions still to be worked out — particularly when other approaches themselves have not necessarily produced elegant analyses of the phenomena that OT has addressed. For that reason, in no case will an inordinate amount of space be devoted to criticizing any one particular analysis within OT, no matter how deficient it might seem. OT syntacticians are intellectual pioneers and therefore should not be 'rewarded' for their risk-taking by being sniped at from the sidelines.

What the paper will attempt to do, then, is to focus on what are arguably problems *inherent* to OT syntax, that is, those that arise from the nature of the research program and would seem to defy solution no matter how many adjustments might be made to this or that facet of that program. Three particularly serious problems are named, identified, and sketched roughly below:

A.   THE RESTRICTIVENESS PROBLEM (§2). One frequently encounters the claim that a particular OT analysis is in some interesting sense 'restrictive', that is, that it follows from the analysis that logically possible, but nonoccurring, language types are excluded. It is argued that such claims tend to be empty.

B.   THE FUNCTIONALITY PROBLEM (§3). There is an interesting attempt to provide substance to the idea that constraints are universal by requiring for each a functional motivation — and in that way to bridge the gap between formal and functional linguistics. It is argued that while identifying a function for each OT constraint is trivially accomplishable, it is an undesirable 'accomplishment'.

C.   THE PROBLEM OF UNMOTIVATED COMPETITION (§4). The view that a grammar consists of forms in competition with each other has created a dynamic necessitating competition sets that fail to respect the autonomy of the different grammatical components. An undesirable consequence is the proliferation of competition sets and the obscuring of important generalizations about grammatical patterning.

Discussion of these three problems will draw from two closely related topics: the first is the interaction of grammar and discourse; the second is the degree to which grammars might be said to be motivated by external functional considerations. Aside from tying in closely with the current research interests of the author of this paper (see Newmeyer 1998b), these topics seem to put OT to the test in ways that a focus on pure syntactic analysis does not.

2.   The restrictiveness problem

It has become standard in the OT syntax literature for a paper in that framework to conclude with words of self-congratulation along the following lines:

> Not only has our analysis provided an elegant treatment of [whatever syntactic phenomenon], but it makes non-trivial predictions as well. As has been demonstrated, no possible ordering of the constraints proposed could yield [whatever ungrammatical sentences]. Hence free constraint ranking admits a much smaller class of languages than might be expected. An OT account thus capable of explaining important facts about the typological distribution of grammatical elements (freely paraphrased from Grimshaw and Samek-Lodovici 1998: 214-215)

Claims of this sort are — in every case that I have examined — without foundation. What they presuppose is that the only relevant constraints are those proposed in the particular research paper. That is, the paper demonstrates that any combination of *those particular constraints* will not yield the targeted impossible forms. But it must be remembered that constraints are universal. If a constraint is necessary for the grammar of *one* language, then, by hypothesis, it exists in the grammars of *all* languages. All it takes is one of these 'additional' constraints, if ranked highly enough, to subvert the claim that one's OT account has explained the nonexistence in principle of some particular set of sentences.

Section 2.1 presents an OT analysis of subject universals, in which a restrictiveness claim is put forward. Section 2.2 illustrates how easily such a claim can be subverted.

2.1   Grimshaw and Samek-Lodovici on subject universals

Let us examine the analysis of subjects in Italian and English in Grimshaw and Samek-Lodovici (1998). Italian has many more possibilities subject-wise than English, given that subjects can be null even in tensed clauses and may occur postposed. Here are the relevant constraints that Grimshaw and Samek-Lodovici propose, presented in ranked order for Italian:

(10)   DROPTOPIC: Leave arguments coreferent with the topic structurally unrealized. [failed by overt constituents which are coreferential with the topic]

(11)   PARSE: Parse input constituents. [failed if an element of the input does not appear in the output]

(12)   ALIGN FOCUS: Align the left edge of focus constituents with the right edge of a maximal projection. [failed by non-aligned foci]

(13)   FULL-INTERPRETATION Parse lexical conceptual structure [failed by expletives and auxiliary *do*]

(14)   SUBJECT: The highest A-specifier in an extended projection must be filled [failed by clauses without a subject in the canonical position]

An interesting result of this particular ordering of these particular constraints is that when the subject is focused, it has to be post-verbal. As Tableau V illustrates, a focused subject can neither be null nor overtly occurring in canonical subject position:

Input: <gridare (x), x=focus, x=Gianni>

Candidates	DROP TOPIC	PARSE	ALIGN FOCUS	FULL -INT	SUBJECT
☞ ha gridato Gianni					*
Gianni ha gridato			*!		
ha gridato,, Gianni			*!		*
ha gridato		*!			*
[*null structure*]		*!*			
expl. ha gridato Gianni				*!	

Tableau V
Italian focused subjects (Grimshaw and Samek-Lodovici 1998)

In English, of course, PARSE is the highest-ranked constraint, with SUBJECT also ranked very high, so a focused overt subject may indeed occur in canonical subject position.

Grimshaw and Samek-Lodovici, however, goes on to make a *universal* claim about subjects, namely that:

> this system of constraints makes the non-trivial prediction that a contrastively focused subject can never be null, i.e., undergo pro-drop. This follows from the fact that a subject will be realized unless it is topic connected and hence under the influence of DROPTOPIC. Its realization will be mandated by PARSE and SUBJECT … This universal prediction … (Grimshaw and Samek-Lodovici 1998: 214)

In other words, languages may have null topics as subjects, but never null contrastively focused subjects.

2.2   Generating null focused subjects

Grimshaw and Samek-Lodovici are correct in their claim that no reranking of the constraints that they propose in their paper will yield null focus in subject position. But why should these be the only relevant constraints? If we add the following constraint to UG, everything changes:

(15)   DROPFOCUS: Leave arguments coreferent with the focus structurally unrealized. [failed by overt constituents which are coreferential with the focus]

If DROPFOCUS is ranked more highly than the other constraints we have looked at, we get just the supposedly prohibited form:

Input: <V (x), x=focus, x=DP>

Candidates	DROP FOCUS	DROP TOPIC	PARSE	ALIGN FOCUS	FULL-INT	SUBJECT
AUX V DP	*!					*
DP AUX V	*!			*!		
AUX V,, DP	*!			*!		*
☞ AUX V			*			*
[*null structure*]			*!*			
expl. AUX V DP	*!				*!	

Tableau VI
How null focused subjects might be derived

In other words, there is a constraint ordering that *does* allow null focused subjects. One's immediate reaction, however, might be to object that DROPFOCUS an absurd constraint. Indeed, one might object that a focused element can *never* be null. But such an objection would be without foundation. Given the right discourse conditions, focus, like topic, can be pragmatically retrievable. The following exchange in English illustrates:

(16)   Q: Who broke the plate?
        A: You know who!

Now imagine a null subject language with the following exchange:

(17)   Q: Who broke the plate?
        A: e broke the plate. (INTERPRETATION: 'You know who broke the plate')

In such a situation, we would have null focus in subject position, presumably licensed by the constraint DROPFOCUS. In fact, we find just such a possibility in Portuguese (Madalena Cruz-Ferreira, personal communication):

(18)   Q: Quem é que partiu o prato?
        A: Partiu o prato?

As long as the reply has question intonation, it can be understood as having a null focus in subject position.
        Serbo-Croatian allows for a similar possibility (Zvjezdana Vrzic, personal communication):

(19)   Q (asked by somebody who most likely did it, but is trying to cover it up):
       Tko je razbio ¶ca¶su?
       Who broke the glass?

A (With an ironic intonation):
Razbio, razbio
Broke, broke (interpretation: 'You know who did it — it was you')

In conclusion, some languages do permit null focused subjects. Grimshaw and Samek-Lodovici are incorrect in claiming both that they *don't* occur and, given the architecture of OT, that they *can't* occur.

The purpose of this section, however, is not to dwell on some particular deficiency in some particular OT syntax paper. Rather, it is to make the general point that given the open-ended number of constraints, one is never on safe ground in claiming that one's particular OT analysis is 'restrictive' in any interesting sense of the term. The limited set of constraints exhibited in the tableau in a particular analysis might well seem to conspire to predict the impossibility of some typologically-dispreferred structure. But since every tableau (when fully expanded) contains every constraint, it seems inevitable that *some* constraint will always lurk in the background which, if ranked highly in some particular grammar, will lead to the generation of that dispreferred structure.

3. The functionality problem

This section deals with the research program that attempts to provide each syntactic constraint of OT with a 'function' and thereby effect a synthesis between formal and functional linguistics. It is argued that such an program is fundamentally wrong-headed. Undoubtedly, some functional motivation can be assigned to any constraint imaginable. However, any theory that assigns to these constraint-function pairings a (synchronic) cognitive reality is both inadequate empirically and brings us no closer to understanding how function helps to shape form as languages evolve through time. Section 3.1 describes this research program, while §3.2 presents a critique of it.

3.1 The proposed synthesis between optimality theory and functional linguistics

The section begins with discussion of the idea that the universality of OT constraints might be rooted in their functional utility, rather than in their putative innateness (§3.1.1). Sections 3.1.2 through 3.1.4 present the attempt to provide an OT-based account of functionally-motivated hierarchies.

3.1.1 On the ontological status of OT constraints

Let us begin by asking how many constraints OT would require to account for the syntax of human language. It seems safe to conclude that the answer is 'thousands'. Only a few dozen or so papers have been published in OT syntax to date, handling only small fragments of only a few dozen languages. Yet well over a hundred constraints have already been proposed. Even for phonology it was noted several years ago that hundreds of markedness constraints alone had been proposed (Blevins 1997: 229). Now, the essence of OT is that constraints are universal. Indeed, such is the very first property of constraints that are mentioned in René Kager's introduction to OT (Kager 1999: 11). OT loses much of its appeal if purely language-particular constraints are allowed. [3]

How might some particular aspect of language get to *be* both universal and an aspect of knowledge of language?[4] There are really not that many possibilities. One is that it could be an instance of innate purely linguistic knowledge. Such has been posited, for example, for the principle of Subjacency in the government-binding theory (Hoekstra and Kooij 1988). Another is that it could be an instance of innate knowledge that is derivative of more general cognitive principles. Haspelmath (1992) considers it plausible that the principles appealed to in grammaticalization research have this property. A third possibility is that the aspect of language in question is learned, but in some sense of the term so 'useful' that all languages will manifest it. An example might be the universal fact that every language has words that refer to the sun and the moon, which presumably arises by virtue of the functional utility of being able to identify the two most prominent celestial bodies.

Interestingly, the standard assumption in OT is that all constraints are innate: 'We propose that the learner start out with an initial ranking of the constraints' (Tesar and Smolensky 1998: 237). But if all (or large percentage) of OT constraints are innate, then we are committed to a degree of innate

---

[3] Along these lines, in their comments on a language-particular *r*-insertion rule proposed in McCarthy (1993), Halle and Idsardi (1997) make the reasonable observation that:

> Conceptually, reliance on arbitrary stipulation that is outside the system of Optimality is equivalent to giving up on the enterprise. Data that cannot be dealt with by OT without recourse to rules are fatal counterexamples to the OT research program. (Halle and Idsardi 1997: 26)

McMahon (2000) argues that many of the allegedly universal constraints of OT phonology are 'effectively language-specific, and that OT cannot in fact function without the addition of parochial rules, or mechanisms which in some way mimic the operations found in traditional derivational models' (ms., p. 10).

[4] Many universals of language of course are not part of linguistic knowledge, but rather an emergent (i.e. epiphenomenal) result of some other principle or principles, which themselves might be innate or learned. For example, it is a universal that no language has prefixes that does not also have suffixes. Speakers do not 'know' this fact; rather, it appears to be an epiphenomenal by-product of the fact that processing considerations favor suffixes over prefixes (Hall 1992).

knowledge that far exceeds anything Chomsky ever dreamed of. Not surprisingly, this fact has led to discomfort among many OT supporters and has engendered a program to ground constraints, directly or indirectly, in their functional utility. Bruce Hayes, an advocate of this program for phonology, takes the position that 'constraints need not necessarily be innate, but only *accessible in some way* to the language learner, perhaps by inductive grounding' (Hayes 1998: 268, emphasis in original).[5] Examples in phonology of potentially inductively grounded constraints might be those prohibiting voiceless segments after nasals or sonorants or prohibiting voiced segments after obstruents.

Recently, we have seen a similar turn in syntax. For example, Haspelmath (1999: 187) suggests that 'the grammatical constraints [of OT] are ultimately based on the constraints on language users' and provides a list of proposed constraints, from both phonology and syntax, that seem to reflect user-functionality. Three of these from the syntactic OT literature are presented in Table I:

name	grammatical constraint	corresponding user constraint
STAY (Grimshaw 1997, Speas 1997)	"Do not move"	Leaving material in canonical position helps the hearer to identify grammatical relationships and reduces processing costs for the speaker
TELEGRAPH (Pesetsky 1998)	"Do not pronounce function words"	Leaving out function words reduces pronunciation costs for the speaker in a way that is minimally disruptive for understanding by the hearer
RECOVERABILITY (Pesetsky 1998)	"A syntactic unit with semantic content must be pronounced unless it has a sufficiently local antecedent"	Omitting a meaning-bearing element in pronunciation makes the hearer's task of extracting the intended meaning from the speech signal very difficult unless it can be inferred from the context

Table I

Some OT constraints and their possible functional motivations (Haspelmath 1999: 185)

---

[5] According, to Hayes, 'the formal system of grammar characteristically reflects principles of good design' (p. 276). But he is clear that not all 'good design' is necessarily 'functional' — for example he feels that the tendency of phonological systems toward symmetry has no evident functional explanation.

Joan Bresnan has in a number of publications also expressed the idea that constraints are functionally-motivated:

> … both phonologists and functional linguists have recognized that linguistic inventories also reflect universal patterns of markedness and are often functionally motivated by perceptual and cognitive constraints. I will argue in support of this conclusion by showing how different inventories of personal pronouns across languages may be formally derived by the prioritizing of motivated constraints in Optimality Theory (Bresnan 1997: ms, p. 1)

In the remainder of this paper the position that the constraints of OT have functionally-motivated user-constraints associated with them will be referred to as 'functionally-based optimality theory', or 'FOT'.

3.1.2   Functionally-motivated hierarchies

The most ambitious attempt to meld OT with the results of functionalist theorizing is the direct importation into OT of hierarchies that appear to have some functional motivation. Historically, the first such hierarchy proposed was the Thematic Hierarchy, one version of which is presented in (20):

(20)   **Thematic Hierarchy** (Bresnan and Kanerva 1989: 23): Agent > Beneficiary > Recipient/ Experiencer > Instrumental > Theme/ Patient > Location

Fillmore (1968) argued that one can predict subject choice in English and other languages by reference to the position on such a hierarchy of the thematic role (or, as he called it, the 'case' role) borne by the NP. Hierarchies of thematic-roles have been appealed to in the explanation of such diverse phenomena as antecedence for reflexivization (Jackendoff 1972), argument structure in morphology (Carrier-Duncan 1985), and the choice of controller in embedded infinitivals (Culicover and Jackendoff 1999).

