

Parametric versus functional explanations of syntactic universals

MARTIN HASPELMATH

February 2006

Abstract: This paper compares the generative principles-and-parameters approach to explaining syntactic universals to the functional-typological approach and also discusses the intermediate approach of Optimality Theory. It identifies some fundamental differences between generative parametric explanations and functional explanations. Most importantly, generative explanations assume that cross-linguistic generalizations are due to the innate Universal Grammar, whereas functional explanations assume that language structure can be influenced by regularities of language use through language change. Despite these differences, both approaches to cross-linguistic similarities and differences seem to be guided by a similar vision: That the superficial structural diversity of languages can be reduced to a few basic patterns once one digs below the surface (macroparameters or holistic types). Unfortunately, the evidence for such reductionist constructs has never been very good, so more recently both generativists and functionalists have shifted their interests away from these ambitious goals. However, I argue that there is one type of cross-linguistic generalization for which there is very good empirical evidence: intra-domain universals relating to prominence scales, most of which are straightforwardly explained functionally in terms of processing difficulty.

1. Introduction: the relevance of observable universals

Many generative and functional linguists share a primary interest in understanding structural similarities and differences between languages. It is clear that the observed language structures represent only a tiny fraction of the logically possible structures, and for a whole range of syntactic domains linguists have a good initial idea of what the limits on structural variation (i.e. the syntactic universals) are. But why should languages be limited in these particular ways in choosing their structures? Generative and functional linguists have given very different answers to this question, but these answers are rarely compared to each other, and the differences in the methodological foundations of the two research programmes are not well understood by most linguists. The present paper aims to elucidate this situation. I will not attempt to hide my personal preference for the functional approach, but I will highlight strengths and weaknesses of both approaches.

I should note at the outset that the goal of explaining language universals is not usually taken as the main goal of generative research. The following quotation is typical:

"The comparative approach in the generative tradition addresses the following questions: (i) what is knowledge of language? (ii) how is this knowledge acquired?... In order to answer these questions we have to identify which linguistic properties can vary across languages and which are constant." (Haegeman 1997:1)

According to this view, the identification of universals is a means to the central goals of characterizing knowledge of language and understanding how it can be acquired. In this contribution, I will not say anything on these more ambitious goals, which I regard as beyond our reach at the moment, or at least as lying outside the realm of linguistics proper. What linguists are primarily responsible for is "observational adequacy", i.e. descriptions of

language systems that are sufficient to predict and explain speaker behaviour. The standard tools of corpus analyses and acceptability judgements are apparently not sufficient to even come close to "descriptive adequacy" (in Chomsky's sense), i.e. descriptions that reflect speakers' mental patterns (their knowledge of language).¹ This view, which in its pessimism differs both from generative linguistics and from cognitive linguistics (e.g. Croft & Cruse 2004), does not preclude the goal of identifying syntactic universals and finding explanations for them.

It has sometimes been asserted that linguistic universals should not be construed as properties of observable systems of speaker behaviour, but as properties of mental systems. For example, Smith (1989:66-67) writes:

"Greenbergian typologists try to make generalizations about *data*, when what they should be doing is making generalizations across rule systems."

This apparently means that Greenbergians generalize over abstract systems (regardless of their cognitive realization), whereas true universals will only be found by generalizing over cognitive systems (see also Coopmans 1983:567(?)). It is easy to agree that a systematic comparison of cognitive systems would be a worthwhile task, but unfortunately it is totally impractical, because we know so little about what these systems might be (see note 1). Even for well-studied languages like English or Japanese, there is hardly anything that cognitively oriented syntacticians agree about, and even within a single research tradition (such as the Chomskyan tradition), there may be new guiding ideas every 10 years or so that lead linguists to very different analyses of basically the same data. Thus, cross-linguistic generalizations over known cognitive systems do not seem to be on the order of the day for the present generation of syntacticians.

Instead, what all comparative linguists do in practice is comparing abstract (Platonic) rule systems that approximately model the observed behaviour of speakers in spontaneous language use and acceptability judgements. This is a much more tractable task that is clearly within our reach. There are hundreds of descriptions of languages from around the world that are accessible to any linguist because they only use widely understood theoretical vocabulary ("basic linguistic theory"), and these are typically the basis for typological generalizations in the Greenbergian tradition. In the Chomskyan tradition, too, almost all comparisons are based on (what are openly acknowledged to be) provisional and largely conventional analyses.² Such analyses may be claimed to be about cognitive rule systems, but in practice this has little relevance for the comparative syntactician's work. Those generative syntacticians that are considering a fairly large number of languages have

¹ The main reason for this assessment is the vast gulf separating linguists' views, as reflected in current textbooks on different cognitively oriented approaches to syntax, e.g. Taylor (2002) (on Cognitive Grammar) and Adger (2003) (on generative grammar). It could of course turn out that one of the two approaches is exactly on the right track and the other approach is totally wrongheaded. But I find it much more likely that while both capture a few aspects of the truth, they are both groping in the dark, equally distant from their common goal of mentally realistic description, which is simply too ambitious.

² The extent to which analyses are based on convention (rather than empirical or conceptual motivation) is rarely explicitly recognized, but in private many linguists will frankly say that the theoretical framework they teach or work in is motivated as much by social convenience as by their convictions. (A personal anecdote: When I was still a Ph.D. student, a well-known generative linguist with careers at MIT and Stanford advised me to state my insights in generative terms, so that my work would be more widely recognized.)

necessarily based their conclusions also on secondary sources from outside the Chomskyan tradition (e.g. Baker 1988, 1996, Ouhalla 1991, Freeze 1992, Cinque 1999, 2005, Johannessen 1999, Julien 2002, Neeleman & Szendrői 2005), and this kind of work has not (at least not widely) been rejected by generative linguists.

Thus, while the relevance of abstract (non-cognitive) rule systems to cognitively oriented linguistics has often been vigorously denied (e.g. Lightfoot 1999:79-82), in practice the difference is not so great, and it makes perfect sense to consider the contribution of generative syntax to the discovery and explanation of phenomenological universals, i.e. universals of grammar as observed in speakers' behaviour, regardless of what cognitive structures might underlie such behaviour (see Haspelmath 2004a for the terminological distinction between "phenomenological" and "cognitive" descriptions and universals).

Many generative (and other cognitively oriented) linguists will immediately object that they are not interested in phenomenological universals, only in the subset of cognitive universals (note that all cognitive universals must be at the same time phenomenological universals, because structures that are not cognitively possible will not show up in speakers' behaviour). There may be all kinds of phenomenological universals that have nothing to do with cognitive structures (e.g. on the lexical level, the fact that all languages have a word for 'moon', which clearly has to do with the fact that the moon is salient in all human habitats). But the problem is that it is impossible to know in advance which of the phenomenological universals are cognitive universals and which ones are not.³ It has been the uniform practice of generative linguistics to simply assume by default that observed universals are relevant theoretically, i.e. that they have a cognitive basis. Thus, I will assume here that all phenomenological universals are interesting both for generative linguists and for functional linguists, because all of them could turn out to be cognitively based or functionally based. Which explanation is correct (the parametric or the functional explanation) is precisely the issue that is the main topic of the paper.

Although the statement by Haegeman (1997) cited at the beginning of this section, which gives primacy to the characterization and acquisition of knowledge of language, is the more orthodox generative position, in practice many generative comparative syntacticians seem to be just as much interested in explaining syntactic universals. Even Noam Chomsky has occasionally highlighted the role of his theory of Universal Grammar in explaining syntactic phenomena:

"The next task is to explain why the facts are the way they are... Universal Grammar provides a genuine explanation of observed phenomena." (Chomsky 1988:61-62)

Most tellingly, generative linguists have usually not refrained from proposing UG-based explanations for universals that are amenable to functional explanations. If their only real interest were the characterization of UG, one would expect them to leave aside all those universals that look functional, and concentrate on those general properties of language that seem arbitrary from a communicative point of view. But this is not what is happening. For

³ See Kiparsky (2004) for an attempt to identify a number of properties that distinguish non-cognitive universals ("typological generalizations") from cognitive universals ("universals").

instance, as soon as topic and focus positions in phrase structure had been proposed, generativists began exploring the possibility of explaining information structure ("functional sentence perspective") in syntactic terms, rather than leaving this domain to pragmaticists. Particularly telling is Aissen's (2003) discussion of Differential Object Marking (DOM), which is known to have a very good functional explanation (e.g. Comrie 1981, Bossong 1985, Filimonova 2005). Aissen does not consider the possibility that this functional explanation makes an explanation in terms of UG superfluous, stating instead that

"the exclusion of DOM from core grammar comes at a high cost, since it means that there is no account forthcoming from formal linguistics for what appears to be an excellent candidate for a linguistic universal." (Aissen 2003:439)

Thus, the reason that DOM was not tackled seriously before by generativists seems to be the fact that the generative tools did not seem suitable to handle it, until Optimality Theory and harmonic alignment of prominence scales entered the stage.