Grammatical relations (or functions) have also been argued to be hierarchically organized. For example, Keenan and Comrie (1977) suggested that we can explain interesting cross-linguistic facts about relativization possibilities by appealing to the following hierarchy:

(21)   **Relational Hierarchy** (Comrie and Keenan 1979: 650): Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparative

Another hierarchy of long-standing is the Animacy Hierarchy (which is often treated separately from its proper sub-hierarchy, the Person Hierarchy).

Silverstein (1976) was possibly the first to notice that a wide variety of grammatical processes in a number of languages, in particular those involving word order, grammatical relation choice, and case marking, seem to be sensitive to the relative degree of animacy of the noun phrases involved. Here are more recent versions of both hierarchies:

(22)   **Animacy Hierarchy** (Dixon 1979: 85): 1st Person Pronoun > 2nd Person Pronoun > 3rd Person Pronoun > Proper Noun > Human Common Noun > Animate Common Noun > Inanimate Common Noun

(23)   **Person Hierarchy** (Croft 1990: 149): Local (1st, 2nd) > 3rd

A related hierarchy pertains to the degree of 'identifiability' of the participants in a discourse and is formulated as the Definiteness Hierarchy:

(24)   **Definiteness Hierarchy** (Comrie 1989: 134-136): Personal Pronoun > Proper Noun > Definite NP > Indefinite Specific NP > Non-specific NP

Finally, a number of studies have pointed to the relative 'discourse prominence' of the phrases in an utterance as a factor in grammatical patterning, where the more 'discourse prominent' element is, in some sense, more at the center of attention of the participants of the discourse. Hence, the following hierarchy:

(25)   **Prominence Hierarchy** (Tomlin 1985): Discourse Prominent > Not Discourse Prominent

The above hierarchies have entered into functionalist theorizing in two separate ways. First, all of the hierarchies, to one degree or another, appear to have functional (or at least external) motivation, in that they are based on facts about language users. Intuitively, at least, all seem to arrange themselves in a decreasing degree of cognitive salience or importance. Human NPs are more central to the human experience than inanimates; agents are more central than instruments; subjects are more central than objects of prepositions; and so on. Second, and more importantly, functionalists formulate grammatical generalizations by means of direct reference to these hierarchies. In the view of Croft (1995), all universal aspects of language are functionally-motivated and representable by implicational hierarchies such as those described above. Grammars of individual languages, in this view, are (to simplify only a little) collections of statements about where particular phenomena fall on particular hierarchies.

3.1.3   Aissen on subject choice

The functionalist program that accords central place to hierarchies of thematicity, animacy, and so on is given an OT realization in two papers by Judith Aissen (Aissen 1999; 2000). The first, which deals with subject choice, is discussed in this sub-section; the second, which treats differential object marking, is discussed in the next.

Languages differ with respect to the degree to which subjects and objects must obey criteria of personhood, thematic role, and prominence. The goal of Aissen (1999) is to capture these different realizations in an optimality-theoretic framework. Aissen, in fact, works with only a proper subset of the relevant hierarchies, as follows:

(26) **Relational:** Subject > Object > Oblique
(27) **Person**: Local (1st, 2nd) > 3rd
(28) **Thematic Role:** Agent > Patient
(29) **Prominence:** Discourse Prominent (X) > Not Discourse Prominent (x)

Following the technique proposed in Prince and Smolensky (1993), she aligns these hierarchies harmonically, interleaving them so that the ranking within each is preserved. Thus we predict that subjects are more likely to be local than 3rd person, more likely to be agents than patients, and more likely to be discourse prominent than not to be. The actual constraints take the form of the prohibition against some grammatical relation being linked to some element on one of the other hierarchies. Hence, they are stated as follows:[6]

(30) a. *Su/Pat
     b. *Su/x
     c. *Obl/Local

In other words, we have universal violable constraints prohibiting subjects from being patients (30a), prohibiting subjects from being non-discourse prominent (30b), and prohibiting oblique forms from being first or second person (30c).

Since the ranking within each hierarchy is preserved under alignment, a ranking such as that depicted in (31a) will be allowed, but not one such as in (31b):

(31) a. *Su/Pat » *Su/Ag
     b. *Su/Ag » *Su/Pat

---

[6] Aissen's alignments always involve the Relational Hierarchy and one other hierarchy. Aligning the other hierarchies with each other would predict peculiar constraints like *Ag/3 » *Pat/3, i. e., one ranking prohibitions against third person agents more highly than the prohibition against third person patients. It is not clear to me what would *prevent* aligning, say, the Thematic Hierarchy with the Person Hierarchy, leading to just such peculiarities.

Agents outrank Patients in the Thematic Hierarchy, so a constraint prohibiting subject Agents cannot outrank a constraint prohibiting subject Patients.[7]

Let us now turn to some of Aissen's particular analyses, beginning with English. The highest ranked relevant constraint is the one forbidding non-prominent subjects. Notice that in Tableau VII, that constraint dominates the one forbidding patient subjects (i. e., *Su/x » *Su/Pat). Passives can be generated, then, if the Patient is prominent and the Agent is non-prominent.

V(Agt/3/x,Pat/1/X)	*Su/x	*Su/Pat	*GR/Pers
ACTIVE Agt/Su/3/x-Pat/Oj/1/X	*!		**
☞ PASSIVE Pat/Su/1/X-Agt/Obl/3/x		*	**

Tableau VII
English (prominent patient) (Aissen 1999: 689)

Fox has no passives at all, a result that Aissen obtains by a high ranking of the prohibition against patient subjects (i.e., *Su/Pat » *Su/x).

V (Agt/3, Pat/1)	*Su/Pat	*GR/Pers
☞ ACTIVE Agt/Su/3-Pat/Oj/1		**
PASSIVE Pat/Su/1-Agt/Obl/3	*!	

Tableau VIII
Fox (Aissen 1999: 687)

Other languages restrict actives and passives to particular persons. Lushootseed, for example, does not allow 1st or 2nd person passive agents:

Agt ⇓ Pat ⇒	1	2	3
1	-	active/*passive	active/*passive
2	active/*passive	-	active/*passive
3	active/passive	active/passive	active/passive

Table II

---

[7] It is important to stress that from the fact that *Su/Ag never outranks *Su/Pat, it does not follow that no language allows agent objects. Such languages indeed exist. As Aissen notes (p. 686): 'Language in which agents are realized as object … will only emerge if there are higher-ranked constraints which force the appearance of object agents'. As noted in §2.2 of this paper, the structure of OT is not conducive to any successful claim that one's analysis restricts the set of possible languages.

Distribution of voice by person in Lushootseed (Aissen 1999: 690-691)

Aissen derives this result by means of a high ranking of *Obl/Local:

V(Agt/1/x, Pat/3/X)	*Obl/Local	*Su/x	*Su/Pat	*GR/Pers
☞ACTIVE Agt/Su/1/x-Pat/Oj/3/X		*		**
PASSIVE Pat/Su/3/X-Agt/Obl/1/x	*!		*	*

Tableau IX
Lushootseed (Aissen 1999: 691)

Lummi is like Lushootseed, except that *actives* are disallowed when the subject is 3rd person and the object is 1st or 2nd (Jelinek and Demers 1983):

Agt ⇓ Pat ⇒	1	2	3
1	-	active/*passive	active/*passive
2	active/*passive	-	active/*passive
3	*active/passive	*active/passive	active/passive

Table III
Distribution of voice by person in Lummi (Aissen 1999: 692)

Aissen derives this generalization by means of a high ranking of *Oj/Local:

V(Agt/3/X, Pat/1/x)	*Obl/Local	*Oj/Local	*Su/x	*Su/Pat	*GR/Pers
ACTIVE Agt/Su/3/X-Pat/Oj/1/x		*!			*
☞PASSIVE Pat/Su/1/x-Agt/Obl/3/X			*	*	**

Tableau X
Lummi (1st person patient) (Aissen 1999: 692)

3.1.4   Aissen on differential object marking

Aissen (2000) provides an OT account of the fact that some languages overtly case mark direct objects and some do not and that whether they do or not is often a function of the degree of animacy or definiteness of that object. In general, the higher in the hierarchy of animacy and/or definiteness the object is, the more likely it is to be case marked. Aissen employs the following hierarchies in this paper:

(32)   **Animacy:** Human > Animate > Inanimate
(33)   **Definiteness:** Personal Pronoun > Proper Noun > Definite NP > Indefinite Specific NP > Non-specific NP

As Aissen notes, the correlation of overt marking with animacy and/or definiteness appears to have a functional explanation: 'It is those direct objects which are most in need of being distinguished from subjects that are overtly case marked' (ms., p. 3).

Employing the technique of harmonic alignment discussed in the previous section, Aissen arrives at the following constraint rankings:

(34)   a. *Oj/Hum » *Oj/Anim » *Oj/Inam
        b. *Oj/Pro » *Oj/PN » *Oj/Def » *Oj/Spec » *Oj/NSpec

In other words, no language will prohibit inanimate objects that does not prohibit human objects; no language will prohibit non-specific objects that does not prohibit pronominal objects; an so on.

Now, then, the goal is to ensure that if in a particular language a direct object with a particular value for animacy or definiteness is case marked, then objects with animacy or definiteness values higher on their respective hierarchies will also be case marked. Aissen first steps to achieving this result are to propose constraint (35) and locally conjoining it with constraints of (34a-b) (see Table IV):

(35)   *$\emptyset_C$ 'Star Zero': Penalizes the absence of a value for the feature CASE.

Local conjunction of *$\emptyset_C$ with the subhierarchy on object animacy (34a)	Local conjunction of *$\emptyset_C$ with the subhierarchy on object definiteness (34b)
*Oj/Hum & *$\emptyset_C$ » *Oj/Anim & *$\emptyset_C$ » *Oj/Inam & *$\emptyset_C$	*Oj/Pro & *$\emptyset_C$ » *Oj/PN & *$\emptyset_C$ » *Oj/Def & *$\emptyset_C$ » *Oj/Spec & *$\emptyset_C$ » *Oj/NSpec & *$\emptyset_C$

Table IV

Local conjunction of *ø$_C$ with object-oriented subhierarchies (Aissen 2000: ms., p. 9)

If nothing more were said, the constraints of Table IV would force case on *all* objects. Therefore, the next step is to propose a constraint that is, in essence, the inverse of (35), namely one that penalizes the *presence* of morphological case:

(36)   *STRUC$_C$: penalizes a value for the morphological category CASE

Aissen is now in a position to propose language-particular tableaux capturing the object case marking generalizations for each language. Consider Hebrew and Turkish. In Hebrew, all and only definite objects (i.e. personal pronouns, proper names, and other definites) are case marked. In Turkish, these objects are case marked, and specifics as well. This result is achieved by the relative ranking of the locally conjoined constraints of Table IV and constraint (36). The competition sets consist of a case marked object and a non case marked object of the same degree of definiteness. Tableaux XI and XII illustrate.

Hebrew specific, indefinite patient

Patient: specific indefinite	*Oj/Def & *ø$_C$	*STRUCC	*Oj/Spec & *øC	*Oj/NSpec & *øC
Oj: specific indefinite CASE: ACC		*!		
? Oj: specific indefinite CASE:			*	*

Tableau XI
Hebrew specific indefinite patients (Aissen 2000: ms., p. 14)

Turkish specific, indefinite patient

Patient: specific indefinite	*Oj/Def & *øC	*Oj/Spec & *øC	*STRUC C	*Oj/NSpec & *øC
? Oj: specific indefinite CASE: ACC			*	
Oj: specific indefinite CASE:		*!		*

Tableau XII
Turkish specific indefinite patients (Aissen 2000: ms., p. 15)

In other words, since in Hebrew the prohibition against case marking dominates the constraint prohibiting non-case marked specifics, it follows that that language will disallow case marked specifics. In Turkish, on the other hand, the prohibition against case marking is dominated by the constraint prohibiting non-

case marked specifics. Therefore, case marked specific objects are allowed in that language.

Finally, it should be pointed out that Aissen's account of differential object marking is much more overtly 'functionalist' than her earlier paper on subject choice. Not only does the former incorporate functionally-motivated hierarchies (as does the latter), but it attempts to provide functional motivation for the constraints themselves. For example, constraint (35), penalizing the absence of case marking, is attributed to the listener-oriented principle 'minimization of perceptual confusion'. The constraints of Table IV are regarded as being rooted in 'iconicity', since they favor morphological marks for marked configurations. Constraint (36), on the other hand, is said to have an economy-based motivation, since it reduces the amount of structure that needs to be processed.

3.2.   The failure of functionally-based optimality theory

Any linguist committed to a rapprochement between formal and functional linguistics will find in the FOT synthesis advocated Haspelmath, Bresnan, and Aissen an immediate intuitive appeal. If this synthesis were able to resolve the most bitter and long-standing division in the field of linguistics, that would be a marvelous result. However, upon close examination, FOT fails to bridge the gap between the two opposing tendencies. The remainder of this section will document that the problems such an approach faces. Section 3.2.1 raises some conceptual difficulties inherent in the idea that each constraint has a functional motivation. Section 3.2.2 takes the position that there is no direct relationship between the constructs that make up particular grammars and functional motivations for those constructs. And §3.2.3 argues that grammars do not contain functionally-motivated hierarchies.

3.2.1   On constraints and functional motivations

3.2.1.1 The content of the claim that all constraints are motivated functionally

How strong is the claim that every constraint has a functional motivation? Perhaps in phonology there is real content to such a claim, but this is unlikely to be the case for syntax. Nobody understands or, in the foreseeable future, is likely to understand the full set of external factors that might combine to account for the properties of syntactic structure. The functionalist literature has mooted

---

[8] Differential case marking with respect to animacy is handled in analogous fashion. Aissen goes on to discuss the handling of 'two-dimensional' differential object marking, in which marking is sensitive to *both* definiteness and animateness. Given the considerable space that would be required to present her treatment and the fact that any critique of her approach to one-dimensional marking carries over ipso facto to that of two-dimensional, the latter will not be discussed here.

literally dozens of potential factors, ranging all the way from economy, iconicity, and parsing to love of metaphor and 'playfulness'. In short, even the plausible external motivations are so numerous, so diverse, and so open-ended that any conceivable rule or constraint in any framework could be provided with one. In the words of Gereon Müller (1999a: 232), 'it is usually not hard to contrive some functional motivation for almost any given constraint'.