I conclude from all this that the topic of this paper, the explanation of observable (phenomenological) syntactic universals, is highly relevant to both mainstream generative linguistics and mainstream functional linguistics.⁴

2. The parametric approach to explaining syntactic universals

2.1. The basic idea: principles and parameters

The principles-and-parameters (P&P) approach to cross-linguistic differences is compelling in its simplicity and elegance:

"We can think of the initial state of the faculty of language as a fixed network connected to a switch box; the network is constituted of the principles of language, while the switches are the options to be determined by experience. When the switches are set one way, we have Swahili; when they are set another way, we have Japanese. Each possible human language is identified as a particular setting of the switches—a setting of parameters, in technical terminology." (Chomsky 2000:8)

This approach does two things at the same time: First, it explains how children can acquire language (Haegeman's second goal), because "they are not acquiring dozens or hundreds of rules; they are just setting a few mental switches" (Pinker 1994:112). Second, it offers a straightforward way of explaining implicational universals, the type of universals that comparative linguists have found the most intriguing, and that are attested the most widely. If a parameter is located at a relatively high position in the network of

⁴ A final potential objection from an innatist point of view should be mentioned: If one thinks that "all languages must be close to identical, largely fixed by the initial state" (Chomsky 2000:122), then one should seek explanations for those few aspects in which languages differ, rather than seeking explanations for those aspects in which languages are alike (see also Baker 1996: ch. 11). Again, while such an approach may be attractive in theory, it is not possible in practice: While it is a meaningful enterprise to list all known universals (as is attempted in the Konstanz Universals Archive, <http://ling.uni-konstanz.de/pages/proj/sprachbau.htm>), it would not be possible to list all known non-universals, and one would not know where to begin explaining them.

categories and principles, in will have multiple consequences in the observed system. The simplest example is the order of verb and object, noun and genitive, adposition and complement, etc.: If the parameter is located at the higher level of head and complement (assuming that verb/object, noun/genitive, adposition/complement can be subsumed under these categories), setting the head-ordering switch just once gives us the order of a range of lower-level categories, thus accounting for the implicational relationships among their orderings. Thus, "small changes in switch settings can lead to great apparent variety in output" (Chomsky 2000:8; see also Baker 2001a:ch.3 for extensive discussion of this point).

This research programme proved very attractive, and it is usually described as "highly successful, leading to a real explosion of empirical inquiry into language of a very broad typological range" (Chomsky 2000:8; see also Newmeyer 2005:40, in an otherwise very critical context). One even reads occasionally that it provided the solution for Plato's Problem (the problem of language acquisition despite the poverty of the stimulus), so that deeper questions can now be asked (Grewendorf 2002:99, Chomsky 2004, Fukui & Zushi 2004:11).⁵

2.3. Criteria for success

According to the principles-and-parameters vision, it should be possible at some point to describe the syntax of a language by simply specifying the settings of all syntactic parameters of Universal Grammar. We would no longer have any need for thick books with titles like *The syntax of Haida* (cf. Enrico's (2003) 1300-page work), and instead we would have a simple two-column table with the parameters in the first column and the positive or negative settings in the second column. There would not be many parameters ("a few mental switches", Pinker 1994:112, "there are only a few parameters", Adger 2003:16), perhaps 20 (Fodor 2003:734), perhaps 50-100 (Roberts & Holmberg 2005), and at any rate a number comparable to the number of chemical elements in the periodic table (Baker 2001a: ch. 6). Reducing the description of the syntax of a language from 1300 pages to one or two pages would truly be a spectacular success that is worth the effort.

Since carrying out the research programme is not an easy task, we may still be a few decades (or more) away from this ultimate goal. But what progress has been made over the last quarter century, since the P&P programme was inaugurated? Certainly one might reasonably expect the following indications of progress:

- (i) More and more reduction of previously isolated phenomena to independently known parameters;
- (ii) more and more parameters for which there is a consensus in the field;
- (iii) parameters proposed on the basis of variation within one language family are found to be sufficient to predict variation also in totally unrelated language families and areas of the world;
- (iv) broad cross-linguistic research involving many languages is increasingly successful, and generative typological research begins

⁵ But one also reads more cautious assessments. Chomsky (2000:8) reminds the reader that the P&P approach "is, of course, a program, and it is far from a finished product...one can have no certainty that the whole approach is on the right track".

to converge with typological research that does not share the P&P vision (cf. Haider 2001:291-292).

Whether progress has indeed been made along these lines is not easy to say, and one's assessment will of course depend on one's perspective. Baker (2001) offers a highly optimistic view of the progress of generative comparative syntax, comparing it to the situation in chemistry just before the periodic table of elements was discovered: "We are approaching the stage where we can imagine producing the complete list of linguistic parameters, just as Mendeleev produced the (virtually) complete list of natural chemical elements" (p. 50). In the following subsection, I will present a more sober assessment (see also Newmeyer 2004, 2005:76-103).

2.3. Assessment of success

It is certainly the case that more and more phenomena have been studied from a P&P point of view over the last 25 years: the Chomskyan approach has been consistently popular in the field, and a large number of diverse languages have been examined by generative syntacticians. But to what extent the observed variation has been reduced to previously known parameters is difficult to say. One might expect someone to keep a record of the parameters that have been proposed (analogous to the Konstanz Universals Archive, which records proposed universals), or at the very least one might expect handbooks or textbooks to contain lists of well-established parameters. But the voluminous *Oxford Handbook of Comparative Syntax* (Cinque & Kayne 2005, 977 pages) contains no such list. It has language, name and subject indexes, but no parameter index. The subject index just allows one to find references to a handful of parameters. Likewise, textbooks such as Culicover 1997, Biloa 1998, Ouhalla 1999 (all of them with "parameter" in their title) do not contain lists of parameters. The only laudable exception is Roberts 1997, who gives a short list of parameters at the end of each chapter. But in general it is very difficult even to find out which parameters have been proposed in the literature.

It seems to me that one reason for this lack of documentation of proposed parameters is that most of them are not easy to isolate from the assumptions about UG in which they are embedded, and these differ substantially from author to author and from year to year. Even if all works stated the parameters they propose explicitly in quotable text (many do not), these proposals would be difficult to understand without the context. The highly variable and constantly shifting assumptions about UG are thus a serious obstacle for a cumulative process of acquisition of knowledge about parameters. One would hope that in general the best proposals about UG are adopted, and that the general changes in assumptions are in the right direction. But so far it seems that there is too little stability in our picture of UG to sustain a cumulative process of parameter detection, where one linguist can build on the discoveries (not merely on the generalizations and ideas) of another linguist.

It is also not clear whether there are more and more parameters for which there is a consensus in the field. In the absence of an authoritative list, I examined 16 textbooks, overview books and overview articles of generative syntax for the parameters that they mention or discuss (Atkinson 1994, Biloa 1998, Carnie 2002, Culicover 1997, Fanselow & Felix 1987, Freidin 1992, Fukui

1995, Haegeman 1997, Haider 2001, Lightfoot 1999, Jenkins 2000, McCloskey 1988, Ouhalla 1999, Pinker 1994, Radford 2004, Smith 1989). The result is shown in Table 1.

Table 1.

Head-directionality parameter	Smith 1989: 69-71, Freidin 1992:118-120, Chomsky & Lasnik 1993:518, Atkinson 1994, Pinker 1994:111-112, Fukui 1995:329, Culicover 1997: 160, Haegeman 1997, Biloa 1998:20, Lightfoot 1999:59, Ouhalla 1999:298ff., Jenkins 2000:77, Carnie 2002:18, 130, Radford 2004:16-25
Pro-drop parameter	Fanselow & Felix 1987: 139-140, Freidin 1992:118-120, Atkinson 1994, Fukui 1995:3334, Haegeman 1997, Ouhalla 1999:ch. 12, Jenkins 2000:77, Haider 2001: 285-287, Radford 2004:16-25
Subjacency parameter	Fanselow & Felix 1987: 138, McCloskey 1988:25, Freidin 1992:118-120, Atkinson 1994, Fukui 1995:333, Ouhalla 1999:298ff., Haider 2001:284-285
<i>Wh</i> -movement parameter	Chomsky & Lasnik 1993:513, Fukui 1995:332-333, Ouhalla 1999:298ff., Haider 2001:289-290, Radford 2004:16-25
Verb-movement parameter	Haegeman 1997, Ouhalla 1999:ch. 16, Radford 2004:164
Verb-second parameter	Lightfoot 1999:83, Ouhalla 1999: ch. 13, Haider 2001:287-288
Configurationality parameter	Fukui 1995:330, Haider 2001:288

Table 1. Some prominent examples of parameters

There are only 7 parameters which are mentioned in these works, presumably because the textbook and overview authors sensibly limited themselves to less controversial parameters. However, it is difficult to argue that the number of parameters for which there is a consensus has increased over the years. Let us look at the main parameters in Table 1 one by one.

The **head-directionality parameter** is the easiest parameter to explain to readers with little background knowledge, so it is the most widely used example for parameters. However, since Travis's work of the 1980s (e.g. Travis 1989) it has been clear that the simplistic version of the parameter that the textbooks mention can at best be part of the story, and Kayne's (1994) wholesale attack on it has been very influential. A big problem is that the Greenbergian word order generalizations are only tendencies, and parameters can only explain perfect correlations, not tendencies (Baker & McCloskey 2005). Another problem is that assumptions about heads and dependents have varied. Determiners used to be considered specifiers (i.e. a type of dependent), but now they are mostly considered heads. Genitives used to be considered complements, but now they are generally considered specifiers.

And since the 1990s, almost all constituents are assumed to move to another position in the course of the derivation, so that it is the landing sites, so that it is the landing sites, not the underlying order of heads and nonheads that determines word order. Finally, Dryer (1992) and Hawkins (1994, 2004) have argued that the factor underlying the Greenbergian correlations is not head-nonhead order, but branching direction. As a result of all this, not many generative linguists seem to be convinced of this parameter.

The **pro-drop parameter** (or null-subject parameter) is not a good example of a successful parameter either, because the various phenomena that were once thought to correlate with each other (null thematic subjects; null expletives; subject inversion; that-trace filter violations; Rizzi 1982) do not in fact correlate very well (see Newmeyer 2005:88-92). Haider (1994:383) notes in a review of Jaeggli & Safir (1989): "The phenomenon "null subject" is the epiphenomenon of diverse constellations of a grammar that are independent of each other...The search for a unitary pro-drop parameter therefore does not seem to be different from the search for a unitary relative clause or a unitary passive." Roberts & Holmberg (2005: [7]) are more optimistic than Haider: In their rebuttal of Newmeyer (2004), they state that they find Gilligan's (1987) result of four valid implications (out of 24 expected implications) "a fairly promising result". However, as Newmeyer notes, three of these four correlations are not noticeably different from the null hypothesis (the absence of overt expletives would be expected anyway because it is so rare), and for the remaining correlation (subject inversion implying that-trace filter violations), Gilligan's numbers are very small.