To illustrate the ease with which any constraint might be 'motivated functionally', let us consider the two constraints that Haspelmath (1999: 186) writes would be 'the first candidates for elimination' if OT constraints had to be rephrased in terms of user-optimality:

(37)   NO LEXICAL HEAD MOVEMENT: A lexical head cannot move

(38)   LEFT EDGE (CP): The first pronounced word in CP is a function word related to the main verb of that CP

It does not seem to be too much of a challenge to provide NO LEXICAL HEAD MOVEMENT with a functional motivation. An iconicity requirement presumably keeps heads and dependents straying too far from each other in order to facilitate semantic interpretation.

LEFT EDGE (CP) is instantiated by the mandatory presence of the *que* complementizer in French and that of *for* initiating English infinitival complements:

(39)   a. Je crois que Pierre a faim.
         b. *Je crois Pierre a faim.
(40)   a. I'm looking for a book to read.
         b. *I'm looking a book to read.

Here too there is a plausible functional motivation. LEFT EDGE (CP) aids the hearer by providing explicit information as to clause boundaries, as well as typing the clause as finite or nonfinite. In other words this constraint plays a role in semantic interpretation. Hence we could add the material in Table Ia below to Table I.[10]

---

[9] Müller, however, goes on to give examples of constraints which seem to lack functional motivation (see below, footnote 11).

[10] Haspelmath says that it is the lack of generality of these two constraints that led him to regard them as lacking corresponding user constraints. He writes in criticism of Table Ia:

> Your constructed functional motivations for these two constraints don't strike me as particularly felicitous. Of course, NO MOVEMENT (=STAY) is functionally motivated by iconicity, but why HEAD, and why LEXICAL? Why not NO FUNCTIONAL MAXIMAL PROJECTION MOVEMENT? And similarly, why not RIGHT EDGE (CP)? (Quite

name	grammatical constraint	corresponding user constraint
NO LEXICAL HEAD MOVEMENT (Grimshaw 1997)	"A lexical head cannot move"	Moving a lexical head away from its dependents makes it difficult for the hearer to pair heads and dependents
LEFT EDGE (CP) (Pesetsky 1998)	"The first pronounced word in CP is a function word related to the main verb of that CP"	Not explicitly marking embedded clauses in terms of their boundaries and finiteness impedes rapid recognition of their meaning and role

Table Ia
Possible functions for NO LEXICAL HEAD MOVEMENT and LEFT EDGE (CP)

By way of illustrating the ease with which any conceivable grammatical construct might be provided with some functional motivation, consider the transformational rules of pre-principles-and-parameters approaches. Even though they formed the target of functionalist assault for their abstractness and seeming divorce from anything that might be considered user-based (see, for example, Givón 1979), they too were argued by certain linguists to have functional motivations. For example, Langacker (1974) classified transformational rules in terms of whether they raised, fronted, or backed grammatical elements and claimed that each formal operation was designed to facilitate a particular discourse function (see also Creider 1979).

3.2.1.2 On correlations between rankings and functions

An advocate of FOT might suggest that there is, in fact, a way of supplying empirical content to the idea that constraints need to be grounded functionally. Perhaps one might expect some direct correlation between the functionality of a particular constraint and its ranking in the grammar of an individual language and/or its typical cross-linguistic ranking. That is, the better motivated the constraint functionally, the higher in general we would expect its ranking to be.

---

generally, asymmetrical alignment constraints cannot be functionally motivated.)
(personal communication)

Haspelmath's point that the constraints of Table Ia lack generality compared to those of Table I is certainly correct. But the reasonable conclusion from that fact is that the functional motivations for the former are weaker than those for the latter, not that the former have no functional motivations at all.

Such a suggestion presents several problems, however. First, given the present limited state of our knowledge of what external factors are the best motivators of syntactic structure, we have no non-circular way of ranking functions. For example, we might be tempted to say that faithfulness constraints tend to be more highly ranked than markedness constraints because they play a more important function. But our only evidence for their serving a more important function is that they are more highly ranked. There exists no theory of functionality from which it follows that faithfulness is more important than markedness.

Even in phonology, where functionality is better understood than in syntax, there is no clear relationship between the importance of the function that a constraint serves and its typical ranking. For example, Hayes (1998) devises an algorithm for showing how some constraints are more phonetically grounded than others, but does not go on to demonstrate how this translates into cross-linguistic constraint rankings. As far as I know, nobody has produced a successful demonstration along those lines.

Let us turn now to the second problem with attempting to correlate ranking-strength with function-strength. Consider two commonly discussed functional motivations — iconicity and parsing. Here are their typical effects:

(41)   a. Iconicity: Syntactic constituents reflect semantic units.
        b. Parsing: Heavy constituents follow light constituents.

Iconicity can be illustrated by the fact that adjectives are generated under the same phrasal node as the noun that they modify. Parsing can be illustrated by the fact that within the verb phrase, heavier sentential complements are positioned after lighter phrasal complements. Now, let us ask which functional motivation leads to higher-ranked constraints for English. The answer is: 'It depends'. In some cases, we have identical grammatical elements in variant orders with no meaning difference, each option corresponding to a different ranking of the relevant constraints. For example, both of the following are grammatical English sentences:

(42)   a. A man who was wearing a silly-looking red hat dropped by today.
        b. A man dropped by today who was wearing a silly-looking red hat.

Sentence (42a) reflects a ranking of an iconicity-dominant constraint over a parsing-dominant one; sentence (42b) the reverse ranking. In some cases, however, only a higher ranking of an iconicity-based constraint is possible. Simple adjective phrases cannot be extraposed from the nouns that they modify, no matter how heavy they are:

(43)   a. An extremely peculiar-looking man dropped by today.

b. *A man dropped by today extremely peculiar-looking.

And in other cases, parsing effects seem to be more important than iconicity effects. When comparatives are used attributively, the adjective is separated from its complement by the head noun, despite the fact that together they serve to modify semantically that head noun (44a-b). Their structural unity can be obtained only in a manner that is consistent with parsing pressure (44c):

(44)   a. That's a more boring book than any I have ever read.
       b. *That's a more boring than I have ever read book.
       c. That's a book more boring than any I have ever read.

So we must conclude that constraints with different functional motivations are 'interleaved' within the grammar — a little iconicity here, a little parsing there, and so on. There is no hope for correlating the strength of a particular functional motivation with the ranking of the constraint that it putatively drives.

There is a third, even more severe, problem with any attempt to correlate the functionality of a constraint with its ranking. In a nutshell, to ask the question: 'Is this constraint functionally motivated?' is to ask the *wrong* question. No rule or constraint has a functional motivation in and of itself, but rather only within the total system in which it occurs. To illustrate, let us consider the two constraints OP-SPEC (repeated below as (45)) and HEAD-RIGHT (46):

(45)   OP-SPEC: Syntactic operators must be in specifier position

(46)   HEAD-RIGHT: Heads uniformly follow their complements and adjuncts

OP-SPEC, when highly ranked, 'forces' *Wh*-Movement. Is this constraint functionally motivated or not? It seems to be, since operators in specifier position play a useful role both in marking scope and reserving a 'special' position for focused elements.[11] What about HEAD-RIGHT? Given that there is a parsing advantage for all heads being on same side of their complements (Hawkins 1994), then this constraint is also functionally motivated. But what is *dysfunctional*, however, is for any language to rank them *both* highly. Most head-final languages do not have *Wh*-Movement and there is a good reason for that. As Hawkins (1995) has observed, any movement of an argument away from its subcategorized position creates the danger of temporary ambiguity. If the verb comes early (as in VO languages), the danger is minimized, since upon hearing

---

[11]Müller (1999a), however, argues that OP-SPEC is *not* functionally motivated. As he points out, *wh*-phrases are not necessarily (semantic) operators, some cases of *Wh*-Movement have to be partly undone semantically, all *wh*-phrases can in principle be interpreted in situ, and *wh*-phrases may respect OP-SPEC by undergoing partial movement to an embedded non-scope position. It is not clear how an advocate of FOT would respond to these points.

the verb the hearer knows which and how many arguments to expect. But in OV languages, the verb comes *after* an object gap, thereby delaying information about the content of that gap. OV languages solve this problem in part by making arguments 'toe the line' — *Wh*-Movement is rare and A-movements are in general disfavored.

In other words, all other things being equal, in grammars in which OP-SPEC is a high-ranked constraint, HEAD-RIGHT is a low-ranked one. But these are just two constraints out of, presumably, thousands. For any pair — or triple, or quadruple, etc., etc. — of constraints one can ask the degree to which that association of constraints is a 'functional' one. After all, only the most ardent anti-structuralist would deny that grammars are tightly-integrated wholes. We are thus led to the conclusion that there is no way of correlating the functionality of a particular constraint with its mean cross-linguistic parochial ranking.[12]

3.2.2. Constraint functionality and speaker knowledge

We have seen that there is little hope in correlating the functionality of a constraint with its ranking. But this section will defend an even stronger position, namely that functional motivations are neither part of grammar nor linked to elements of grammar. Rather, they are wholly external to grammar and make their presence felt in language use and acquisition — and therefore in language change. FOT errs in advocating a tight constraint-function pairing.

Let us begin with a couple simple questions about the grammar of Mary Miller, a native speaker of English. One is: 'Why do subjects precede objects?' The other is: 'Why aren't there null subjects?' We could supply very functionalist-sounding answers to those questions: 'Subjects precede objects because they have cognitive prominence over objects and cognitive prominence is iconically represented'; and 'There are no null subjects because agreement is too weak to license them'.

But those are the wrong answers. Mary Miller's grammar has those properties because the grammars of her parents and peers have them. Except in unusual historical circumstances, one's grammar reflects to an extremely high degree the grammars of those in one's speech community. The factor that best explains why a person's grammar has the properties that it has is *conventionality*. Grammars differ only slightly from generation to generation. As noted by

---

[12] David Odden (personal communication) has suggested that local conjunction of constraints might provide a way of insuring the ranking dissociation of OP-SPEC and HEAD-RIGHT. For example, if OP-SPEC&HEAD-LEFT universally dominated OP-SPEC&HEAD-RIGHT, then we would predict the (general) absence of *Wh*-Movement in verb-final languages. Such a solution is well worth exploring. My main reservation about such an approach comes from the fact that literally dozens of properties correlate with direction of headedness and therefore (presumably) dozens of constraints would need to be locally conjoined. The net result of such an approach would be the explosion in the number of constraints needed by several orders of magnitude.

William Croft (1995: 522), this stability in a sense has a functional motivation, since it is rooted in mental routinization and social convention. Along these lines, Croft has recently made the perspicacious observation that:

> … a central aspect of a speaker's use of language is convention. When I say *Who did you meet yesterday?*, I put the interrogative pronoun *Who* at the beginning of the sentence because that is the convention of my speech community. I know the conventions of my speech community, and my use of language will serve its purpose best most of the time if I conform to the conventions of my speech community. It may be that the initial position of *Who* is partly motivated by pragmatic universals of information structure, or partly specified by an innate Universal Grammar. In fact, one (or both) of those factors may be the motivation for the origin of the convention. But that is not why I have put it there in that utterance. (Croft 2000: ms., pp. 9-10)

'Conforming to the conventions of one's speech community' is not, of course, the sort of functional motivation that has been claimed to underlie constraints. Models of grammar such as FOT that see constraints as being tied synchronically to motivations such as parsing and iconicity are thus empirically off-base.

Grammars do reflect the effects of motivations such as parsing pressure and pressure towards iconicity, of course. But these effects make themselves felt at the level of language change, via the acquisition process, and are not 'registered' internally to the grammars themselves. (This point is made forcefully with respect to OT phonology and phonology in general in Hale and Reiss 2000 and Buckley 2000). In a nutshell, the forces (functional or otherwise) that bring a construction into a language are not necessarily the same ones that keep it there. To give one example in support of this claim, consider the Modern English genitive. It may either precede or follow the noun it modifies:

(47)   a. GEN-N: Mary's mother's uncle's lawyer
        b. N-GEN: the leg of the table

The GEN-N ordering is unexpected, since English is otherwise almost wholly a right-branching language. So why do English-speaking children acquire the GEN-N ordering? The short — and 100% correct answer — is 'conventionality'. They learn that ordering because they detect it in the ambient language of their speech community. But the long answer is very interesting and drives home the great divide between the functional explanation of a grammatical change and force of conventionality that leads to the preservation of the effects of that change.

Old English 1000 years ago was largely left-branching with dominant orders of OV and GEN-N.[13] This is the correlation motivated by parsing efficiency. The shift to VO order in the Middle English period was matched by a shift to N-GEN order. A text count of 85% N-GEN has been reported for Middle English in Kirby (1998) and Fischer (1992).[14] We don't know details of why this happened. Lightfoot (1991) suggests that as tokens of VO order in main clauses increased, cues that English was OV declined, leading English to be reanalyzed as VO underlyingly. But then, after a certain time, everything started to reverse itself, with the text count of GEN-N order increasing dramatically. Why did this reversal occur? According to Kroch (1994), it may have been a result of the two genitives becoming 'functionally differentiated'. The GEN-N construction became favored for animates while the N-GEN construction has tended to be reserved for inanimates (see also Wedgwood 1995 and Kirby 1998).

Now, then, what would the relation be between the OT constraints that license these two orders in Modern English and the functional motivations that gave rise to them? The answer is that it is so indirect as to be uninteresting. The current state of the English genitive is a product of over a thousand years of changes, many functionally motivated when they occurred, but preserved in the language primarily by the force of conventionality. Yes, it was undoubtedly parsing pressure that led Old English to be predominately GEN-N. That pressure no longer exists, but the order does. If the need for 'functional differentiation' is part of the explanation for why that order was preserved, one challenges any advocate of FOT to demonstrate that that particular functional force is a motivating factor in the grammars of English speakers today and to identify the particular constraints to which this factor is linked. Among other problems that would need to be addressed is the fact that the functional differentiation is only partial. That is, inanimates can occur in the GEN-N construction (*the table's leg* is not horribly unacceptable) and animates can occur in the N-GEN construction (*that book of Mary's*).

The point is that languages are filled with structures that arose in the course of history to respond to some functional pressure, but, as the language as a whole changed, ceased to be very good responses to that original pressure. Such facts are challenging to any theory like FOT, in which the sentences of a language are said to be a product of constraints that must be functionally motivated.

Hale and Reiss (2000) and Buckley (2000) point out that FOT confounds what we know with how what we know *got to be* what we know. Parsing ease, desire for functional differentiation, pressure for an iconic relationship between

---

[13] Lightfoot (1999: 117) notes that in limited circumstances, N-GEN order and split genitive order were possible as well.