The **subjacency parameter** was famous in the 1980s, but it seems that it always rested exclusively on a contrast between English and Italian. It has not apparently led to much cross-linguistic research. According to a recent assessment by Baltin (2004:552), "we have no clear account of the parameter, much less whether the variation that is claimed by the positing of the parameter exists."

The **configurationality parameter** is the only other parameter that is a "deep" parameter in the sense that its setting can "lead to great apparent variety in the output". As with the other parameters, there is a great diversity of approaches among generative linguists (see Hale 1983, Jelinek 1984, the papers in Marácz & Muysken 1989, as well as Baker 2001b for an overview). The lack of consensus for this parameter is perhaps less apparent than in the case of the head-directionality and pro-drop parameters, but this also has to do with the fact that far fewer linguists work on languages that are relevant to this parameter.⁶

The **wh-movement parameter** is not problematic, but it is not a deep parameter and does not entail correlations, so it is not relevant to the explanation of universals. Likewise, the **verb movement parameters** do not entail correlations, unless one assumes that verb movement is conditioned by particular morphological properties of the verb (e.g. Vikner 1997). But this latter point is known to be very problematic and controversial (see Roberts 2003 for an overview of verb movement).

The verb movement parameters also illustrate another widespread limitation of parametric explanations: Very often the range of data considered

⁶ The polysynthesis parameter, Baker's (1996) prime example of a macroparameter, did not make it on the list because it is hardly discussed in the literature apart from Baker's work. This is a great pity, because in my view, the correlations reported by Baker in his ch. 11 constitute the most interesting typological hypothesis in the entire generative literature.

come from a single language family or subfamily, and there is no confirming evidence from other families (this is also deplored in Baker & McCloskey 2005). Thus, Vikner (1997) considers only Germanic languages plus French (another Indo-European language with close historical ties to Germanic), and Roberts & Holmberg's (2005) showcase example of a parameter that captures valid cross-linguistic correlations concerns exclusively the Scandinavian branch of the Germanic subfamily of Indo-European. Kayne (2000) discusses a number of parameters that are relevant exclusively to Romance.

Another widespread approach is to compare just two widely divergent languages, and to try to connect the differences observed between them by proposing a small number of parameters. The most typical example of this approach is the language pair Japanese/English (see, e.g., Kuroda 1988, Fukui 1995) (but also French and English, Kayne 1981).

So far it seems that sophisticated parameters such as those that were developed on the basis of a single family or a pair of languages have rarely (if ever) been confirmed by evidence coming from unrelated families in different parts of the world.

Finally, I see no signs of an increasing amount of broad cross-linguistic research in the principles-and-parameters approach. Works such as those by Johannessen (1999), Julien (2002), and Cinque (2005) remain exceptional and have not been particularly influential in the field. Baker (1988) was really the only truly influential work that adopts a broad cross-linguistic approach, and although Baker (1996) proposed a much more interesting parameterization hypothesis, it was much less well received in the field.

2.4. Abandoning deep parameters

The divergence between the promises of the research programme as sketched in §2.1 and the actual research results has not only been noticed by outside observers, but also by practitioners of generative syntax. Baker (1996:7) notes:

"One might expect that more and more parameters comparable to the Pro-Drop Parameter would be discovered, and that researchers would gradually notice that these parameters...themselves clustered in nonarbitrary ways...It is obvious to anyone familiar with the field that this is not what has happened."

Pica (2001:v-vi), in the introduction to a new yearbook on syntactic variation as studied from a Chomskyan perspective, states quite bluntly:⁷

"Twenty years of intensive descriptive and theoretical research has shown, in our opinion, that such meta-parameters [e.g. the Null-Subject Parameter, or the Polysynthesis Parameter] do not exist, or, if they do exist, should be seen as artefacts of the 'conspiracy' of several micro-parameters."

Kayne (1996) justifies his shift to microparameters on somewhat different grounds: He does not explicitly doubt the existence of macroparameters, but he simply finds small-scale comparative syntax studying closely related languages "more manageable" because the space of hypotheses is smaller.⁸ And Newmeyer (2004, 2005) advocates abandoning parameters altogether.

⁷ A similar quote can be found in Baltin (2004:551): "I have never seen convincing evidence for macroparameters."

⁸ Curiously, Kayne does not seem to be concerned that by limiting himself to closely related languages, he runs the risk that he will discover shared innovations that have a purely

The shift from macroparameters that explain clusterings of properties to microparameters is often connected with the claim that parameters are exclusively associated with lexical items, and this is regarded as a virtue, because it restricts the range of possible parameters and thus reduces the burden of the language learner (see Baker 1996:7-8 and Newmeyer 2005:53-57 for discussion). According to Chomsky (1991:26), this view opens up the possibility that "in an interesting sense, there is only one possible human language". This conclusion may be welcome from the point of view of learnability, but it also ultimately means that generative syntax abandons the claim that it can contribute to understanding cross-linguistic diversity and regularity.

Of course, not everyone in the field shares this approach (see Baker 2006), but it seems fair to say that the enthusiasm for parameters that was widespread in the 1980s has been replaced by a more sober attitude, and that this is not only due to the fact that Chomsky has lost interest in them, but also to the fact that the results of parametric research over the last two decades were much more modest than had been hoped.

2.5. UG-based explanation without parameters

Deep parameters are not the only way in which explanations for syntactic universals can be formulated in a generative approach that assumes a rich Universal Grammar. Syntactic universals could of course be due to nonparameterized principles of UG. For instance, it could be that all languages have nouns, verbs and adjectives because UG specifies that there must be categories of these types (Baker 2003). Or it could be that syntactic phrases can have only certain forms because they must conform to an innate X-bar schema (Jackendoff 1977). Or perhaps all languages conform to the Linear Correspondence Axiom (dictating uniform specifier-head-complement order), so that certain word orders cannot occur because they cannot be generated by leftward movement (Kayne 1994).

Such nonparametric explanations are appropriate for unrestricted universals, whereas (deep) parameters explain implicational universals. What both share is the general assumption that empty areas in the space of logically possible languages should be explained by restrictions that UG imposes on grammars: Nonexisting grammars are not found because they are in conflict with our genetic endowment for acquiring grammars. They are computationally impossible, or in other words, they cannot be acquired by children, and hence they cannot exist. This widespread assumption has never been justified explicitly, as far as I am aware, but it has rarely been questioned by generative linguists (but see now Hale & Reiss 2000 and Newmeyer 2005). We will see in §4 that it is not shared by functionalists.

But before moving on to functionalist explanations of universals, I will discuss Optimality Theory, which is in some ways intermediate between the Chomskyan P&P approach and functionalist approaches.

historical explanations, rather than properties that are shared because of the same parameter setting.

3. The constraint-ranking approach to explaining syntactic universals

3.1. Constraint-ranking vs. parameters

What standard Optimality Theory (OT) shares with the Chomskyan principles-and-parameters approach is the idea that universals should fall out from the characterization of UG: Unattested languages do not exist because they are incompatible with UG and cannot be acquired. Observed universals are assumed to be cognitive universals. Instead of principles, OT has constraints, and instead of open parameters, OT has cross-linguistically variable ranking of constraints. Thus, the difference between English *it rains* (**rains*) and Italian *piove* (**esso piove*) can be explained by different rankings of the same universal constraints SUBJECT ("the highest specifier position must be filled") and FULLINT(ERPREDICTION) ("lexical items must contribute to the interpretation of the structure") (Grimshaw & Samek-Lodovici 1998, Legendre 2001). In English, SUBJECT is ranked above FULLINT (so that the expletive *it* is allowed and required), while in Italian, FULLINT is ranked higher than SUBJECT (so that the expletive *esso* is disallowed and the sentence can remain subjectless).

The possible objection that OT is thus not radically different from P&P is countered by advocates of OT by pointing out that "a parameter that is "off" is completely inactive. But a constraint that is crucially dominated can still be active" (McCarthy 2002:242). For example, in Italian the dominated constraint SUBJECT is not always irrelevant: When a subject does not have a topic antecedent, the subject appears overtly to satisfy the constraint SUBJECT. Thus, OT constraints can account for between-language variation and for within-language variation, unlike parameters, which only govern between-language variation (McCarthy 2002:110).

A big difference in practice between OT and P&P is that OT analyses must be explicit about which constraints are assumed and how they are ranked. The constraints are given catchy names (to fit the column heading of a constraint tableau), and their content is typically spelled out in a separate paragraph of the text, as in (5) (from Legendre 2001:5):

(5) FULLINT: Lexical items must contribute to the interpretation of a structure.

Since the constraints have conveniently short (and typographically highlighted) names, it is easy to list them and index them, and some books even have a separate index of constraints (e.g. Kager 1999, Müller 2000). Thus, it is relatively straightforward to determine how many and which constraints have been posited. This is in contrast to parameters, which as we saw in §2.3 are rarely listed, and often not even explicitly spelled out.

The forced explicitness of OT about its constraints reveals that a very large number of them are apparently needed: certainly many hundreds, if not many thousands. This is in striking contrast with low estimates for parameters ("a few mental switches"), though it has been observed that the literature on parameters also contains a fairly large number of them.⁹ Perhaps the OT

⁹ According to Lightfoot (1999:259), "if there are only 30-40 structural parameters, then they must look very different from present proposals...a single issue of the journal *Linguistic Inquiry* may contain 30-40 proposed parameters" (see also Newmeyer 2005:81-83).

approach simply forces its practitioners to be more honest about the assumptions they make about UG.