[14] Lightfoot (1999: 119) gives examples of split genitives in Middle English. Unlike in Old English, the element to the right in the split genitive does not have to bear a thematic relation with the element to the left, a state of affairs that Lightfoot ties to Case theory.

form and meaning, and so on are indeed forces that shape grammars. These forces influence adult speakers, in their use of language, to produce variant forms consistent with them. Children, in the process of acquisition, hear these variant forms and grammaticalize them. In that way, over time, certain functional influences leave their mark on grammars. There is no place — indeed no *need* — for the functional forces to match up in a one-to-one fashion with particular constraints internal to any particular grammar. As Hale and Reiss stress with respect to FOT approaches to phonology:

> Phonology is not and should not be grounded in phonetics since the facts that phonetic grounding is meant to explain can be derived without reference to *phonology*. Duplication of the principles of acoustics and acquisition inside the grammar violates Occam's razor and thus must be avoided. (Hale and Reiss 2000: 162)

They go on to observe that 'Computational theories aren't inherently functionalist; people are functionalist' (p. 166). The computational system provided by UG specifies what a possible human language is and characterizes the formal properties of individual languages. An explanation for why some grammars are more common than others is not provided by UG. Nor should it be, given the existence of perfectly reasonable function-based explanations for the typological distribution of grammatical elements (for more remarks along these lines, see Newmeyer 1998a).

3.2.3. Functionality, universality, and grammatical hierarchies

3.2.3.1 Functional motivation vs. innateness

If one takes the mainstream OT position that the full set of constraints is innate, then one has 'solved' the learnability problem for constraints. There *is* no problem, though, of course, explaining the acquisition of language-particular rankings still presents more than a minor challenge. But what about FOT? It is by no means clear how the requirement that each grammatical constraint have an accompanying 'user constraint' can be made to be compatible with their putative universality. The central theme of the functionalist literature is that functional motivation is an *alternative* to innateness. Functionalists argue that by showing how grammars are rooted in language-external functional and cognitive human attributes, the idea of an innate UG can be dispensed with:

> Indeed, because the functionalist working hypothesis about innate linguistic abilities is that they are part of a more general cognitive-learning schema, the avenue of 'explaining' some phenomenon by declaring the

rule describing it to be innate is not available… (Foley and Van Valin 1984: 13)

It would be nothing less than bizarre to claim that all constraints are *both* innate and functionally motivated. As Haspelmath (1999: 184) has trenchantly observed, '[functional] justifications are irrelevant in a theory that assumes innate constraints'. But if some constraints are functionally motivated but *not* innate, then how could the full set of constraints possibly be universal? For example, consider a language L in which one never finds operators (i.e. *wh-* and other quantifier-like elements) in specifier position. Presumably in the grammar of L, OP-SPEC would have to be a very low ranked constraint. If OP-SPEC is simply innate, then the child learning L would, in the acquisition process, keep 'demoting' it to a lower and lower ranking. But if constraints are functionally-motivated, there would be no reason for the child acquiring L to acquire OP-SPEC at all! Hence, given FOT, constraints cannot be universal.

3.2.3.2 Some general remarks on functionally-motivated hierarchies

The grammatical hierarchies discussed by Aissen and incorporated into her version of OT are interesting in that they are plausibly both innate and functionally-motivated. Consider the statement that humans are higher ranked than nonhumans, agents are higher ranked than patients, and so on. It is not implausible that we are born with some sort of knowledge of 'cognitive importance', i.e., that animates are more central to human cognition than inanimates, that agents are more central than patients, and so on. It would seem that FOT might be on safe ground, then, in positing that these (externally motivated) hierarchies are part of grammar. If the person and other hierarchies are innate attributes of human cognition, then the dilemma discussed above with respect to parochial constraints such as OP-SPEC does not arise.

The remainder of this section, however, will argue that these hierarchies are *not* in the grammar. Let us begin by asking how would one go about testing for whether they are in the grammar, rather than just being an external functional influence on the overall shape of grammars. In principle, it is not really that difficult. We would simply need to find some grammatical process in some language that refers to several positions on the hierarchy. It is not clear however that such a process exists. Consider the Animacy Hierarchy, repeated below:

(48)   1st Person Pronoun > 2nd Person Pronoun > 3rd Person Pronoun > Proper Noun > Human Common Noun > Animate Common Noun > Inanimate Common Noun

To establish the Animacy Hierarchy as a grammatical construct, one would need to find for some language a grammatical generalization that involves a mapping from distinct positions on the hierarchy to some aspect of grammatical form. I know of no such case. To be sure, certain grammatical generalizations seem sensitive to *relative* degree of animacy. In Navajo, for example, higher-animacy arguments must precede lower-animacy ones (Witherspoon 1977). And other languages have chosen some cut-off point in the hierarchy to make some grammatical distinction. In Kharia, pronouns and nouns referring to animate beings have a number distinction not found among common nouns referring to inanimate things (Biligiri 1965). Such situations are easily handled by means of reference to a binary feature distinction (whose roots might well lie in some hierarchically-arranged cognitive categorization of animacy distinctions). But apparently, however, no language has grammaticalized the Animacy Hierarchy itself. We only see bits of its shadow in the actual grammars of particular languages.[15]

It is interesting to observe that none of the tableau in Aissen (1999) involve linking more than two positions on one hierarchy with more than two positions on another. I doubt that any attempt at a more extensive harmonic alignment would succeed. For example, indirect objects are generally considered to be low on the Relational Hierarchy. But this grammatical function is almost always realized by a human (or at least higher animate) NP. The harmonic alignment of the full Relational Hierarchy and the full Animacy Hierarchy would therefore make incorrect predictions.

Let us now examine some of the hierarchies, one-by-one.

3.2.3.3 The Thematic Hierarchy

There is reason for strong doubt that there exists a Thematic Hierarchy—even as a cognitive, as opposed to a strictly grammatical, construct. That seems to be the best explanation for the fact that after over three decades of investigation, nobody has proposed a hierarchy of theta-roles that comes close to working. Here is just a sampling of the versions of the Thematic Hierarchy that have been proposed over the years:

---

[15] William Croft (personal communication) suggests two possible counterexamples to such a claim. In K'iché, only pronouns and human nouns inflect for plural, and only animate NPs trigger 3pl absolutive verb agreement. Croft has described this situation in terms of two different cutoff points in two different constructions/rules. But he notes that one could argue that it actually involves reference to three different points on the Animacy Hierarchy: human Ns with noun pl and 3pl agr; animate Ns with no noun pl, but 3pl agr; inanimate Ns with no noun pl, no 3pl agr.

Croft's second example was discussed in Croft (1988). In the Australian language Kunparlang verb agreement is quantitatively correlated with animacy, so that humans trigger agreement at a higher proportion than animates, and inanimates at a still lower proportion. In other words, in this case it might be necessary to refer to more than one point on the hierarchy.

(49)   Versions of the Thematic Hierarchy:
a.   Fillmore (1968: 33)
      Agent > Instrumental > Objective

b.   Jackendoff (1972)
      Agent > Location/ Source/ Goal > Theme

c.   Ostler (1980)
      (relational predicates) Theme > Goal > Source > Path
      (actional predicates) Source > Path > Theme > Goal

d.   Givón (1984: 139)
      Agent > Dative/ Beneficiary > Patient > Locative > Instrument/
Associative > Manner

e.   Kiparsky (1985: 20)
      Agent > Source > Goal > Instrument > Theme > Locative

f.   Carrier-Duncan (1985: 7)
      Agent > Theme > Goal/ Source/ Location

g.   Larson (1988: 382-383)
      Agent > Theme > Goal > Location (and other obliques)

h.   Wilkins (1988: 211)
      Agent > Patient > Location/ Source/ Goal > Theme

i.   Randall (1988: 138) (for effects of lexical rules on argument structure)
      Theme > Agent > Instrument/ Source/ Path/ Goal/ Location/ …

j.   Bresnan, and Kanerva (1989: 23); Bresnan and Moshi (1990: 169); Alsina
(1996: 688)
      Agent > Beneficiary > Recipient/ Experiencer > Instrumental > Theme/
Patient > Location

k.   Baker (1989: 544)
      Agent > Instrument > Patient/ Theme > Goal/ Location

l.   Grimshaw (1990: 8)
      Agent > Experiencer > Goal/ Source/ Location > Theme

m.   Jackendoff (1990: 261)

Agent > Patient/Beneficiary > Theme > Source/ Goal/ Reference Object > Identificational Goal/ Reference Object

n.   Langacker (1990)
Agent > Instrument > Patient/ Mover/ Experiencer ('Energy Flow Hierarchy' for subject choice)
Agent > Experiencer > Other ('Initiative Hierarchy')

o.   Speas (1990: 16)
Agent > Experiencer > Theme > Goal/ Source/ Location > Manner/ Time

p.   Dowty (1991); Rugemalira (1994)
Proto-Agent > Proto-Patient

q.   Kiefer (1993)
Actor > Agent > Beneficiary > Theme/Patient > Instrument

r.   Van Valin and Lapolla (1997: 127) (continuum in terms of LS (=logical structure) argument positions) — this is an elaboration of Foley and Van Valin 1984:59:
Agent > Effector/ Mover/ User/ etc. > Location/ Perceiver/ Cognizer/ Experiencer/ etc. > Theme/ Stimulus/ Implement/ etc. > Patient/ Entity

3.2.3.4 Against proto-thematic roles

The theory of proto-thematic roles argued for in Dowty (1991) is often taken as the solution to the problem of too many thematic hierarchies, and, in fact, Aissen (1999) adopts it. Dowty argued that the proliferation of theta-roles seen in the literature can be dispensed with in favor of two cluster-concepts called 'Proto-Agent' and 'Proto-Patient'. The particular role borne by a particular argument is determined by the number of entailments that the verb gives it. Dowty outlined these entailments as follows:

Contributing properties for the Agent Proto-Role (Dowty 1991):
    a.   volitional involvement in the event or state
    b.   sentience (and/or perception)
    c.   causing an event or change of state in another participant
    d.   movement (relative to the position of another participant)
    e.   exists independently of the event named by the verb
Contributing properties for the Patient Proto-Role:
    a.   undergoes change of state
    b.   incremental theme
    c.   causally affected by another participant

d.　　stationary relative to movement of another participant
　　　e.　　does not exist independently of the event, or not at all

Dowty suggests that the argument whose predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object.

　　　Unfortunately for the proto-theta-role approach, it is not difficult to find transitive verbs whose subjects and objects do not measure up to Dowty's thematic criteria for them. Consider such stative predicates as *receive, inherit, undergo,* and *sustain.* To illustrate the problem, let us measure sentence (50) against Dowty's proto-agent entailments (51):

(50)　　John received a letter from Mary

(51)　　a.　VOLITION: *Mary*
　　　　　b.　SENTIENCE/PERCEPTION: d.n.a.
　　　　　c.　CAUSATION: *Mary*
　　　　　d.　MOVEMENT: *the letter*
　　　　　e.　INDEPENDENT EXISTENCE: *John, Mary*

John is a 'Proto-Agent' only by one-half of one test, while *Mary* passes two and one half. Hence, one would predict that *Mary,* not *John,* should be subject.

　　　One might, on the other hand, choose to interpret Dowty's criteria as governing the *necessary* properties of the roles associated with each predicate, rather than focusing on individual sentences (such as (50)) in which that predicate occurs. By this interpretation, the verb *receive* fares no better, as Table V illustrates.

Proto-Agt properties	[recipient]	[theme]	[source]
Volition	no	no	no
Sentience	no	no	no
Cause	no	no	no
Movement	no	yes	no
Independent existence	yes	yes	yes
Total	**1**	**2**	**1**
Proto-Pat properties	**[recipient]**	**[theme]**	**[source]**
Change of state	no	no	no
Incremental theme	no	no	no
Causally affected	yes?	yes	no
Relatively stationary	yes	no	no
No independ. existence	no	no	no
Total	**2**	**1**	**0**

Table V
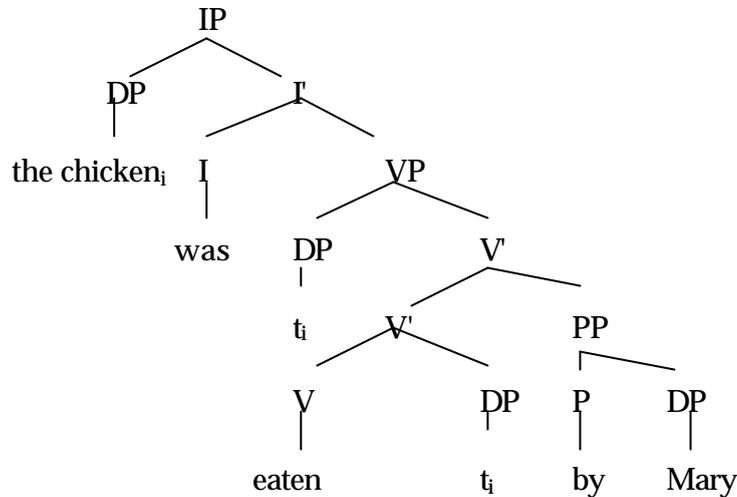Proto-Agent and Proto-Patient properties of the verb *receive*

If the assignments in Table V are correct, then for the verb *receive*, the theme should be the subject and the recipient should be the direct object, i.e., we should have sentences like *A package received John from Mary.* As far as subject properties are concerned, the theme necessarily moves and has an independent existence, while the only necessary property of the recipient is its independent existence. (Note that sentences like *The wall received a coat of paint* illustrate that the recipient need not be sentient.) As far as object properties are concerned, recipients are, I believe, always causally affected by another participant and stationary relative to the movement of another participant.

It is worth pointing out that passive sentences (and most likely A-movements in general) pose a lethal problem for the attempt to link any approach that incorporates the UTAH (Baker 1988) and the hypothesis of proto-thematic-roles.[16] Consider the active-passive pair (52a-b). For the past two decades, or longer, it has been assumed that (52b) has a derivation schematically represented as (53) (though the intermediate trace in [Spec, VP] is a more recent innovation):

---

[16] According to the UTAH (Uniformity of Theta Assignment Hypothesis), identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure. Virtually all principles-and-parameters approaches adopt the UTAH, I believe. See Baker (1997; to appear) for an attempt to effect a linkage between the UTAH and Dowtyan proto-roles and for criticism of this attempt, see Newmeyer (to appear).

(52)   a. Mary ate the chicken.
            b. The chicken was eaten by Mary.

(53)

```
                        IP
                    ┌────┴──────┐
                   DP           I'
                    |        ┌───┴────────┐
            the chicken_i    I            VP
                             |        ┌───┴──────┐
                            was      DP          V'
                                     |       ┌───┴──────────┐
                                    t_i      V'             PP
                                          ┌──┴────┐      ┌──┴───┐
                                          V       DP     P      DP
                                          |       |      |      |
                                        eaten    t_i    by     Mary
```

Such a derivation, however, violates the UTAH. *Mary* has identical theta–roles in (52a) and (52b) and, therefore, by the UTAH, would have to be in identical D-structure positions in those two sentences. But agent phrases of passives have not been analyzed as originating in deep subject ([Spec, IP]) position since before Emonds (1970) — and the Theta-Criterion, of course, explains *why* they cannot originate there. So we have a fundamental conflict between the UTAH and the Theta-Criterion. Worse, however, is the fact that (53) leads to the assignment of the wrong theta–roles, given a Dowtyan analysis. The argument with the most proto-Agent properties is not a deep subject; indeed the deep subject, being empty, has no thematic properties at all. That presents no problems for a semanticist like Dowty, who rejects null elements in syntactic derivations, but is highly problematic for the standard principles-and-parameters approaches. Even worse, the agent phrase can be omitted:

(54)   The chicken was eaten.