According to McCarthy (2002:1), "one of the most compelling features of OT...is the way that it unites description of individual languages with explanation of language typology". It is standardly assumed that the constraints are universal, and that the cross-linguistic variation exclusively derives from variation in constraint ranking. Thus, "if the grammar of one language is known exhaustively, then in principle the grammars of all possible human languages are known. Analysis of one language can never be carried out without consideration of the broader typological context" (McCarthy 2002:108). This is actually quite similar in spirit to the P&P approach, whose vision is the exhaustive description of a language by simply determining the settings of all the parameters, once they have been discovered (see §2.2). The difference is that P&P has never come particularly close to this vision, not even in the description of small fragments of a grammar, and it is common to find P&P works that make very little reference to parameters. Although in principle every P&P analysis of a particular language makes predictions about other possible languages, the analyses are typically so complicated that it is difficult to see what exactly is predicted. OT is more straightforward: "To propose a constraint ranking for one language in OT is to claim that all possible rerankings of these constraints yield all and only the possible human languages" (Legendre 2001:15).

However, in practice this has not led to a large number of broad cross-linguistic syntactic studies in the OT framework, so the success of this approach is not yet apparent. Legendre (2001:15), in the introduction to Legendre et al. (eds.) 2001, mentions only a single paper in the entire volume as an example of "typology by reranking". And Sells (2001:10-11), in the introduction to Sells (ed.) (2001), states: "Typological predictions are brought out most directly in the papers here by Choi, Morimoto and Sharma, but are implicit in most of them." So in actual practice, OT tends to be like P&P in that authors mostly aim at insightful descriptions of particular languages. There is of course nothing bad about limiting oneself to a language one knows well, but the skeptical reader cannot help being struck by the contrast between the ambitious rhetoric and the typically much more modest results.

3.2. Towards functionalism

But there are two ways in which OT approaches have gone beyond the Chomskyan thinking and practice, and have moved towards the functionalist approach that will be the topic of §4. One is that they have sometimes acknowledged the relevance of discoveries and insights coming from the functional-typological approach. Thus, Legendre et al. (1993) refer to Comrie, Givón, and Van Valin; Aissen (1999) refers to Chafe, Croft, Kuno and Thompson; and Sharma (2001) refers to Bybee, Dahl, Greenberg, and Silverstein. Such references are extremely rare in P&P works. Thus in OT syntax, one sometimes gets the impression that there is a genuine desire for a cumulative approach to explaining syntactic universals, i.e. one that allows researchers to build on the results of others, even if these do not share the same outlook or formal framework.

More importantly, the second way in which some OT authors have moved towards functionalists is that they have allowed functional considerations to play a role in their theorizing. For instance, Bresnan (1997) provides a

functional motivation for the constraint PROAGR ("Pronominals have agreement properties"): "The functional motivation for the present constraint could be that pronouns...bear classificatory features to aid in reference tracking, which would reduce the search space of possibilities introduced by completely unrestricted variable reference." Aissen (2003) explicitly links the key constraints needed for her analysis to two of the main factors that functionalists have appealed to: iconicity and economy. The idea that OT constraints should be "grounded" in functional considerations is even more widespread in phonology (e.g. Boersma 1998, Hayes et al. 2004), but the trend can be seen in syntax, too. Another reflection of this trend is the appearance of papers that explicitly take a functionalist stance and use OT just as a formalism, without adopting all the tenets of standard generative OT (e.g. Nakamura 1999, Malchukov 2005).

I believe that the affinity between OT and functionalist thinking derives from the way in which OT accounts for cross-linguistic differences: through the sanctioning of constraint violations by higher-ranked constraints. The idea that cross-linguistic differences are due to different weightings of conflicting forces has been present in the functionalist literature for a long time (cf. Haspelmath 1999a:180-181). OT's contribution is that it turned these functional forces into elements of a formal grammar and devised a way of integrating the idea of conflict and competition into the formal framework.

3.3. Explaining the constraints

However, an important question has not been properly answered by OT proponents: Why are the constraints the way they are? Or in other words: What are the constraints on the constraints? After all, only a small part of the logical possibilities seem to be realized as constraints. For example, the set of constraints does not seem to contain the constraints EMPTYINT ("Lexical items must not contribute to the interpretation") or OBJECT ("The complement of the V position must be filled"). One possible answer is given by McCarthy (2002:240): "If the constraints are universal and innate, we cannot presuppose that there are "constraints on constraints", except for physical limitations imposed by human biology and genetics." This is similar to classical P&P, where no strong need was felt to explain why the principles of UG are the way they are (at least until Chomsky inaugurated the Minimalist Programme). But the P&P principles were conceived of as few in number and highly abstract, whereas OT constraints are numerous and often fairly concrete. Moreover, it is not uncommon for OT authors to explicitly justify constraints by external (sometimes functional) considerations. But if it is simultaneously assumed that constraints are part of UG and externally conditioned, then something crucial is missing: an account of how the external factors can impinge on UG, which is presumably part of the human genotype (see Haspelmath 1999a for further discussion).

Moreover, if the constraints can be explained functionally, one wonders what the relation between functional explanation and explanation by the formal system is. A striking example of a certain kind of incoherence of argumentation is found in Aissen (2003), one of the most widely cited papers in OT syntax, which deals with Differential Object Marking (DOM). Aissen assumes an elaborate system of constraints that she associates with concepts derived from functional linguistics: the animacy and definiteness scales, iconicity, and economy. The animacy constraints are *OJ/HUM ("an object

must not be human"), *OJ/ANIM ("an object must not be animate"), and so on. The iconicity constraint is *Ø_C (or STAR ZERO CASE): "A value for the feature CASE must not be absent", and the economy constraint is *STRUC_C (or STAR STRUCTURE CASE): "A value for the morphological category CASE must not be present." Now to get her system to work, Aissen needs to make use of the mechanism of local constraint conjunction, creating constraints such as *OJ/HUM & *Ø_C ("a human object must not lack a value for the feature CASE"). However, as she acknowledges herself (in note 12, p. 447-448), unrestricted constraint conjunction can generate "undesirable" constraints such as *OJ/HUM & *STRUC_C ("a human object must lack a value for the feature CASE"). If such constraints existed, they would wrongly predict the existence of languages with object marking only on nonhuman objects. If they cannot be ruled out, this means that the formal system has no way of explaining the main explanandum, the series of DOM universals ("Object marking is the more likely, the higher the object is on the animacy and definiteness scales"). Aissen is aware of this, and she tries to appeal to functional factors outside the formal system:

"Although constraints formed by conjunction of the subhierarchies with *STRUC_C might exist, grammars in which they were active would be highly dysfunctional since marking would be enforced most strenuously exactly where it is least needed."

But if certain nonoccurring languages are ruled out directly by non-formal functional considerations, then it is unclear why the elaborate system of indirectly functionally derived constraints is needed in the first place. The DOM universals are amenable to a straightforward and complete explanation in functional terms, so adopting a UG-based explanation instead only makes sense if one takes an antifunctionalist stance for independent reasons. Aissen's intermediate approach, with some formalist (though functionally-derived) elements and some direct functionalist components, seems unmotivated.¹⁰

In the next section, we will see what a fully functionalist approach to the explanation of syntactic universals looks like.

4. The functionalist approach to explaining syntactic universals

There are many different authors who use the label "functionalist", so I will concentrate on my own approach here (e.g. Haspelmath 1999a, 2004b, 2005, to appear), which is largely compatible with the approaches of other authors such as Greenberg (1966), Comrie (1981), Givón (1995), Dryer (1992, 2006), Stassen (1997), Croft (2001, 2003), Cristofaro (2003, to appear), Hawkins (2004). I will also discuss other nongenerative approaches to typology here that do not necessarily claim to have explanations for universals.

¹⁰ Aissen also refers to Jäger (2004), noting that his evolutionary approach "may explain more precisely why such grammars are dysfunctional and unstable". But Jäger's functionalist approach entirely dispenses with the formalist elements of Aissen's system and is thus not compatible with her analysis.

4.1. The fundamental difference between the functionalist and the generative approach

The difference between functional and generative linguistics has often been framed as being about the issue of autonomy of syntax or grammar (Croft 1995, Newmeyer 1998), but in my view this is a misunderstanding (see Haspelmath 2000). Instead, the primary difference (at least with respect to the issue of explaining universals) is that generative linguists assume that syntactic universals should all be derived from UG, i.e. that unattested language types are computationally (or biologically) impossible (cf. §2.5), whereas functionalists take seriously the possibility that syntactic universals could also be (and perhaps typically are) due to general ease of processing factors. For functionalists, unattested languages may simply be improbable, but not impossible (see Newmeyer 2005, who in this respect has become a functionalist).

This simple and very abstract difference between the two approaches has dramatic consequences for the practice of research. Whereas generativists put all their energy into finding evidence for the principles (or constraints) of UG, functionalists invest a lot of resources into describing languages in an ecumenical, widely understood descriptive framework (e.g. Givón 1980, Haspelmath 1993), and into gathering broad cross-linguistic evidence for universals (e.g. Haspelmath et al. 2005). Unlike generativists, functionalists do not assume that they will find the same syntactic categories and relations in all languages (cf. Dryer 1997a, Croft 2000b, 2001, Cristofaro to appear), but they expect languages to differ widely and show the most unexpected idiosyncrasies (cf. Dryer 1997b, Plank 2003). They thus agree with Joos's (1957:96) notorious pronouncement that "languages can differ from each other without limit and in unpredictable ways". However, some of these ways are very unlikely, so strong probabilistic predictions are possible after all.

Since for functionalists, universals do not necessarily derive from UG, they do not regard cross-linguistic evidence as good evidence for UG (cf. Haspelmath 2004a:§3), and many are skeptical about the very existence of UG (i.e. the existence of genetically fixed restrictions specifically on grammar). For this reason, they do not construct formal frameworks that are meant to reflect the current hypotheses about UG, and instead they only use widely understood concepts in their descriptions ("basic linguistic theory", Dixon 1997). The descriptions provided by functionalists are not intended to be restrictive and thus explanatory. Description is separated strictly from explanation (Dryer 2006).

4.2. "Deep" implicational universals

Implicational universals, which describe correlations between features, were first highlighted in Greenberg's (1963) work on word order universals and other grammatical correlations. But they were implicit in many earlier writings on holistic "language types", going back to 19th century authors such as Humboldt, Schleicher and Müller. Saying that a language is "agglutinating" means that this trait permeates its entire grammar, so that not only its verb inflection, but also its noun inflection is agglutinating (Plank 2000). Languages that show agglutinating verb inflection but flecnal noun inflection are unexpected on this view, and implicitly a bidirectional implicational universal is posited ("If a language has agglutinating verb

inflection, it will (tend to) have agglutinating noun inflection, and vice versa").

Until Greenberg started working with systematic language sampling, hypotheses about typological correlations were typically formulated implicitly in terms of idealized holistic language types (e.g. Finck 1909, Sapir 1921, Skalička 1966), and sometimes fairly strong hypotheses about correlations were made. The idea was that languages are systems *où tout se tient* ("where everything hangs together", Meillet 1903:407), so that connections between quite different parts of the grammar can be discovered by linguists. This view is expressed in very poetic terms by Georg von der Gabelentz (1901:481), in the passage that contains the first mention of the term *typology*:¹¹

"Aber welcher Gewinn wäre es auch, wenn wir einer Sprache auf den Kopf zusagen dürften: Du hast das und das Einzelmerkmal, folglich hast du die und die weiteren Eigenschaften und den und den Gesamtcharakter! - wenn wir, wie es kühne Botaniker wohl versucht haben, aus dem Lindenblatte den Lindenbaum konstruieren könnten. Dürfte man ein ungeborenes Kind taufen, ich würde den Namen Typologie wählen."

Plank (1998) documents a large number of mostly forgotten early attempts to link phonological properties of languages with nonphonological (morphological and syntactic) properties. Klimov (1977, 1983) (summarized in Nichols 1992:8-12) proposed a typology that links a large number of diverse lexical, syntactic and morphological properties. And while Greenberg (1963) was conservative in the kinds of implications that he proposed, others following him made bolder claims on correlations involving word order patterns. Thus, Lehmann (1973, 1978) links the VO/OV distinction to phonological properties. For example, he claims that "the effect of phonological processes is "progressive" in OV languages, "anticipatory" in VO. OV languages tend to have vowel harmony and progressive assimilation; VO languages tend to have umlaut and anticipatory assimilation" (Lehmann 1978:23).

Thus, quite analogously with the generative notion of "deep" parameters (or "macroparameters"), the nongenerative literature is full of claims of "deep" implicational universals. While the nongenerative literature does not make a connection to ease of language acquisition, it seems that the idea that there are just a few basic blueprints of human languages, and that linguists could discover the key to all the tight interconnections between structural features, is attractive and hard to resist independently of the Chomskyan way of thinking about language.

4.3. Assessment of success; abandoning deep implications

The deep implications mentioned in the preceding subsection have not been more successful than the deep parameters discussed in §2. Most claims of holistic types from the 19th century and the pre-Greenbergian 20th century have not been substantiated and have fallen into oblivion. In many cases these

¹¹ "But what an achievement would it be were we to be able to confront a language and say to it: 'you have such and such a specific property and hence also such further properties and such and such an overall character'—were we able, as daring botanists must have tried, to construct the entire lime tree from its leaf. If an unborn child could be baptized, I would choose the name typology."

holistic types were set up on the basis of just a few languages and no serious attempt was made to justify them by means of systematic sampling. When systematic cross-linguistic research became more common after Greenberg (1963) (also as a result of the increasing availability of good descriptions of languages from around the world), none of the holistic types were found to be supported. For example, the well-known morphological types (agglutinating/flectional) were put to a test in Haspelmath (1999b), with negative results. It is of course possible that we have simply not looked hard enough. For example, I am not aware of a cross-linguistic study of vowel harmony and umlaut whose results could be correlated with word order properties (VO/OV) of the languages.

But there does seem to be a widespread sense in the field of (nongenerative) typology that cross-domain correlations do not exist and should not really be expected. After the initial success of word order typology, there have been many attempts to link word order (especially VO/OV) to other aspects of language structure, such as comparative constructions (Stassen 1985), alignment types (Siewierska 1996), indefinite pronoun types (Haspelmath 1997), and (a)symmetric negation patterns (Miestamo 2005). But such attempts have either failed completely or have produced only weak correlations that are hard to distinguish from areal effects. Nichols's (1992) large-scale study is particularly telling in this regard: While she starts out with Klimov's hypotheses about correlating lexical and morphosyntactic properties, she ends up finding geographical and historical patterns instead, rather than particularly interesting correlations. The geographical patterning of typological properties was also emphasized by works such as Dryer (1989), Dahl (1996), and Haspelmath (2001), and the publication of *The World Atlas of Language Structures* (Haspelmath et al. 2005) documents the shift of interest to geographical patterns. According to a recent assessment by Bickel (2005), typology has turned into

"a full-fledged discipline, with its own research agenda, its own theories, its own problems. The core quest is no longer the same as that of generative grammar, the core interest is no longer in defining the absolute limits of human language. What has taken the place of this is a fresh appreciation of linguistic diversity in its own right... Instead of asking "what's possible?", more and more typologists ask "what's where why?."

In Bickel's formulation, this shift is entirely positive. The rhetoric is reminiscent of what one hears from Minimalist critics of macroparameters. Pica (2001:vi), quoted in §2.4, continues: "This move [i.e. the Minimalist abandonment of macroparameters] allows a radical simplification of the nature and design of UG..." It is not easy to admit defeat, but I believe that in both cases the primary motivation for the shift of interests was the realization that the earlier goals were not reachable. It is illusory to think that linguistic diversity can be captured by a few holistic types, or a few word-order types, or a few parametric switches. Languages are not neat systems "où tout se tient". Like biological organisms, they have a highly diverse range of mutually independent properties, many of which are due to idiosyncratic factors of history and geography.

However, as we will see in the next subsection, this does not mean that there are no typological implications at all.

4.4. Intra-domain implications

In generative grammar, the distinction between "deep" (or macro-) parameters and more shallow (or micro-) parameters is now widely known, but there is no widespread corresponding nongenerative terminology for typological correlations or implications. In the preceding two subsections, I used the terms "deep implications" and "holistic types". Only the latter has antecedents in the earlier literature (e.g. Comrie 1990). I now want to introduce a new term pair: **cross-domain** implications and **intra-domain** implications. The term "domain" is somewhat vague, but I mean it in a narrower sense than "level" (phonology, morphology, syntax), approaching the sense of "construction". There is obviously a continuum from the most extreme cross-domain implications (e.g. semantic organization of kinship terms correlating with the allophonic realization of trill consonants) to the most extreme intra-domain implications (e.g. a velar voiced plosive implying a labial voiced plosive, or trial number in personal pronouns implying dual number). Since my claims about the nature of these two types of implications are not very precise and are not based on systematic evidence, there is no need to define the two types in a more precise fashion.

My basic claim is this: Since languages are not systems *où tout se tient*, we hardly find cross-domain implications. However, a large number of intra-domain implications have been found, and they are usually amenable to functional explanation. Let me give a few examples of such implications.

- (1) If a language lacks overt coding for transitive arguments, it will also lack overt coding for the intransitive subject (Greenberg 1963, Universal 38).
- (2) Grammatical Relations Scale: **subject > object > oblique > possessor**
If a language can relativize on a position on the Grammatical Relations Scale, it can also relativize on all higher positions (Keenan & Comrie 1977).
- (3) Alienability Scale: **body-part terms > personal sphere terms > others**
If a language allows juxtaposition to express a possessive relation with a possessed item from one of the positions on the Alienability Scale, it also allows juxtaposition with all higher positions (Nichols 1988)
- (4) Animacy Scale: **human > nonhuman animate > inanimate**
Definiteness Scale: **definite > indefinite specific > nonspecific**
If a language has overt case marking for an object on a position on one of these scales, it also has overt object case marking for all higher positions (Silverstein 1976, Comrie 1981, Bossong 1985, Aissen 2003).
- (5) Spontaneity Scale:
transitive > unergative > unaccusative costly > unaccusative automatic
If a language has overt causative marking for noncausative bases in one position of the scale, it also has overt causative marking for all higher positions (Haspelmath 2006).
- (6) Ditransitive Person-Role Scale:
Recipient 1st/2nd + Theme 3rd > Recipient 1st/2nd + Theme 1st/2nd or Recipient 3rd + Theme 3rd > Recipient 3rd + Theme 1st/2nd

If a language has bound-pronoun combinations with one position on this scale, it also has bound-pronoun combinations with all higher positions (Haspelmath 2004b).

- (7) If a language uses a special reflexive pronoun for an adnominal possessor that is coreferential with the subject, then it also uses a special reflexive pronoun for the object.

These implicational universals are all restricted to a single constructional domain: relative clauses, adnominal possessive constructions, object case-marking, causative verb formation, ditransitive pronominal objects, reflexive marking. They typically refer to prominence scales, often with more than two positions. Such universals are very different from the universals discussed and explained by parametric approaches, which cannot easily accommodate scales, regardless of whether they take a macro or a micro perspective.