As Dowty makes clear in his paper, in sentences with only one DP argument, that argument is assigned the subject position by default. But in (52b), *the chicken* needs to be an underlying direct object.

Even if something like proto-thematic roles could be motivated for subject and object choice, more fine-grained roles would still be necessary for other facets of the grammar. This fact undercuts any theoretical parsimony argument for reducing the number of roles to two. For example, a tradition in word-formation studies going back at least as far as Carrier-Duncan (1985) states morphological generalizations in terms of a variety of theta-role labels. Or consider the analyses of control phenomena in Sag and Pollard (1991) and Culicover and Jackendoff (1999). In the latter analysis, the controller is identified

differently for different verb classes, verbs in each class picking a particular theta-role as controller. The controller might be an Agent (55a), an Addressee (55b), a Patient (55c), a Recipient (55d), a Source (55e), or a Holder (55f) (the controller is identified in bold-face):

(55)   a. **Mary** attempted to take a vacation.
        b. Bill told **Mary** to come visit him.
        c. I persuaded **John** to see the doctor.
        d. We taught **the dog** to roll over.
        e. **Sam** promised Mary to write more often.
        f. Tom required **Alice** to hand in the assignment.

No way suggests itself for handling the control facts by reference only to the roles Proto-Agent and Proto-Patient.[17]

    In short, while grammatical processes might well have to refer to individual thematic roles, there is little evidence for a Thematic Hierarchy forming an integral part of UG.

3.2.3.5 The Relational Hierarchy

    We now have the Relational Hierarchy to consider. It is repeated below as (56):

(56)   **Relational Hierarchy:** Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparative

Are grammatical relations innate constructs? Frameworks that accord them the status of grammatical primitives, such as relational grammar and lexical-functional grammar (LFG) have tended to answer this question in the affirmative. For example, Steven Pinker, writing from an LFG perspective, has remarked:

> The child is assumed to know, prior to acquiring a language, the overall structure of the grammar, the formal nature of the different sorts of rules it contains, and the primitives from which those rules might be composed. (Pinker 1984: 31)

Pinker goes on to suggest that the child uses semantic bootstrapping to identify agentive arguments of active action predicates as subjects and then expects

---

[17] See Primus (1999) for arguments that even for the linking of grammatical relations and thematic roles, more proto-roles are needed than Proto-Agent and Proto-Patient. Primus motivates the need for the role 'Proto-Recipient'.

'without further learning' (p. 44) that entities labeled SUBJ will be the leftmost NP daughter of S, can be controlled by an argument of a matrix predicate, and so on.

Such a view seems consistent with standard versions of OT. If constraints involving reference to grammatical relations are innate, then it follows that grammatical relations are innate. But, for the reasons discussed at the beginning of this section, this view seems highly problematic from the point of view of FOT. Somehow, in order to preserve the essence of both standard OT and FOT, an advocate of the latter would have to find some way to defend the claim that grammatical relations (and their hierarchy) are universal without being innate.

It is not clear how such a claim can be defended. In the standard functionalist view of grammatical relations (e.g. Croft 1991), they are not primitive terms at all, but rather 'cluster concepts', prototypically reflecting such cognitive notions as prominence, topicality, animacy, agency, and so on. Such a view, whatever its merits, entails a trivialization of the FOT appeal to hierarchy alignment to capture grammatical generalizations. The harmonic alignment of the relational hierarchy with those of animacy, prominence, and so on can be of little theoretical interest if the constructs composing the former are defined in terms of the constructs composing the latter. To be sure, there is no obligation for advocates of FOT to adopt the functionalist approach to grammatical relations, but if they do not they would seem to have no alternative but to land on innateness as a most uncomfortable fallback position.

In any event, there is reason to doubt that grammatical relations are universal concepts.[18] There is an interesting discussion in Van Valin and Lapolla (1997) of the evidence that might lead a learner to posit that his or her language has distinct grammatical relations such as 'subject', 'direct object', and so on. The best evidence, they argue, is what they call the 'neutralization of semantic oppositions for morphosyntactic purposes'. So take English, which, like most languages, has exactly this type of neutralization. Consider a typical active – passive pair:

(57) a. John kills the ducklings.
      b. The ducklings are killed by John.

---

[18] In mainstream transformational approaches, grammatical relations have generally been defined derivatively in terms of configurational structure, so the question of their innateness/universality has been discussed only in th e context of the innateness/universality of particular structural configurations. A subject, for example, was originally defined simply as an NP immediately dominated by S. The situation is more complicated now, given the VP-internal subject hypothesis. In principles -and-parameters accounts, there are several 'subject positions', each correlated with different subject properties, and not all which need be filled in every language (see McCloskey 1996 for discussion).

It is the agreement pattern that is of interest to us. The verb agrees in (57a) with the agent and in (57b) with the patient. In other words, agreement is not purely semantically-determined; rather it needs to make reference to some morphosyntactic category, in this case the 'subject'. Languages differ as to which processes exhibit neutralization of semantic oppositions for morphosyntactic purposes and therefore provide evidence for grammatical relations. In those languages in which agreement is subject to purely semantic conditions, there is no argument derivable from this process for the existence of distinct grammatical relations in that language.

Now, as it turns out, there exist languages in which there is *never* restricted neutralization of semantic roles. Based on the work of Mark Durie (1985; 1987), Van Valin and Lapolla identify Acehnese as one such language and Primus (1999) makes the same point with respect to Guarani and Tlingit. No learner of these languages would be led to posit a distinction between subjects and objects. Therefore, grammatical relations are not universal, nor can any hierarchy involving them be universal. FOT is forced either to give up the universality of constraints or, more seriously, to explain how a speakers can have in their grammars functionally-motivated constraints for which they have no evidence.

3.2.3.6 On Aissen on subject choice

This sub-section and the following will examine the hierarchy alignment approach of Aissen (1999; 2000) and argue that it contains serious deficiencies. It must be stressed that Aissen's work is being singled out for critical treatment not because it is a particularly poor exemplar of its genre, but for precisely the opposite reason. Aissen's approach represents the most subtle and compelling argumentation within the general framework of FOT. If such work cannot motivate the need for explanations based on interleaved hierarchies, then it must be concluded that none can do so.

Let us begin with her ranking for English subject choice in the 1999 paper:

(58)   *Su/x » *Su/Pat » *GR/Pers

In this account, the 'worst' subjects are posited to be non-prominent ones. But contra Aissen, subjects in English *can* be non-prominent. Given the definition of prominence in terms of attention (see §3.1), *the candy* and *your garage* are prominent in examples (59-60) respectively:

(59)   Q: What happened to the candy?
        A: Bill ate it.

(60)   Q: What happened to your garage?

A: Well, a gasoline can exploded there and blew it to kingdom come.

The replies violate her proposed constraint ordering for English and hence should be impossible. But they are fine.

Now let us turn to Aissen's ranking for Fox:

(61)   *Su/Pat » … » {*GR/Pers, *GR/χ}

Here we have a problem of a different sort. The dominant ranking of *Su/Pat is designed to capture the fact that in Fox all clauses are active. But the generalization for Fox is broader than that. That language presents no evidence for A-movements / relation-changing rules at all.[19] A constraint simply saying that subjects cannot be patients is much too narrow. There is no doubt that the correct generalization for Fox can be stated in optimality-theoretic terms, but there is little evidence for the existence of interleaved hierarchies emerging from the 'optimal' analysis.

Based on the discussion in Jelinek and Demers (1983), Aissen proposes the following constraint ranking for Lushootseed:

(62)   *Obl/Local » *Su/x » *Su/Pat » *GR/Pers

The high ranking of *Obl/Local captures Jelinek and Demers' claim that passive agents cannot be in the first or second person. Dawn Bates (personal communication) has suggested that this prohibition, at least where the clitic (or affixal) forms are concerned, might be due to prosodic factors. If so, then a constraint like *Obl/Local is unneeded, since it would follow from something broader. And there seems to be uncertainty among Salishists as to whether the prohibition even holds at all when 'independent pronouns', i. e., those playing an emphatic focusing role, are involved.[20]

Finally, Aissen proposes the following ranking for Lummi:

(63)   *Obl/Local » *Oj/Local » *Su/x » *Su/Pat » *GR/Pers

This ranking captures the impossibility of 1st and 2nd person patients with 3rd person agents in the active voice and their possibility in the passive. Here it is not clear if Aissen's constraint rankings for Lummi capture the full set of relevant generalizations. According to Richard Demers and Eloise Jelinek (personal communication), the prohibition against local patients with 3rd person agents in the active voice is a function of the split-ergative system of that language in main

---

[19] Thanks to Ives Goddard (personal communication) for this information.
[20] I would like to thank Dawn Bates, Henry Davis, and Thom Hess for discussing the Lushootseed data with me.

clauses. 1st and 2nd person are NOM/ACC, while 3rd person is ERG/ABS. It follows that the internal argument can be a 3rd person ERG or a 1st & 2nd person ACC. So the 3rd ERG and the 1st/2nd ACC compete for the same syntactic position, leading, as it turns out, to *neither* being acceptable. Interestingly, there is no split ergativity in *embedded* clauses and there one does, in fact, find 1st and 2nd person patients of active verbs. It is unclear if and how these facts might emerge from Aissen's analysis.[21]

3.2.3.7 On Aissen on differential object marking

The principal objection that can be raised against the treatment of differential object marking in Aissen (2000) is not based on its empirical inadequacy, but rather on its conceptual complexity. A vast battery of high-tech machinery is necessary, given the FOT framework in which it is situated, when only a hand tool or two suffices to get the job done.

Despite Aissen's analysis, the following is *all* that the grammars of Hebrew and Turkish need to say about object case marking:

(64)   a. Hebrew: Case-mark definite objects
        b. Turkish: Case-mark specific objects

That is the sum total of grammatical information pertinent to object case marking in those languages. But what about the fact that Hebrew and Turkish fall on different points on the functionally-motivated hierarchy of definiteness? That is an interesting fact, but it is no more a fact to be encoded in grammars than is the fact that any language that allows sentences with three levels of center-embedding will also allow sentences with two levels. Grammars have been shaped, in part, to reflect functional motivation. As Aissen points out, citing the rich functionalist literature on the topic, there is functional utility in signaling the rarity of a high definiteness object by means of a special marking. Hence, as Hebrew and Turkish were shaped over time, it would have been odd for speakers to begin marking objects of one particular degree of definiteness without at the same time marking those rarer objects of a higher degree. The result is the appearance that speakers of these languages are 'following' a hierarchy that they have 'internalized'. They are doing no such thing. All that a

---

[21] Judith Aissen (personal communication) helpfully corrected an earlier critique of mine of her analysis of the Lummi data. Both the facts and the analysis are so intricate that I fear that misunderstandings might remain in my exposition. It is worth pointing out that Aissen's analyses depart from those of others in OT (for example, Ackema and Neeleman 1998a) in positing that actives and passives compete in the same candidate set. Such a treatment, while crucial to achieving the desired results, departs from the standard OT assumption that competitors share argument structure (assuming that actives and passives differ in argument structure).

speaker of Hebrew knows is that definite objects are case marked and all that a speaker of Turkish knows is that specific objects are case marked.

The amount of machinery required by a wholly grammar-internal OT account of these facts is staggering. One needs one constraint penalizing the presence of case marking and another constraint penalizing its absence. One needs to harmonically align hierarchies and then to locally conjoin the product of this alignment with the constraint penalizing the absence of case marking. And most cumbersome of all, because in OT without competition between rival candidates there is no way for the correct form to emerge, one needs to set up a proliferation of otherwise unneeded candidate sets, simply to ensure that a 'winner' results. It is never clear from a reading of the OT literature how literally (in terms of claims about I-Language) one is to take the tableaux that form the centerpiece of every OT analysis. If tableaux like XI and XII are simply metaphorical in nature, then one would have to object that such metaphors do little to promote understanding. But a literal interpretation seems remote, to say the least. Are we really to believe that for each language, for each degree of definiteness, case marked and non-case marked objects are in a separate competition in speakers' heads with each other? If OT is incapable of formulating the following generalization in so many words — that in language L objects with property x are case marked and objects with property y are not case marked — then one would have to conclude that that framework is less than appealing as an approach to UG.

To conclude, there seems to be little evidence for a central hypothesis of FOT, namely the idea that functionally motivated hierarchies are central to grammatical description.

4.   The problem of unmotivated competition

There are two ways that sentences with the same lexical items, argument structure, and tense and aspect morphemes might be considered to be 'optional variants' of each other. Where no obvious differences in information structure distinguish the two sentences, we have a case of 'discourse-independent optionality'. Where there are information structure differences, we have 'discourse-dependent optionality'. (65a-b) illustrate the former, (66a-b) the latter:

(65)   a. I believe that it will rain tomorrow.
        b. I believe it will rain tomorrow.
(66)   a. You will never get me to eat fruit-flavored tofu.
        b. Fruit-flavored tofu you will never get me to eat.

This section will argue that the core notion of OT, the idea that particular well-formed structures emerge as 'winners' as a result of 'competition' with their less successful rival structures, leads inevitably to uninsightful analyses of both types

of optionality. Section 4.1 gives a brief historical overview of the handling of optionality within grammatical theory and reviews some of the less successful approaches to optionality within OT. Sections 4.2 and 4.3 discuss what appear to be the standard approaches to discourse-independent and discourse-dependent optionality within OT, respectively. In both cases it is argued that OT needs to resort to the setting up of ad hoc competition sets. Worse, to handle discourse-dependent optionality, discourse constructs need to be imported into the syntax, leading to analyses which are both empirically deficient and violate the principle of syntactic autonomy.

4.1 The treatment of optionality in grammatical theory: An overview

In early transformational grammar, pairs of sentences exhibiting discourse-independent and discourse-dependent optionality were related by means of optional transformational rules. Indeed, in Chomsky (1957) even sentences with different cognitive meanings were so related (declarative-interrogative pairs, affirmative-negative pairs, and so on). The Katz-Postal Hypothesis (Katz and Postal 1964), with its requirement that all information necessary for semantic interpretation be present in underlying syntactic structure, reduced to some degree the amount of optionality in the transformational component. Negative elements, for example, could no longer be inserted, as they were relevant to semantic interpretation. Likewise, question and imperative transformations became obligatory, subject to the presence of abstract markers in underlying structure. The move away from optional transformations was taken further by the generative semantics tendency, which dominated syntactic theorizing in the late 1960s and early 1970s (for discussion, see Newmeyer 1986). Since discourse-relevant constructs like 'focus' and 'presupposition' were deemed to be present in underling syntactic structure, discourse variants could no longer be related by means of optional transformational rules.