Scales have been adopted into generative approaches only by Optimality Theory, starting with Prince & Smolensky's 1993 [2004] device of "harmonic alignment" (originally introduced to capture sonority scale effects in syllable structure). This formal device was prominently applied to syntax by Aissen (1999, 2003) and related work. In Aissen's papers, prominence scales are taken over from the functionalist and other nongenerative typological literature and are used to create a series of constraints with fixed ranking that can describe all and only the attested systems. However, as I pointed out in §3, OT cannot say why the constraints are the way they are, and Aissen does not succeed in explaining how prominence scales such as those in (1)-(7) come to play a role in the makeup of constraints. One gets the impression that Aissen creates the constraints because they work, and they work because they are derived from the valid intra-domain implications that we have seen. Aissen's approach does not bring us closer to understanding why the implicational universals should hold.

4.5. Functional explanations of intra-domain implications

In my view, the most interesting and convincing functional explanations are those that derive universal implications in language structure from scales of processing difficulty, and I will limit myself to such explanations here. In phonology, such functional explanations are straightforward and no longer controversial: Labial voiced plosives are easier to produce than velar voiced plosives because there is more space between the larynx and the lips for airflow during the oral closure. Thus, it is not surprising that in language structure, too, there is a preference for labial voiced plosives, and many languages lack velar voiced plosives (Maddieson 2005).

In morphology and syntax, processing difficulty play a role in various ways. Perhaps the most straightforward way is the Zipfian economy effect of frequency of use: More frequent items tend to be shorter than rarer items, and if one member of an opposition is zero, processing ease dictates that it should be the more frequent one. Such a simple Zipfian explanation applies in the case of (1), (3), (4) and (5) above. The case of the intransitive subject is more frequent than the case of the transitive subject or object,¹² so it tends to be

¹² In many texts, the number of intransitive and transitive clauses is about equal. But the case of the intransitive subject (nominative or absolutive) is virtually always also used for one of

zero-coded (Comrie 1978). Inalienable nouns such as body-part terms and kinship terms tend to occur with a possessor, so languages often restrict possessive marking to other possessed nouns, which occur with a possessor much more rarely. Objects tend to be inanimate and indefinite, so languages tend to restrict overt case marking to the rarer cases, definite and human objects.

Another general effect of frequency that is ultimate due to processing constraints is compact expression as a bound form (see Bybee 2003). This explains (6) above: The more frequent person-role combinations (especially Recipient 1st/2nd + Theme 3rd) tend to occur as bound forms, whereas the rarer combinations often do not (see Haspelmath 2004b for frequency figures).

Rarity of use may also lead to the need for a special form to help the hearer in the interpretation. Thus, the object of a transitive verb is relatively rarely coreferential with the subject, so many languages have a special reflexive pronoun signaling non-coreference. But an adnominal possessor is far more often coreferential with the subject, so there is less of a need for a special form, and many languages (such as English) that have a special object reflexive pronoun lack a special possessive reflexive pronoun (see Haspelmath to appear).

The universal in (2) is explained in processing terms in Hawkins (2004:ch. 7), who invokes a general domain minimization principle for filler-gap dependencies.

In comparing these functional explanations of syntactic universals to generative explanations, several points can be made:

- (i) The functional explanations have nothing to do with the absence of an autonomy assumption. Syntax and grammar are conceived of as autonomous, and a strict competence-performance distinction is made. But there is no assumption that competence grammars are totally independent and isolated: Grammars can be influenced by performance.
- (ii) The functional explanations do not contradict the idea that there is a Universal Grammar, even though they do not appeal to it. Functional explanation and Universal Grammar are largely irrelevant to each other.
- (iii) The functional explanations do not presuppose a cognitively real (i.e. descriptively adequate) description of languages. Phenomenological (i.e. observationally adequate) descriptions of languages are sufficient to formulate and test the universals in (1)-(7) (Haspelmath 2004a).
- (iv) The explanations do not involve the notion of "restrictiveness" of the descriptive framework, appealing instead to factors external to the description and the framework (cf. Newmeyer 1998:... on extrnal vs. internal explanations).
- (v) The functional explanations have no implications for language acquisition, but otherwise they are more easily tested than the generative explanations.

While functional explanations avoid the assumption of a hypertrophic Universal Grammar that contains a vast number of innate principles and parameters or OT constraints, they do presuppose a mechanism for incorporating processing preferences into grammars. It is generally

the transitive arguments, so it is always more frequent than the case used for the other transitive argument.

recognized that this mechanism lies in diachronic change (Bybee 1988, Keller 1994, Kirby 1999, Nettle 1999, Haspelmath 1999a, Croft 2000a), even though the precise way in which processing preferences become grammatical rules is perhaps not totally clear yet. Here is not the place for an in-depth discussion of this question, but the basic idea is that novel structures always arise through language change, and language change is influenced by language use. When novel variants are created unconsciously in language use, the easily processable variants are preferred (Croft 2000a); when innovations spread, the easily processable structures are preferred (Haspelmath 1999a); and when language is acquired, the easily processable structures are preferred (Kirby 1999). Not all of these hypotheses may be needed, and telling which ones are the right ones is not easy, but there is a consensus among functionalists that language change is the key mediating mechanism that allows performance factors to shape grammars.

Thus, functional explanations in linguistics can be seen as parallel to functional evolutionary explanations in biology (Haspelmath 1999a, 2004a:§2): Just as the diachronic evolutionary Darwinian theory of variation and selection did not presuppose detailed knowledge of the genetic basis for variation in organisms, functionalism in linguistics does not presuppose cognitively real descriptions. Phenomenological universals can be explained functionally.

5. Summary

Explaining syntactic universals is a hard task on which there is very little agreement among comparative syntacticians. I have reviewed two prominent approaches to this task here, the generative parametric approach and the functionalist approach in the Greenbergian tradition (as well as the generative Optimality Theory approach, which diverges in interesting ways from the parametric approach). The disagreements between them start with the kinds of descriptions that form the basis for syntactic universals: Generativists generally say that cognitively real descriptions are required, whereas functionalists can use any kind of observationally adequate description. Generativists work with the implicit assumption that all universals will find their explanation in Universal Grammar, whereas functionalists do not appeal to UG and do not even presuppose that languages have some of the same categories and structures. Functionalists attempt to derive general properties of language from processing difficulty, whereas generativists see no role for performance in explaining competence.

But I have identified one common idea in parametric generative approaches and nongenerative approaches to syntactic universals:¹³ the hope that the bewildering diversity of observed languages can be reduced to very few fundamental factors. These are macroparameters in generative linguistics, and holistic types in nongenerative linguistics. The evidence for both of these constructs was never overwhelming, and in both approaches, not only external critics, but also insiders have increasingly pointed to the discrepancy between the ambitious goals and the actual results. Generativists now tend to focus on microparameters (if they are interested in comparative grammar at

¹³ Interestingly, this idea seems to be much less present in Optimality Theory.

all), and nongenerative typologists now often focus on geographical and historical particularities rather than world-wide universals.

However, I have argued that there is one type of implicational universal that is alive and well: intra-domain implications that reflect the relative processing difficulty of different types of elements in closely related constructions. Such implicational universals cannot be explained in terms of parameters, and the only generative explanation is an OT explanation in terms of harmonic alignment of prominence scales. However, this is a crypto-functional explanation, and unless one is a priori committed to the standard generative assumptions, there is no reason to prefer it to the real functional explanation.

References

- Adger, David. 2003. *Core syntax: a minimalist approach*. Oxford: Oxford University Press.
- Aissen, Judith. 1999. "Markedness and subject choice in Optimality Theory." *Natural Language and Linguistic Theory* 17: 673-711.
- Aissen, Judith. 2003. "Differential object marking: Iconicity vs. economy." *Natural Language and Linguistic Theory* 21.3: 435-483.
- Atkinson, Martin. 1994. "Parameters." In: Asher, Ronald E. (ed.) *The encyclopedia of language and linguistics*, vol. 6. Oxford: Pergamon, 2941-2942.
- Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Baker, Mark C. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.
- Baker, Mark C. 2001a. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.
- Baker, Mark C. 2001b. "Configurationality and polysynthesis." In: In: Martin Haspelmath & Ekkehard König & Wulf Oesterreicher & Wolfgang Raible (eds.) *Language typology and language universals: An international handbook*. (Handbücher zur Sprach- und Kommunikationswissenschaft) Vol. 2. Berlin: de Gruyter, 1433-1441.
- Baker, Mark C. 2003. *Lexical categories: verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.
- Baker, Mark C. 2006. "The macroparameter in a microparametric world." To appear in: Theresa Biberauer & Anders Holmberg (eds.) *The structure of parametric variation*.
- Baker, Mark C. and Jim McCloskey. 2005. "On the Relationship of Typology to Theoretical Syntax." Paper presented at the LSA Symposium on Typology in the U.S. Downloadable at <http://www.rci.rutgers.edu/mabaker/papers.html>.
- Baltin, Mark. 2004. "Remarks on the relation between language typology and Universal Grammar: Commentary on Newmeyer." *Studies in Language* 28.3: 549-553.
- Bickel, Balthasar. 2005. "Typology in the 21st century: major current developments." Paper presented at the 2005 LSA Meeting (Symposium on typology in the United States)
- Bilola, Edmund. 1998. *Syntaxe générative: la théorie des principes et des paramètres*. Munich: Lincom.
- Boersma, Paul. 1998. *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Bossong, Georg. 1985. *Differenzielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Bossong, Georg. 1991. "Differential Object Marking in Romance and Beyond." In D. Wanner and D. Kibbee (eds.), *New Analyses in Romance Linguistics: Selected Papers from the XVIII Linguistic Symposium on Romance Languages, Urbana-Champaign, April 7-9, 1988*. Amsterdam: Benjamins, 143-170.
- Bossong, Georg. 1998. "Le marquage différentiel de l'objet dans les langues d'Europe". In: Feuillet, Jack (ed.) *Actance et valence dans les langues de l'Europe*. Berlin: Mouton de Gruyter, 193-258.
- Bresnan, Joan (1997): "The emergence of the unmarked pronoun: Chichewa pronominals in Optimality Theory". *Berkeley Linguistics Society* 23
- Bybee, Joan L. 1988. "The diachronic dimension in explanation." In: John A. Hawkins (ed.) *Explaining language universals*. Oxford: Blackwell, 350-379.