With the publication of Chomsky (1973), however, the pendulum began to swing back to an increased amount of optionality in the transformational component. The approach initiated by this paper led by the late 1970s to one single movement rule, Move-α, supplemented by constraints acting as filters on overgeneration. Optionality, then, could result if two or more S-structures derived from the same D-structure survived the filtering process. Discourse-dependent optionality was not handled in a uniform fashion in the government-binding theory. Sentences like (65b) with topicalized NPs were posited to have different D-structures from simple declaratives like (66a). But in other cases, two discourse variants were derived from different applications of Move-α to the same D-structure, and distinguished in their information content by interpretive principles applying at LF.

The minimalist program has seen a shift away from optionality. In the MP all movement is in effect obligatory, since it takes place only where it is mandated by some feature. So apparent optionality has to result from two derivations arising from different candidate sets, as in the treatment of 'optional' *Wh*-Movement in Denham (2000). That is, different elements need to be chosen from the numeration for each derivation. The problem with such an approach, as pointed out in Sternefeld (1997), is that reducing competition on candidate sets has an unwanted consequence — many ill-formed derivations that would have otherwise been blocked will now survive because the more economical derivation is not part of the same candidate set anymore (for discussion of this problem, see Müller and Sternefeld 1996).

Optionality — or apparent optionality — presents the same sort of dilemma for OT that it does for the MP. Since the leading idea of OT is competition, winners, and losers, the possibility of optionality is threatening to the enterprise. How might OT handle this problem? The remainder of this subsection will outline various possibilities, saving discussion of the most popular attempted solution for the following two subsections.

One way to defuse any potential optionality problem is to claim that the seemingly optional variants belong to separate grammars. For some time, Anthony Kroch has been advocating the idea that many individuals have multiple grammars, each grammar corresponding to a different stylistic register (Kroch 1989; Kroch and Taylor 1997). Positing multiple grammars seems to solve some long-standing problems in history of English syntax. Such a move has been applied to OT syntax in Legendre (2000a) in her analysis of French *wh*-constructions. Both (67a) and (67b) below mean 'Where did Peter go?', but the first is far more informal than the second:

(67) a. Pierre est parti oú?
      b. Oú Pierre est-il parti?

Legendre posits different grammars with different rankings of STAY and OPSPEC.

Undoubtedly, in some cases it is correct to posit multiple grammars for a single individual, each grammar corresponding to a different register. But such an approach raised to the level of general solution to the problem of register variation has unpleasant implications for much of the work in generative syntax that has taken place over the past 45 years. Countless conclusions about the nature of grammars have depended on register variants being handled in the *same* grammatical system. Sentences with *do so* and with sentential subjects, for example, have typically been treated in the same data set as their more casual variants and sentences with *wanna*-contraction and stranded prepositions have been considered to be part of the same grammars as their more formal counterparts. The crucial data for Kaisse (1985), a work devoted to the understanding of rapid and informal speech, are sentences found only in the least formal registers. However, these

48

sentences are not distinguished in terms of the grammar they belong to from more stylistically elevated ones.
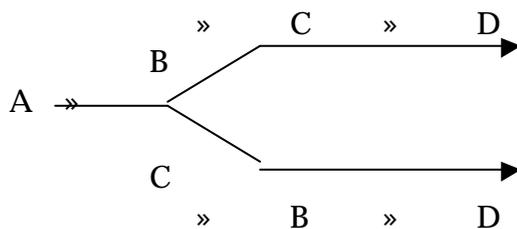
Most attempts therefore to deal with (apparent) optionality within OT have assumed one grammar, but with special assumptions about how constraints can interact to yield the *appearance* of optionality. The following few pages will outline these attempts, following closely the lucid extended discussion in Müller (1999b) and the clear overview in Legendre (2000a: ms, 12-13). Here are the various possibilities:

A.   True optionality

True optionality is a function of two candidates satisfying and violating the same constraints. Examples are the treatment in Grimshaw (1997) of the optionality of complementizer *that* in English and some related cross-linguistic correlations in Vikner (2000). As noted in Legendre et al. (1995) and Keer and Bakovic (1997), given the architecture of OT, true optionality will almost never obtain, since it is so easily sabotaged by the existence of some low-ranked constraint that distinguishes the two variants. And in fact, Grimshaw (1999) abandons her earlier analysis.

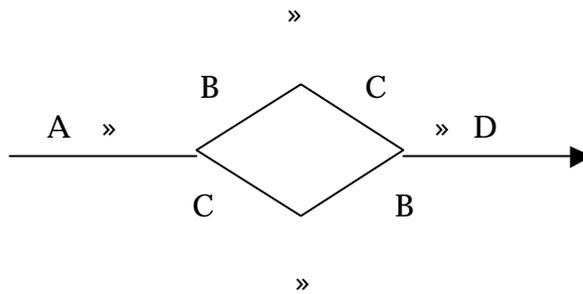B.   Ordered global constraint ties

One can think of this option as one of partial ordering of constraints (or 'floating constraints'). Anttila (1997) and Nagy and Reynolds (1997) take this option for phonology, as do Sells, Rickford, and Wasow (1996) and Ackema and Neeleman (1998b) for syntax. In a departure from strict ranking, one constraint can both precede or follow another, leading to more than one output from the same input (and hence optionality). Schematically:

```
                    »    C    »    D
              B  ────────────────────►
A  -»──────────
              C  ────────────────────►
                    »    B    »    D
```

 Legendre (2000a: ms, 12) says that ordered global constraint ties are problematic from a learnability standpoint (see Tesar and Smolensky 1998) and predict unattested typologies. Müller (1999b) notes that they vastly increase the number of possible grammars.
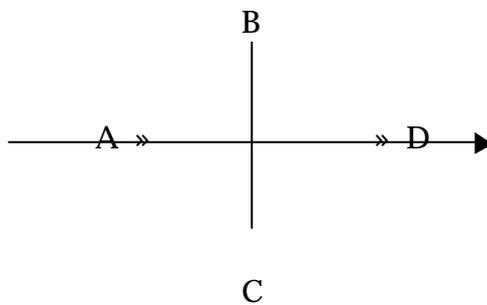
C.   Ordered local constraint ties

Ordered local constraint ties are a more constrained version of ordered global constraint ties. A » B and B »A, as with the latter, but the tie is not an abbreviation for two constraint orders. Rather, one can think of the tie itself as a constraint, so the violation of one cancels out the violation of the other. This approach to optimality has been mooted in Pesetsky (1998) and Legendre (2000b) and can be represented schematically as follows:

```
                              »

            B              C
   A   »                       »   D
            C              B

                              »
```

Müller (1999b) argues that ordered local constraint ties will usually fail to deliver the desired results, because all it takes is some lower ranked constraint to undermine the existence of two optimal candidates from one candidate set. Also, in the view of Keer and Bakovic (1997: 268), there is too much descriptive looseness in such an approach and too little explanatory pay-off.

D.   Conjunctive local constraint ties

Conjunctive local constraint ties embody crucial non-ranking of constraints, an approach that has been explored in Legendre et al. (1995) and Müller (1997). Ties are treated as ordinary constraints, but with no resolution of the tie into suborders:

```
                      B
                      |
         A »          |          » D
                      |
                      C
```
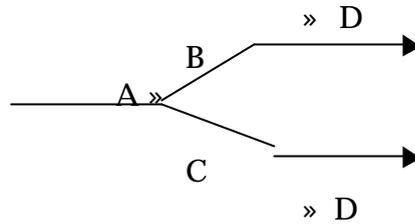
As noted by Müller (1999b), conjunctive local constraint ties have *both* the lower-ranked constraint problem and the learnability problem.

E.   Disjunctive ties:

```
                      B
         A »                 » D
E'. Local:
                      C
```

E". Global ——— A » B » D

C » D
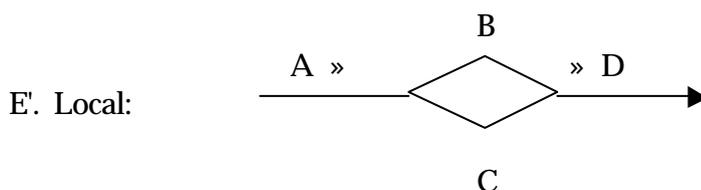
Again, following Müller (1999b), the former has the problem of a possible lower-ranked constraint sabotaging the optionality; the latter the problem of learnability.

4.2   The proliferation of candidate sets and discourse-independent optionality

The most appealed to solution to the (apparent) optionality problem within OT is to give the *illusion* of competition by an otherwise unmotivated creation of a distinct candidate set for each seemingly optional candidate. We will call any such analysis an instance of 'unmotivated competition'.

Let us look first at discourse-independent optionality, where no difference in information content distinguishes the variants. The examples below are taken from Bakovic and Keer (2000):

optional complementizer deletion:
(68)   a. I think [that the coat doesn't fit him].
          b. I think [the coat doesn't fit him].
(69)   a. The coat [that he always wears *t*] doesn't fit him.
          b. The coat [he always wears *t*] doesn't fit him.
(70)   a. The guy [who she dates *t*] doesn't respect her.
          b. The guy [she dates *t*] doesn't respect her.

To handle such examples, Bakovic and Keer make two otherwise unmotivated moves, though probably necessary ones, given the structure of OT. First, they expand the input to go beyond semantically discernable (lexical) contrasts to include semantically inert (functional) contrasts. So the input now includes functional features like [±COMP] and [±WH]. Second, they posit a set of highly-ranked faithfulness constraints sensitive to these new input features. So both (68a) and (68b) can now be generated as follows. (68a) will be the winner in the candidate set where faithfulness involves a high ranking of the feature [+COMP] (Tableau XIII); (68b) will be generated in the candidate set where faithfulness involves a high ranking of the feature [-COMP] (Tableau XIV):

Input: [+COMP]	FAITH[COMP]	MARKEDNESS CONSTRAINTS
☞ I think that the coat does not fit him		(*)
I think the coat does not fit him	*!	(*)

Input: [-COMP]	FAITH[COMP]	MARKEDNESS CONSTRAINTS
I think that the coat does not fit him	*!	(*)
☞ I think the coat does not fit him		(*)

Tableaux XIII and XIV
FAITH[COMP] and discourse-independent optionality (Bakovic and Keer 2000)

The fundamental problem with such an approach can be appreciated by means of an analogy. Suppose that Smith and Jones are about to compete in the 100-yard dash and we have psychic knowledge that they will finish in a dead heat. We don't like that — we like clear winners. So here is what we do. We have Smith wear blue shorts and Jones wear red shorts. Now we run two races involving Smith and Jones. First prize in one goes to the fastest runner wearing blue shorts and first prize in the other goes to the fastest runner wearing red shorts. Now Smith and Jones each emerge winners in one of the two races and losers in the other!

Clearly, the 'optimal' thing to do is to let them run in the same race and finish in a dead heat. Analogously, the 'optimal' solution to optional complementizer deletion is to let (68a) and (68b) compete and both emerge winners. But we saw in the previous section how difficult it is to effect such a solution.

As noted in Müller (1999b: 7), solutions to the optionality problem within OT like that of Bakovic and Keer's increase complexity in several domains. First, more candidate sets have to be considered for the same surface form. Second, the candidate sets are larger. And third, this type of approach creates vacuous ambiguities in well-formed sentences — there ends up being an optimal candidate that is faithful to the input and there is another optimal candidate in a (minimally) different candidate set that is not. These problems seem to strike at the heart of the OT enterprise.

4.3   The proliferation of candidate sets and discourse-dependent optionality

Let us turn now to discourse-sensitive optionality, that is, cases where we have sentences with identical lexical items, identical argument structure, and identical tense and aspect morphemes, but which differ in their information structure. The typical approach within OT has been parallel that for handling

discourse-independent optionality in the sense that the result is an increase in the number of candidate sets. Typically, some discourse feature is built into the input, generally as a feature on a lexical item, and constraints are formulated that are sensitive to the presence, absence, or position of that feature. Examples of such discourse features from the OT literature are 'focus' (Costa 1998, Grimshaw and Samek-Lodovici 1998); 'topic' (Grimshaw and Samek-Lodovici 1998); 'old information' and 'new information' (Hall 1998); 'active' and 'identifiable' (Legendre 1999); and 'discourse prominent' (Aissen 1999).[22]

This approach can be illustrated by reference to Legendre's analysis of French Stylistic Inversion (Legendre 1999). In French there are two ways of saying 'What message did John send to Mary?', one without inversion of the subject (71a) and one with inversion (71b):

(71)   a. Quel message est-ce que Jean a envoyé à Marie?
       b. Quel message a envoyé Jean à Marie?

Sentences (71a) and (71b) are not interchangeable in discourse. According to Legendre, the non-inversion sentence is more 'neutral', in that the subject represents active old information. ('Active' here means very recently mentioned — see Chafe 1987). In the inversion sentence, the subject is 'identifiable', but 'inactive', where 'identifiable' means 'stored in hearer's long-term memory' (Lambrecht 1994). Legendre builds the discourse status of the subject into the input and formulates the following constraints, the first of which refers to that discourse status:

(72)   ALIGN-INACTIVE: Align [-active] NPs with the right edge of V^0

(73)   NOM: Subjects must be realized in SpecIP

The following tableaux illustrate how Legendre captures the desired relationship between discourse status and word order:

Input: V = transitive; Jean = extarg. [-new]; wh = inner intarg

'What message did John send to Mary?'	ALINACT	NOM
[$_{IP}$quel message$_j$ a [$_{VP}$t$_i$[V'[V'envoyé t$_j$/Jean$_i$] à M]]]		*
☞ [$_{CP}$quel message$_j$ est-ce que [$_{IP}$Jean$_i$ a [$_{VP}$t$_i$ [$_{V'}$[$_{V'}$envoyé t$_j$] à M]]]]		

Input: V = transitive; *Jean* = extarg. [-active]; wh = inner intarg

'What message did John send to Mary?'	ALINACT	NOM
☞ [$_{IP}$quel message$_j$ a [$_{VP}$t$_i$[$_{V'}$[$_{V'}$envoyé t$_j$/Jean$_i$] à M]]]		*
[$_{CP}$quel message$_j$ est-ce que [$_{IP}$Jean$_i$ a [$_{VP}$t$_i$ [$_{V'}$[$_{V'}$envoyé t$_j$] à M]]]]	*!	

Tableaux XV and XVI
Capturing discourse-dependent optionality by syntactic constraints referring to discourse features (Legendre 1999)

The goal here is not to question the mechanics of Legendre's analysis. As far as one can determine, it achieves the desired result. However, two general points are in order. The first is that the need to build discourse constructs into the syntax in such a manner follows from the general structure of OT. The second is that, in general, serious problems arise from not distinguishing purely syntactic generalizations from generalizations relating syntactic structure and information structure.

The necessity for OT syntactic analyses to incorporate features indicating discourse status follows from the difficulty of engineering the constraints and their rankings so that discourse-dependent optional variants like *Quel message a envoyé Jean à Marie* and *Quel message est-ce que Jean a envoyé à Marie* might emerge as winners in the same competition set. Hence, it is necessary for them to 'compete separately', which entails distinguishing them by virtue of their information content.

The blurring of the boundary between purely syntactic generalizations and those governing the information status of syntactic forms is also encouraged by the difficulty within OT of handling interactions between grammatical components. The problem, as observed in McMahon (2000: ms., p. 32), is most serious when principles from two components are interleaved, as in lexical phonology. OT works well when it is a matter of ranking constraints in a candidate set where only one grammatical component is involved. But it is less than obvious how OT should go about handling mappings from one component to another. One possibility might be that languages rank constraints on input-output pairings, so that, for example, *Subject/non-topic would be a high-ranked constraint in languages where subjects have to be topics. However, in general the interfaces between components are too complex to allow solutions of that sort to work, as will be argued in the remainder of this section with respect to the syntax-information structure interface.

There are serious general theoretical and methodological problems with any analysis that builds discourse constructs like 'active' into the syntactic analysis. The first point to make is that such analyses run afoul of the principle of the autonomy of syntax, namely the idea that the principles characterizing well-formedness of sentences structurally are distinct from those characterizing their appropriate use in discourse. Autonomy, of course, is an empirical issue — it

could be incorrect. But as has been argued at length (most recently in Newmeyer 1998b) it is *not* incorrect. And as we will see, the indirectness of the relationship between syntactic constituency and the information status of participants supports autonomy.

Let us examine the evidence that information structure, in general, cannot be marked as a feature on a single constituent.[23] First, note that focus (which is capitalized in the following examples) can be discontinuous:[24]

(74)   Q: Did Mary wash the car?
        A: No, TOM washed the WINDOWS.

The topic can be discontinuous as well:

(75)   Q: Are Mary and Tom friends?
        A: No, Mary HATES Tom.

The topic can be a non-constituent:

(76)   Q: WHO did the man see?
        A: The man saw THE WOMAN

The topic need not have a syntactic representative:

(77)   Q: What's the weather?
        A: It's RAINING.

Chomsky (1971: 205) even provided an example of where the focused element is a prefix:

(78)   John is more concerned with AFfirmation than with CONfirmation.

Examples (74-78) lead to the conclusion that neither 'topic' nor 'focus' are properties that can be marked as simple features on syntactic constituents. Rather, they call for complex principles of interpretation relating syntactic structure and information structure.

---

[23] Many syntacticians have proposed syntactic constituents with names like 'Topic Phrase', 'Focus Phrase', and so on (see especially Rizzi 1997). To the extent that there is good syntactic motivation for such constituents, then no objection can be raised. But it is important to stress that because something is *called* a (syntactic) 'Topic Phrase', it does not necessarily follow that the material occupying that phrase is the information structure topic. More often than not, it will not be.
[24] Thanks to Jeanette Gundel for discussion and for providing me with several of the examples to follow.

Interestingly, the issue of the relationship between syntactic structure and information content was at the heart of one of the earliest debates between Chomsky and George Lakoff. Chomsky (1971), taking a position that was similar to that of Halliday (1967: 8), argued that focus and presupposition (topic, essentially) were read off of surface structure constituency. Chomsky wrote: 'The focus is the phrase containing the intonation center, and the presupposition is determined by replacement of the focus by a variable…' (p. 200). That surface structure appeared to be implicated in focus and presupposition was suggested by the fact that transformationally derived structures are focusable:

(79)   a. John is certain to LOSE.
        b. John is likely not even to be NOMINATED.

In reply, Lakoff (1971: 262) argued that information structure differences do not necessarily correspond in any simple manner to structural differences. There is no evident disparity in information content, for example, between sentences which differ with respect to whether or not Particle Movement has applied:

(80)   a. John looked up a girl who he had once met in Chicago.
        b. John looked a girl up who he had once met in Chicago.

Lakoff showed that the relation between surface structure and notions like focus and presupposition is very indirect. For example, he called attention to the following sentence:

(81)   The TALL girl left

As Lakoff noted, *tall* should be the information structure focus in the Halliday-Chomsky account. That is, it should be new information. But it is presupposed, analogously to the relative clause *who was tall*. He wrote:

> In [this example] it is presupposed that some girl left and it is
> presupposed that some girl is tall. The new information is that the girl
> who was presupposed to have left is coreferential with the girl who was
> presupposed to be tall. (Lakoff 1971: 261)

An example of an even more complex and indirect relationship between syntactic and information structure is provided in Bach (1974: 271-272) (who cites Asa Kasher). A restrictive relative clause presupposes the existence of entities of which the description given in the relative clause is not true. So the following sentence presupposes the existence of at least one man that I did not see:

(82)   The man that I saw took the number 75 bus.

To complicate the interaction between syntax and discourse still further, formal marking need not correspond directly to information status. For example, a formally definite NP need not represent information known to the hearer (Hawkins 1978 and Gundel 1988: 214):

(83)   I couldn't sleep last night because the neighbor's dog kept me awake.

Evidence for the lack of a simple mapping between syntax and information structure is further evidenced by the many-many relationship between the two. For example, Prince (1998) argues that the process of left-dislocation has three functions: to simplify discourse processing, to trigger a poset (partially-ordered set) inference, and to amnesty island violations. As far as simplifying discourse processing is concerned, Prince writes:

> A 'simplifying' Left-Dislocation serves to simplify the discourse processing of discourse-new entities by removing the NPs evoking them from a syntactic position disfavored for NPs evoking discourse-new entities and creating a separate processing unit for them. Once that unit is processed and they have become discourse-old, they (or, rather, the pronouns that represent them) may comfortably occur in their canonical positions within the clause. (p. 286)

The following narrative illustrates:

"I know what this piece of equipment's raised to do. Any **company$_i$**, if they're worth 150 million dollars you don't need to think for a minute **they$_i$**'re not gonna know what you're doin'. They didn't get there that way." (from Studs Terkel's *Working,* p. 46)

Secondly, 'a "poset" Left-Dislocation serves to trigger an inference on the part of the hearer that the entity represented by the initial NP stands in a salient partially ordered set relation to some other entity or entities already evoked in the discourse-model' (Prince 1998: 289). For example:

"She had an idea for a project. She's going to use three groups of mice$_{ijk}$. **One$_i$**, she'll feed **them$_i$** mouse chow, just the regular stuff they make for mice. **Another$_j$**, she'll feed **them$_j$** veggies. And the third$_k$ she'll feed e$_k$ junk food."

And finally, we see an example of left-dislocation amnestying island violations:

"**My first book$_i$**, I paid half of each trick to the person who gave **it$_i$** to me." (from Studs Terkel's *Working,* p. 95) [to avoid ***My first book$_i$**, which I paid half of each trick to the person who gave **e$_i$** to me]

Along the same lines, Prince argues that topicalization has two discourse functions: triggering poset inferences and something like the traditional topic-marking role associated with it.

Not only may syntactic and information structure be out of alignment, but *semantic* structure can diverge from each. Consider, for example, the following data discussed in Vallduví (1992: 113). He points out that two structurally different sentences may have the same information structure:

(84)　　a. The boss hates BROCCOLI.
　　　　b. It is BROCCOLI that the boss hates.

Broccoli is in focus in both sentences and both have the same grounding. But the situation is more complicated, and more interesting, because of the following contrasts:

(85)　　a. She saw JOHN at the party.
　　　　b. It's JOHN that she saw at the party.
(86)　　a. She saw NOBODY at the party.
　　　　b. *It's NOBODY that she saw at the party.

This contrast suggests that the *semantics* of simple SVO and *it*-cleft sentences are different. In focus-ground sentences, the ground is assumed by the speaker to be believed by the hearer. So in the (a) sentences, 'she saw *x* at the party' is assumed to be believed by the hearer. But the cleft in (86b) cannot entail or semantically presuppose 'she saw *x* at the party', since that would be a semantic contradiction.

An analogous example is provided in Ward and Birner (1995) as part of their discussion of the definiteness effect. An old observation is that the definite article normally sounds very strange in sentences beginning with existential *there,* as in (87):

(87)　　?There's the dog running loose somewhere in the neighborhood.

Syntacticians have generally opted for a purely syntactic explanation for the oddness of the definite article (Milsark 1977; Safir 1985; Reuland 1985). However, Ward and Birner have argued convincingly that the correct explanation is a pragmatic one: NPs in such sentences are required to represent a hearer-new entity. Hence (88) is impeccable:

(88)　　There were the same people at both conferences.

Here again we have a generalization involving formal structure (the *there*-construction), semantics (the notion of 'definiteness'), and discourse (the idea of 'hearer-new'). Can OT account for such intercomponent generalizations? Possibly, but it is hard to imagine how it might do so.

To summarize, the built-in need of OT to regard each grammatical sentence as a winner in its own competition set has led to analyses that import aspects of information structure into the syntactic derivation. But the complexity and indirectness of the interface between syntactic structure and information structure shows that such a move is ill-advised.

5.   Conclusion

This paper has presented a critique of optimality-theoretic syntax and has concluded that they have several inherent limitations. The first is a consequence of the hypothesis that constraints are universal, that is, that if a constraint is needed for the grammar of one language, then that constraint occurs in the grammars of all languages. It follows from this fact that a much encountered claim that an OT approach to a particular problem is 'restrictive', in the sense that it limits the class of possible languages, will almost never be correct. There is likely to be some further constraint, not considered in the particular analysis, which could potentially sabotage the restrictiveness claim. The second limitation is particular to the trend within OT that advocates linking each constraint with a function-based 'user constraint'. The paper has argued, however, that the requirement that each constraint be functionally motivated misrepresents the manner in which grammars are responses to external motivation. Finally, OT falters in its treatment of optionality, whether discourse-independent or discourse-dependent. In both cases, it needs to posit otherwise unneeded competition sets and in the latter case it requires the blurring of the distinction between purely syntactic generalizations and those relating syntactic structure to information content.

REFERENCES

Ackema, Peter and Ad Neeleman (1998a). Conflict resolution in passive formation. Lingua 104: 13-29.

Ackema, Peter and Ad Neeleman (1998b). Optimal questions. Natural Language and Linguistic Theory 16: 443-490.

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

Aissen, Judith (2000). Differential object marking: Iconicity vs. economy. Unpublished ms., UCSC.

Alsina, Alex (1996). Passive types and the theory of object asymmetries. Natural Language and Linguistic Theory 14: 673-723.

Anttila, Arto (1997). Variation in Finnish phonology and morphology. Unpublished Ph. D. dissertation, Stanford University.

Bach, Emmon (1974). Syntactic theory. New York: Holt, Rinehart and Winston.

Baker, Mark C. (1988). Incorporation: A theory of grammatical function changing. Chicago: University of Chicago Press.

Baker, Mark C. (1989). Object sharing and projection in serial verb constructions. Linguistic Inquiry 20: 513-553.

Baker, Mark C. (1997). Thematic roles and syntactic structure. Elements of grammar: Handbook of generative syntax, ed. by Liliane Haegeman. Dordrecht: Kluwer, pp. 73-137.

Baker, Mark C. (to appear). Phrase structure as a representation of 'primitive' grammatical relations. Objects and other subjects: Grammatical functions, functional categories, and configurationality, ed. by William Davies and Stanley Dubinsky. Dordrecht: Kluwer, pp. xxx-xxx.

Bakovic, Eric and Edward Keer (2000). Optionality and ineffability. Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, .

Biligiri, H. S. (1965). Kharia phonology, grammar, and vocabulary. Poona. Deccan College.

Blevins, Juliette (1997). Rules in optimality theory: Two case studies. Derivations and constraints in phonology, ed. by Iggy Roca. Oxford: Clarendon Press, pp. 227-260.

Bresnan, Joan W. (1997). The emergence of the unmarked pronoun: Chichewa pronominals in optimality theory. Berkeley Linguistics Society, 23, ???

Bresnan, Joan W. (2000a). Explaining morphosyntactic competition. The handbook of contemporary syntactic theory, ed. by Mark Baltin and Chris Collins. Oxford: Blackwell, pp. 11-44.

Bresnan, Joan W. (2000b). Optimal syntax. Optimality theory: Syntax, phonology, and acquisition, ed. by J. Dekkers, F. van der Leeuw, and J. van de Weijer. Oxford: Oxford University Press, pp. ???

Bresnan, Joan W. and Jonni Kanerva (1989). Locative inversion in Chichewa: a case study of factorization in grammar. Linguistic Inquiry 20: 1-50.

Bresnan, Joan W. and Liobi Moshi (1990). Object asymmetries in comparative Bantu syntax. Linguistic Inquiry 21: 147-185.

Buckley, Eugene (2000). What should phonology explain? Unpublished ms., University of Pennsylvania.

Carrier-Duncan, Jill (1985). Linking of thematic roles in derivational word formation. Linguistic Inquiry 16: 1-34.

Chafe, Wallace L. (1987). Cognitive constraints on information flow. Coherence and grounding in discourse, vol. 11 ed. by Russell Tomlin. Amsterdam: John Benjamins, pp. 21-52.

Choi (2000). ? Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, pp. ?

Chomsky, Noam (1957). Syntactic structures. Janua Linguarum Series Minor 4. The Hague: Mouton.

Chomsky, Noam (1971). Deep structure, surface structure, and semantic interpretation. Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology, ed. by Danny Steinberg and Leon Jakobovits. Cambridge: Cambridge University Press, pp. 183-216.

Chomsky, Noam (1973). Conditions on transformations. A festschrift for Morris Halle, ed. by Steven Anderson and Paul Kiparsky. New York: Holt Rinehart & Winston, pp. 232-286.

Comrie, Bernard (1989). Language universals and linguistic typology: 2nd edition. Chicago: University of Chicago Press.

Comrie, Bernard and Edward L. Keenan (1979). Noun phrase accessibility revisited. Language 55: 649-664.

Costa, João (1998). Parameters vs. soft constraints in the analysis of discourse-configurationality. CLS 34, Part 2: Papers from the panels, ed. by M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson, and Tamra Wysocki. Chicago: Chicago Linguistic Society, pp. 18-30.

Costa, João (2000). ? Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, pp. ?

Creider, Chet (1979). On the explanation of transformations. Syntax and semantics, volume 12: Discourse and syntax, ed. by Talmy Givón. New York: Academic Press, pp. 3-22.