- Bybee, Joan L. 2003. "Mechanisms of change in grammaticization: The role of repetition." In Janda, Richard & Joseph, Brian (eds.) *Handbook of historical linguistics*. Oxford: Blackwell.
- Carnie, Andrew. 2002. *Syntax: a generative introduction*. Malden, MA: Blackwell.
- Chomsky, Noam & Lasnik, Howard. 1993. "The theory of Principles and Parameters". In: Jacobs, Joachim et al. (eds.) *Syntax: An international handbook of contemporary research*, vol. 1. Berlin: de Gruyter, 506-569.
- Chomsky, Noam A. 1991. "Some notes on economy of derivation and representation." In: Freidin, Robert (ed.) *Principles and parameters in comparative grammar*. Cambridge/MA: MIT Press, 417-454.
- Chomsky, Noam A. 1988. *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.
- Chomsky, Noam A. 2000. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
- Chomsky, Noam A. 2004. "Beyond explanatory adequacy." In: *Structures and beyond: the cartography of syntactic structures*, ed. A. Belletti. Oxford: Oxford University Press.
- Cinque, Guglielmo & Richard Kayne (eds.) 2005. *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Cinque, Guglielmo. 2005. "Deriving Greenberg's Universal 20 and Its Exceptions". *Linguistic Inquiry* 35.3
- Comrie, Bernard. 1978. "Ergativity." In: Lehmann, Winfred P. (ed.) *Syntactic typology*. Austin: University of Texas Press, 329-94.
- Comrie, Bernard. 1981. *Language universals and linguistic typology*. Oxford: Blackwell.
- Comrie, Bernard. 1990. "Holistic versus partial typologies". In Werner Bahner, Joachim Schildt, and Dieter Viehweger, eds.: *Proceedings of the Fourteenth International Congress of Linguists*, volume 1, 139-148. Berlin: Akademie-Verlag.
- Coopmans, Peter. 1983. "Review of *Language universals and linguistic typology* by Bernard Comrie." *Journal of Linguistics* 19: 455-474.
- Cristofaro, Sonia. To appear. "Language universals and linguistic knowledge." To appear in Jae Jung Song (ed.) *Handbook of Typology*.
- Croft, William. 1995. "Autonomy and functionalist linguistics." *Language* 71: 490-532.
- Croft, William. 2000a. *Explaining language change: an evolutionary approach*. London: Pearson.
- Croft, William. 2000b. "Parts of speech as language universals and as language-particular categories", in: Vogel, Petra & Comrie, Bernard (eds.) *Approaches to the typology of word classes*. Berlin: Mouton de Gruyter, 65-102.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.
- Croft, William. 2003. *Typology and universals*. 2nd edition. Cambridge: Cambridge University Press.
- Croft, William & Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Culicover, Peter W. 1997. *Principles and parameters: an introduction to syntactic theory*. Oxford: Oxford University Press.
- Dahl, Östen. 1996. "Das Tempussystem des Deutschen im typologischen Vergleich", in: Lang, Ewald & Zifonun, Gisela (eds.), *Deutsch – typologisch*, Berlin: de Gruyter, 359-368.
- Dixon, R.M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dryer, Matthew. 1989. "Large linguistic areas and language sampling." *Studies in Language* 13:257-92.
- Dryer, Matthew. 1992. "The Greenbergian word order correlations." *Language* 68: 81-138.
- Dryer, Matthew. 1997. "Why statistical universals are better than absolute universals." *Chicago Linguistic Society* 33: 123-145.
- Dryer, Matthew. 1997a. "Are grammatical relations universal?" In: Bybee, Joan & Haiman, John & Thompson, Sandra A. (eds.) *Essays on language function and language type*. Amsterdam: Benjamins, 115-143.
- Dryer, Matthew. 1997b. "Why statistical universals are better than absolute universals." *Chicago Linguistic Society* 33: *The Panels*, 123-145.
- Dryer, Matthew. 2006. "Descriptive theories, explanatory theories, and basic linguistic theory." In: Felix K. Ameka & Alan Dench & Nicholas Evans (eds.) *Catching Language: The Standing Challenge of Grammar Writing*. Berlin: Mouton de Gruyter.
- Enrico, John. 2003. *Haida syntax*. 2 vols. Lincoln: University of Nebraska Press.
- Fanselow, Gisbert & Felix, Sascha. 1987. *Sprachtheorie*, vol. 1. Tübingen: Francke.

- Filimonova, Elena. 2005. "The noun phrase hierarchy and relational marking: problems and counterevidence." *Linguistic Typology* 9.1: 77-113.
- Finck, Franz Nikolaus 1909, *Die Haupttypen des Sprachbaus*. Leipzig: B.G. Teubner [Reprinted: Darmstadt, Wissenschaftliche Buchgesellschaft, 1965]
- Finck, Franz Nikolaus. 1909. *Die Haupttypen des Sprachbaus*. Leipzig.
- Fodor, Janet Dean. 2003. "Setting syntactic parameters." In: Baltin, Mark & Chris Collins (eds.) *The handbook of contemporary syntactic theory*. Oxford: Blackwell, 730-767.
- Freeze, Ray. 1992. Existentials and other locatives. *Language* 68(3):553-595.
- Freidin, Robert. 1992. *Foundations of generative syntax*. Cambridge, MA: MIT Press.
- Fukui, Naoki. 1995. "The principles-and-parameters approach: a comparative syntax of English and Japanese." In: Shibatani, Masayoshi & Bynon, Theodora (eds.) *Approaches to language typology*. Oxford: Clarendon Press, 327-372.
- Fukui, Naoki & Mihoko Zushi. 2004. "Introduction." In: Noam Chomsky. 2004. *The generative enterprise revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui and Mihoko Zushi*. Berlin: Mouton de Gruyter, 1-25.
- Gabelentz, Georg von der (1901): *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel.
- Gilligan, Gary M. 1987. *A cross-linguistic approach to the pro-drop parameter*. Ph.D. dissertation, University of Southern California.
- Givón, T. (1995): *Functionalism and grammar*. Amsterdam: Benjamins.
- Givón, Talmy. 1980. *Ute Reference Grammar*. Ignacio, Colorado: The Ute Press.
- Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements." In: Greenberg, Joseph H. (eds.) *Universals of grammar*, 73-113. Cambridge, Mass.: MIT Press.
- Greenberg, Joseph. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.
- Grewendorf, Günther. 2002. *Minimalistische Syntax*. Tübingen: Francke.
- Grimshaw, Jane & Samek-Lodovici, Vieri (1998): "Optimal subjects and subject universals." In: Barbosa, Pilar et al. (eds.) *Is the best good enough?* Cambridge/MA: MITWPL & MIT Press.
- Haegeman, Liliane. 1997. "Introduction: on the interaction of theory and description in syntax." In: Haegeman, Liliane (ed.) *The new comparative syntax*. London: Longman, 1-32.
- Haider, Hubert. 1994. "(Un-)heimliche Subjekte -- Anmerkungen zur Pro-drop Causa, im Anschluß an die Lektüre von Osvaldo Jaeggli & Kenneth J. Safir, eds., *The Null Subject Parameter*." *Linguistische Berichte* 153: 372-85.
- Haider, Hubert. 2001. "Parametrisierung in der Generativen Grammatik." In: Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.) *Language typology and language universals*. (Handbücher zur Sprach- und Kommunikationswissenschaft) Berlin: de Gruyter, 283-293.
- Haiman, John. 1985. *Natural syntax*. Cambridge: Cambridge University Press.
- Hale, Kenneth L. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1.1:5-47.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. (Mouton Grammar Library, 9.) Berlin: Mouton de Gruyter
- Haspelmath, Martin. 1997. *Indefinite pronouns*. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press
- Haspelmath, Martin. 1999a. "Optimality and diachronic adaptation." *Zeitschrift für Sprachwissenschaft* 18.2: 180-205.
- Haspelmath, Martin. 1999b. "The Agglutination Hypothesis: a belated empirical investigation." Paper presented at the ALT Conference, Amsterdam.
- Haspelmath, Martin. 2000. "Why can't we talk to each other? A review article of [Newmeyer, Frederick. 1998. *Language form and language function*. Cambridge: MIT Press]." *Lingua* 110.4: 235-255.
- Haspelmath, Martin. 2001. "The European linguistic area: Standard Average European." In: Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (eds.) *Language typology and language universals*. (Handbücher zur Sprach- und Kommunikationswissenschaft) Berlin: de Gruyter, 1492-1510.
- Haspelmath, Martin. 2004a. "Does linguistic explanation presuppose linguistic description?" *Studies in Language* 28.3: 554-579 (special issue guest edited by Martina Penke and Anette Rosenbach)
- Haspelmath, Martin. 2004b. "Explaining the Ditransitive Person-Role Constraint: a usage-based account." *Constructions* 2004/2, 49 pp. (free online journal, www.constructions-online.de)