Croft, William (1988). Agreement vs. case marking and direct objects. Agreement in natural language: Approaches, theories, descriptions, ed. by Michael Barlow and Charles A. Ferguson. Stanford, CA: Center for the Study of Language and Information, pp. 159-179.

Croft, William (1990). Typology and universals. Cambridge: Cambridge University Press.

Croft, William (1991). Syntactic categories and grammatical relations. Chicago: University of Chicago Press.

Croft, William (1995). Autonomy and functionalist linguistics. Language 71: 490-532.

Croft, William (2000). Explaining language change: An evolutionary approach. London: Longman.

Culicover, Peter W. and Ray Jackendoff (1999). Control is mostly semantic. Unpublished ms., Ohio State University and Brandeis University.

Denham, Kristin (2000). Optional Wh-Movement in Babine-Witsuwit'en. Natural Language and Linguistic Theory 18: xxx-xxx.

Dixon, R. M. W. (1979). Ergativity. Language 55: 59-138.

Dowty, David (1991). Thematic proto-roles and argument selection. Language 67: 547-619.

Durie, Mark (1985). A grammar of Acehnese. Dordrecht: Foris.

Durie, Mark (1987). Grammatical relations in Acehnese. Studies in Language 11: 365-399.

Emonds, Joseph E. (1970). Root and structure-preserving transformations. Unpublished Ph. D. dissertation, MIT.

Fillmore, Charles J. (1968). The case for case. Universals in linguistic theory, ed. by Emmon Bach and Robert Harms. New York: Holt, Rinehart and Winston, pp. 1-90.

Fischer, Olga C. M. (1992). Syntax. The Cambridge history of the English language, ii. 1066-1476, ed. by N. Blake. Cambridge: Cambridge University Press, pp. 207-408.

Foley, William A. and Robert D. Van Valin (1984). Functional syntax and universal grammar. Cambridge Studies in Linguistics 38. Cambridge: Cambridge University Press.

Givón, Talmy (1979). On understanding grammar. New York: Academic Press.

Givón, Talmy (1984). Syntax: A functional-typological introduction, vol. 1. Amsterdam: John Benjamins.

Grimshaw, Jane (1990). Argument structure. Cambridge, MA: MIT Press.

Grimshaw, Jane (1997). Projection, heads, and optimality. Linguistic Inquiry 28: 373-422.

Grimshaw, Jane (1999). Clause structure and competition in syntax. Unpublished ms., Rutgers University.

Grimshaw, Jane and Vieri Samek-Lodovici (1998). Optimal subjects and subject universals. Is the best good enough?: Optimality and competition in syntax, ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. Cambridge, MA: MIT Press, pp. 193-219.

Gundel, Jeanette K. (1988). Universals of topic-comment structure. Studies in syntactic typology, ed. by Michael Hammond, Edith Moravcsik, and Jessica Wirth. Amsterdam: John Benjamins, pp. 209-239.

Hale, Mark and Charles Reiss (2000). 'Substance abuse' and 'dysfunctionalism': Current trends in phonology. Linguistic Inquiry 31: 157-169.

Hall, Bruce (1998). Grice, discourse representation, and optimal intonation. CLS 34, Part 2: Papers from the panels, ed. by M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson, and Tamra Wysocki. Chicago: Chicago Linguistic Society, pp. 63-78.

Hall, Christopher J. (1992). Morphology and mind: A unified approach to explanation in linguistics. London: Routledge.

Halle, Morris and William J. Idsardi (1997). *r*, hypercorrection, and the Elsewhere Condition. Derivations and constraints in phonology, ed. by Iggy Roca. Oxford: Clarendon Press, pp. 331-348.

Halliday, M. A. K. (1967). Some aspects of the thematic organization of the English clause (= (Research Bulletin RM-5224-PR). Santa Monica: The RAND Corporation.

Haspelmath, Martin (1992). Grammaticalization theory and heads in morphology. Morphology now, ed. by Mark Aronoff. Albany: SUNY Press, pp. 69-82, 194-198.

Haspelmath, Martin (1999). Optimality and diachronic adaptation. Zeitschrift für Sprachwissenschaft 18: 180-205.

Hawkins, John A. (1978). Definiteness and indefiniteness. London: Croom Helm.

Hawkins, John A. (1994). A performance theory of order and constituency. Cambridge Studies in Linguistics 73. Cambridge: Cambridge University Press.

Hawkins, John A. (1995). Argument-predicate structure in grammar and performance: A comparison of English and German. Insights in Germanic linguistics I, ed. by Irmengard Rauch and Gerald F. Carr. Berlin: Mouton de Gruyter, pp. 127-144.

Hayes, Bruce P. (1998). Phonetically driven phonology: The role of optimality theory and inductive grounding. Functionalism and formalism in linguistics, vol. 1: General papers ed. by Michael Darnell, Edith Moravcsik, Frederick J. Newmeyer, Michael Noonan, and Kathleen Wheatley. Amsterdam: John Benjamins, pp. 243-285.

Hoekstra, Teun and Jan G. Kooij (1988). The innateness hypothesis. Explaining language universals, ed. by John A. Hawkins. Oxford: Blackwell, pp. 31-55.

Jackendoff, Ray (1972). Semantic interpretation in generative grammar. Cambridge, MA: Cambridge University Press.

Jackendoff, Ray (1990). Semantic structures. Cambridge, MA: MIT Press.

Jelinek, Eloise and Richard Demers (1983). The agent hierarchy and voice in some Coast Salish languages. International Journal of American Linguistics 49: 167-185.

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

Kaisse, Ellen M. (1985). Connected speech: The interaction of syntax and phonology. Orlando: Academic Press.

Katz, Jerrold and Paul Postal (1964). An integrated theory of linguistic descriptions. Cambridge, MA: MIT Press.

Keenan, Edward L. and Bernard Comrie (1977). Noun phrase accessibility and universal grammar. Linguistic Inquiry 8: 63-99.

Keer, Edward and Eric Bakovic (1997). Have FAITH in syntax. WCCFL 16, 255-269.

Kiefer, Ferenc (1993). Thematic roles and compounds. Proceedings of the XVth International Congress of Linguists, ed. by . Québec: Presses de l'Université Laval, pp. 167-170.

Kiparsky, Paul (1985). Morphology and grammatical relations. Unpublished ms, Stanford University.

Kirby, Simon (1998). Function, selection and innateness: The emergence of language universals. Oxford: Oxford University Press.

Kroch, Anthony (1989). Reflexes of grammar in patterns of language change. Language variation and change 1: 199-244.

Kroch, Anthony (1994). Morphosyntactic variation. Papers from the 30th Regional meeting of the Chicago Linguistic Society, vol. 2: The parasession on variation in linguistic theory ed. by Katharine (and 5 others) Beals. Chicago: Chicago Linguistic Society, pp. 180-201.

Kroch, Anthony and Anne Taylor (1997). The syntax of verb movement in Middle English: Dialect variation and language contact. Parameters of morphosyntactic change, ed. by Ans van Kemenade and Nigel Vincent. Cambridge: Cambridge University Press, .

Lakoff, George (1971). On generative semantics. Semantics: an interdisciplinary reader in philosophy, linguistics, and psychology, ed. by Danny Steinberg and Leon Jakobovits. New York: Cambridge University Press, pp. 232-296.

Lambrecht, Knud (1994). Information structure and sentence form. Cambridge: Cambridge University Press.

Langacker, Ronald W. (1974). Movement rules in functional perspective. Language 50: 630-664.

Langacker, Ronald W. (1990). Concept, image and signal: The cognitive basis of grammar. Berlin: Mouton de Gruyter.

Larson, Richard K. (1988). On the double object construction. Linguistic Inquiry 19: 335-392.

Legendre, Géraldine (1999). Why French stylistic inversion is optimal. unpublished ms., Johns Hopkins University.

Legendre, Géraldine (2000a). An introduction to optimality theory in syntax. Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, pp. ??????

Legendre, Géraldine (2000b). Masked second-position effects and the linearization of functional features. Optimality-theoretic syntax, ed. by

Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, .

Legendre, Géraldine, Paul Smolensky, and Colin Wilson (1998). When is less more? Faithfulness and minimal links in *wh*-chains. Is the best good enough?: Optimality and competition in syntax, ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. Cambridge, MA: MIT Press, pp. 249-290.

Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer, and William Raymond (1995). Optimality and *Wh*-Extraction. University of Massachusetts Occasional Papers 18: Papers in Optimality Theory, ed. by J. Beckman, L. Walsh Dickey, and Urbanczyk. Amherst, MA: GLSA, University of Massachusetts, .

Lightfoot, David (1999). The development of language: Acquisition, change, and evolution. Blackwell/Maryland Lectures in Language and Cognition 1. Oxford: Blackwell.

Lightfoot, David W. (1991). How to set parameters: Arguments from language change. Cambridge, MA: MIT Press.

McCarthy, John (1993). A case of surface constraint violation. Canadian Journal of Linguistics 38: 169-195.

McCloskey, James (1996). Subjecthood and subject positions. A handbook of theoretical syntax, ed. by Liliane Haegeman. Dordrecht: Kluwer, pp. 197-236.

McMahon, April M. S. (2000). Change, chance, and optimality. Oxford: Oxford University Press.

Milsark, Gary (1977). Toward an explanation of certain peculiarities of the existential construction in English. Linguistic Analysis 3: 1-30.

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

Müller, Gereon (1999a). On common-sense justifications of optimality-theoretic constraints. Zeitschrift für Sprachwissenschaft 18: 230-234.

Müller, Gereon (1999b). Optionality in optimality-theoretic syntax. GLOT International 4.5: 3-8.

Müller, Gereon and Wolfgang Sternefeld (1996). A-bar chain formation and economy of derivation. Linguistic Inquiry 27: 480-511.

Nagy, Naomi and Bill Reynolds (1997). Optimality theory and variable word-final deletion in Faetar. Language Variation and Change 9: 37-55.

Newmeyer, Frederick J. (1986). Linguistic theory in America: Second edition. New York: Academic Press.

Newmeyer, Frederick J. (1998a). The irrelevance of typology for linguistic theory. Syntaxis 1: 161-197.

Newmeyer, Frederick J. (1998b). Language form and language function. Cambridge, MA: MIT Press.

Newmeyer, Frederick J. (to appear). Grammatical functions, thematic roles, and phrase structure: Their underlying disunity. Objects and other subjects: Grammatical functions, functional categories, and configurationality, ed. by William Davies and Stanley Dubinsky. Dordrecht: Kluwer, pp. xxx-xxx.

Ostler, Nicholas (1980). A theory of case linking and agreement [truncated version of unpublished 1979 MIT Ph. D. dissertation entitled Case linking: A theory of case and verb diathesis, applied to Classical Sanskrit]. Bloomington: Indiana University Linguistics Club.

Pesetsky, David (1998). Some optimality principles of sentence pronunciation. Is the best good enough?: Optimality and competition in syntax, ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. Cambridge, MA: MIT Press, pp. 337-384.

Pinker, Steven (1984). Language learnability and language development. Cambridge, MA: Harvard University Press.

Primus, Beatrice (1999). Cases and thematic roles: Ergative, accusative, and active. Tübingen: Max Niemeyer.

Prince, Alan and Paul Smolensky (1993). Optimality theory: Constraint interaction in generative grammar. RuCCS Technical Report 2. Piscataway, NJ: Rutgers University Center for Cognitive Science.

Prince, Ellen F. (1998). On the limits of syntax, with reference to left-dislocation and topicalization. Syntax and Semantics, volume 29: The limits of Syntax, ed. by Peter Culicover and Louise McNally. San Diego: Academic Press, pp. 281-302.

Randall, Janet H. (1988). Inheritance. Syntax and semantics, vol. 21: Thematic relations, ed. by Wendy K. Wilkins. New York: Academic Press, pp. 129-146.

Reuland, Eric J. (1985). Representation at the level of logical form and the definiteness effect. Grammatical representation, ed. by Jacqueline Guéron, Hans-Georg Obenauer, and Jean-Yves Pollock. Dordrecht: Foris, pp. 327-362.

Rizzi, Luigi (1997). The fine structure of the left periphery. Elements of grammar: Handbook of generative syntax, ed. by Liliane Haegeman. Dordrecht: Kluwer, pp. 281-337.

Rugemalira, Josephat M. (1994). The case against the thematic hierarchy. Linguistic Analysis 24: 62-81.

Safir, Kenneth J. (1985). Syntactic chains. Cambridge: Cambridge University Press.

Sag, Ivan and Carl Pollard (1991). An integrated theory of complement control. Language 67: 63-113.

Samek-Lodovici, Vieri (2000). ? Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, pp. ?

Sells, Peter (2000). Form and function in the typology of grammatical voice systems. Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, pp. ?

Sells, Peter, John Rickford, and Thomas Wasow (1996). An optimality theoretic approach to variation in negative inversion in AAVE. Natural Language and Linguistic Theory 14: 591-627.

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

Speas, Margaret J. (1990). Phrase structure in natural language. Studies in Natural Language and Linguistic Theory 21. Dordrecht: Kluwer.

Speas, Margaret J. (1997). Optimality theory and syntax: null pronouns and control. Optimality theory: An overview, ed. by Diana Archangeli and D. Terence Langendoen. Oxford: Blackwell, .

Sternefeld, Wolfgang (1997). Comparing reference sets. Economy in linguistic theory, ed. by M. Bierwisch, H.M. Gärtner, and C. Wilder. Berlin: Akademieverlag, pp. 81-114.

Tesar, Bruce and Paul Smolensky (1998). Learnability in optimality theory. Linguistic Inquiry 29: 229-268.

Tomlin, Russell S. (1985). Interaction of subject, theme, and agent. Beyond the sentence: Discourse and sentential form, ed. by Jessica Wirth. Ann Arbor: Karoma, pp. 61-80.

Vallduví, Enric (1992). The informational component. New York: Garland.

Van Valin, Robert D. and Randy J. Lapolla (1997). Syntax: Structure, meaning, and function. Cambridge: Cambridge University Press.

Vikner, Sten (2000). V-to-I movement and 'do' insertion in optimality theory. Optimality-theoretic syntax, ed. by Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, MA: MIT Press, .

Ward, Gregory and Betty Birner (1995). Definiteness and the English existential. Language 71: 722-742.

Wedgwood, Daniel (1995). Grammaticalization by reanalysis in an adaptive model of language change: A case study of the English genitive constructions. Unpublished MA thesis, University of Edinburgh.

Wilkins, Wendy K. (1988). Thematic structure and reflexivization. Syntax and semantics, vol. 21: Thematic relations, ed. by Wendy K. Wilkins. New York: Academic Press, pp. 191-214.

Witherspoon, Gary (1977). Language and art in the Navajo universe. Ann Arbor: University of Michigan Press.