- Haspelmath, Martin. 2005. "Argument marking in ditransitive alignment types." *Linguistic Discovery* 3.1:1-21 (<http://linguistic-discovery.dartmouth.edu>)
- Haspelmath, Martin. 2006. "Universals of causative verb formation." Paper presented at the Symposium Current Topics in Typology, Freie Universität Berlin, January 2006.
- Haspelmath, Martin. To appear. "A frequentist explanation of some universals of reflexive marking".
- Haspelmath, Martin & Matthew S. Dryer & David Gil & Bernard Comrie (eds.) *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hayes, Bruce & Kirchner, Robert & Steriade, Donca (eds.) 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press
- Jackendoff, Ray. 1977. *X bar syntax: a study of phrase structure*. Cambridge: MIT Press.
- Jaeggli, Osvaldo & Kenneth L. Safir (eds.) 1989. The null subject parameter. Dordrecht: Foris.
- Jäger, Gerhard. 2004. "Learning constraint sub-hierarchies: The Bidirectional Gradual Learning Algorithm." In R. Blutner & H. Zeevat (eds.), *Pragmatics in OT*. Palgrave MacMillan, 251-287.
- Jelinek, Eloise. 1984. "Empty Categories, Case, and Configurationality." *Natural Language and Linguistic Theory* 2: 39-76.
- Jenkins, Lyle. 2000. *Biolinguistics: Exploring the biology of language*. Cambridge: Cambridge University Press.
- Johannessen, Janne Bondi. 1998. *Coordination*. New York: Oxford University Press.
- Joos, Martin (ed.) 1957, *Readings in linguistics: The development of descriptive linguistics in America since 1925*. New York: American Council of Learned Societies.
- Julien, Marit. 2002. *Syntactic heads and word-formation*. New York: Oxford University Press.
- Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.
- Kayne, Richard. 1984. *Connectedness and binary branching*. Dordrecht: Foris.
- Kayne, Richard. 1996. "Microparametric syntax: some introductory remarks." In: James R. Black & Virginia Montapanyane (eds.) *Microparametric syntax and dialect variation*. John Benjamins, Amsterdam, ix-xvii. (Reprinted in Kayne 2000:3-9)
- Kayne, Richard. 2000. *Parameters and universals*. New York: Oxford University Press.
- Keenan, Edward L. & Comrie, Bernard. 1977. "Noun phrase accessibility and universal grammar." *Linguistic Inquiry* 8: 63-99.
- Keller, Rudi. 1994. *Language change: The invisible hand in language*. London: Routledge.
- Kiparsky, Paul. 2004. "Universals constrain change, change results in typological generalizations." Ms., Stanford University.
- Kirby, Simon (1997): "Competing motivation and emergence: Explaining implicational hierarchies". *Linguistic Typology* 1.1: 5-5-31.
- Kirby, Simon (1999): *Function, selection and innateness: The emergence of language universals*. Oxford: Oxford University Press.
- Kirby, Simon. 1999. *Function, selection, and innateness: The emergence of language universals*. Oxford: Oxford University Press.
- Klimov, Georgij A. 1977. *Tipologija jazykov aktivnogo stroja*. Moskva: Nauka.
- Klimov, Georgij A. 1983. *Principy kontensivnoj tipologii*. Moskva: Nauka.
- Kuroda, S.-Y. 1988. "Whether We Agree or Not: A Comparative Syntax and English and Japanese." *Linguisticae Investigationes* 12: 1-47.
- Legendre, Géraldine & Grimshaw, Jane & Sten Vikner (eds.) 2001. *Optimality-theoretic syntax*. Cambridge, MA: MIT Press.
- Legendre, Géraldine & William Raymond & Paul Smolensky. 1993. "An Optimality-Theoretic typology of case and grammatical voice systems." *Berkeley Linguistics Society* 19: 464-478.
- Legendre, Géraldine. 2001. "An introduction to Optimality Theory in syntax." In: Legendre, Géraldine, Grimshaw, Jane and Sten Vikner (eds.) 2001. *Optimality-theoretic syntax*. Cambridge, MA: MIT Press, 1-27.
- Lehmann, Winfred P. 1973. "A Structural Principle of Language and its Implications." *Language* 49: 47-66.
- Lehmann, Winfred P. 1978. "The great underlying ground-plans." In: Lehmann, Winfred P. (ed.) 1978. *Syntactic Typology: Studies in the Phenomenology of Language*. Austin: University of Texas Press, 3-55.
- Lightfoot, David. 1999. *The development of language: Acquisition, change, and evolution*. (Maryland Lectures in Language and Cognition, 1.) Malden, MA: Blackwell.

- Maddieson, Ian. 2005. "Voicing and gaps in plosive systems." In: Haspelmath, Martin & Matthew S. Dryer & David Gil & Bernard Comrie (eds.) *The World Atlas of Language Structures*. Oxford: Oxford University Press, 26-29.
- Malchukov, Andrej. 2005. "Case Pattern Splits, Verb Types and Construction Competition." In: M. Amberber & H. de Hoop (eds.) *Competition and Variation in Natural Languages: The Case for Case*. Amsterdam: Elsevier, 73-118.
- Marácz, László & Muysken, Pieter (eds.) 1989. *Configurationality: the typology of asymmetries*. Dordrecht: Foris.
- McCarthy, John J. 2002. *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.
- McCloskey, James. 1988. "Syntactic theory." In: Newmeyer, Frederick (ed.) *Linguistics: The Cambridge Survey*, vol. 1 (*Linguistic theory: Foundations*), 18-59.
- Meillet, Antoine. 1903. *Introduction à l'étude comparative des langues indo-européennes*. Paris: Hachette.
- Miestamo, Matti. 2005. *Standard negation: the negation of declarative verbal main clauses from a typological perspective*. Berlin: Mouton de Gruyter.
- Müller, Gereon. 2000. *Elemente der optimalitätstheoretischen Syntax*. Tübingen: Stauffenburg.
- Nakamura, Wataru. 1999. "Functional Optimality Theory: Evidence from split case systems." In: Darnell, Michael et al. (eds.) *Functionalism and formalism in linguistics*. Vol. II, 253-276. Amsterdam: Benjamins.
- Neeleman, Ad & Kriszta Szendrői. 2005. "Radical pro-drop and the morphology of pronouns." Paper presented at the conference on the Structure of Parametric Variation, University of Newcastle, September 2005.
- Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.
- Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2004 "Against a parameter-setting approach to language variation." *Linguistic Variation Yearbook* 4: 181-234.
- Newmeyer, Frederick. 1998. *Language form and language function*. Cambridge: MIT Press.
- Newmeyer, Frederick. 2005. *Possible and probable languages: a generative perspective on linguistic typology*. Oxford: Oxford University Press.
- Nichols, Johanna, 1988. "On alienable and inalienable possession." In: Shipley, William (ed.), *In honor of Mary Haas*, pp. 557-609. Berlin: Mouton de Gruyter.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: The University of Chicago Press.
- Ouhalla, Jamal. 1991. *Functional Categories and Parametric Variation*. London: Routledge.
- Ouhalla, Jamal. 1999. *Introducing Transformational Grammar: From Principles and Parameters to Minimalism*. London: Hodder Arnold.
- Pica, Pierre. 2001. "Introduction." *Linguistic Variation Yearbook* 1: v-xii.
- Pinker, Steven. 1994. *The language instinct*. New York: HarperPerennial.
- Plank, Frans. 1998. "The co-variation of phonology with morphology and syntax: A hopeful history." *Linguistic Typology* 2.2: 195-230.
- Plank, Frans. 1999. "Split morphology: How agglutination and flexion mix." *Linguistic Typology* 3.3: 279-340.
- Plank, Frans. 2003ff. "Das grammatische Raritätenkabinett: a leisurely collection to entertain and instruct." (Online collection of rare features in the world's languages: <http://ling.uni-konstanz.de:591/universals/intro.html>)
- Prince, Alan & Smolensky, Paul. 2004. *Optimality theory: constraint interaction in generative grammar*. Oxford: Blackwell. (Originally circulated in 1993)
- Radford, Andrew. 2004. *Minimalist syntax*. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Roberts, Ian & Anders Holmberg. 2005. "On the role of parameters in Universal Grammar: a reply to Newmeyer." In: Hans Broekhuis, N. Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster (eds.) *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*. Berlin: Mouton de Gruyter, 538-553.
- Roberts, Ian. 1997. *Comparative syntax*. London: Arnold.
- Roberts, Ian. 2003. "Head movement." In: Baltin, Mark & Chris Collins (eds.) *The handbook of contemporary syntactic theory*. Oxford: Blackwell, 113-147.
- Sells, Peter. (ed.) 2001. *Formal and empirical issues on Optimality Theoretic syntax*. Stanford: CSLI Publications.
- Sells, Peter. 2001. "Introduction." In: Sells, Peter (ed.) *Formal and empirical issues on Optimality Theoretic syntax*. Stanford: CSLI Publications, 1-16.

- Sharma, Devyani. 2001. "Kashmiri case clitics and person hierarchy effects." In: Sells, Peter (ed.) *Formal and empirical issues on Optimality Theoretic syntax*. Stanford: CSLI Publications, 225-256.
- Siewierska, Anna. 1996. "Word order type and alignment type." *Sprachtypologie und Universalienforschung* 49.2: 149-176.
- Skalička, Vladimír. 1966. "Ein typologisches Konstrukt". *Travaux du Cercle Linguistique de Prague* 2: 14–28. (Reprinted in: *Typologische Studien*. ed. by Peter Hartmann. Braunschweig: Vieweg, 335–341)
- Smith, Neil. 1989. *The twitter machine: reflections on language*. Oxford: Blackwell.
- Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford: Blackwell
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Oxford University Press.
- Taylor, John R. 2002. *Cognitive Grammar*. Oxford: Oxford University Press.
- Travis, Lisa. 1989. "Parameters of Phrase Structure." in M. R. Baltin and A. S. Kroch (eds.) *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press, 263-279.
- Vikner, Sten. 1997. "V °-to-I ° movement and inflection for person in all tenses." In: Haegeman, Liliane (ed.) *The new comparative syntax*. London: Longman, 189-213.