

Formal Generative Typology

Mark C. Baker
Rutgers University

Abstract: This article lays out an approach that combines a formal-generative perspective on language, including tolerance of abstract analyses, with a typological focus on comparing unrelated languages from around the world. It argues that this can be a powerful combination for discovering linguistic universals and patterns in linguistic variation that are not detected by other means.

Key words: typology, universals, Universal Grammar, generative linguistics, syntax, object, subject, verb

1. Introduction

The approach to linguistic research that I describe here can be called “Formal Generative Typology” (FGT). It does not have an extensive body of official doctrine, its own entrenched name, distinctive modes of representation, or a special choice of data. It is a relatively opportunistic approach—ready to borrow insights from other linguistic approaches in the course of pursuing its goals.

Its distinctive content comes primarily from its goals and methods. FGT is firmly rooted in the Chomskian tradition of linguistics, as that has developed over the past 50 years (Chomsky 1957, Chomsky 1965). It grows most directly out of the “Principles and Parameters” stage of that tradition, which crystallized in the 1980s when Chomsky and his followers started to turn their attention more seriously to issues of crosslinguistic comparison (Chomsky 1981, Chomsky and Lasnik 1993). It is in this sense that the approach is “formal” and “generative”. Yet as this line of research has been pursued in its own terms, it has come to a degree of convergence of interest and method with linguistic typology as practiced by linguists like Joseph Greenberg (Greenberg 1963), Bernard Comrie (Comrie 1981), and others. That is why it is “typology”.

To some contemporary linguists, the term “formal generative typology” might seem like an oxymoron. It is often thought that typology can only be properly pursued from a functionalist perspective, and that formal-generative techniques are inherently narrow and hard to generalize to new languages. In this article, I seek to show that this need not be so. In so doing, I lay out the basic goals and concepts that make FGT a fruitful and consistent enterprise, and to present a few examples that illustrate its hopes and its promise.

I confess that this is a somewhat personal vision. Although virtually all of its important ingredients are borrowed from other linguists, it may be that I am the only one who tries to do exactly what is described here. Nevertheless, I believe that many other talented linguists have research programs that overlap significantly with what I describe, and that they have compatible if not identical visions. And anyone is welcome to join the enterprise to whatever degree they feel inspired to do so.

2. Some key questions for linguistic research

The central questions that FGT seeks answers to are the following:

- (1) What properties of natural human languages are genuinely universal, inherent to the human species as such?
- (2) What properties of natural human languages vary from one human language to another?
- (3) Which aspects of variation are patterned, systematic, and grammatical in nature, and which aspects of variation are random, idiosyncratic, and lexical in nature?

To these three, we can add a fourth question, usually not a focus of direct inquiry, but always on the horizon, giving the enterprise its possible grander significance:

- (4) What do the answers to (1)-(3) imply about the nature and origins of the human mind, of which language is a part and a reflection?

In this particular chapter, I discuss these questions only as they apply to the *syntax* of human languages, where syntax is construed broadly as including the entire compositional system by which complex expressions are built up out of simple morphemes and words. Thus defined, syntax includes not only syntax proper, but also productive morphology and those aspects of semantics that involve composing and calculating linguistic representations internal to the mind of the language user. This limitation is purely a practical one. There is no reason not to think that these four questions are equally applicable to the domains of phonology and the lexicon. However, pursuing them in those domains is outside my expertise.

There is nothing especially novel about these four questions. Any linguist could ask them, and many do. They can, for example, be seen as elaborations of questions that Chomsky has often articulated. For example, Chomsky and Lasnik (1993) state the following as the first two questions on their list of “classical problems in the study of language”:

- (5)
 - a. What does Jones know when he has a particular language?
 - b. How did Jones acquire this knowledge?

Questions (1)-(3) can be derived from these “classical” questions simply by varying Jones and his linguistic background—in other words by comparing the knowledge Jones has of his language and how he acquired it with the knowledge that Li has of her language, and how she got it, and with the knowledge that Mohammed has of his language, and how he got it. Those bits that are shared for all choices of Jones/Li/Mohammed will constitute the answers to question (1). Those bits that differentiate their languages from one another will be the answers to question (2). And for each element in the answer to question (2), the further issue arises as to whether it is an elemental difference or a difference that derives from some other, more fundamental difference. In other words, the question arises as to whether people’s knowledges of language differ in many little ways or in a few big ways. This is question (3).

How does FGT relate to the Minimalist Program? It is easy to get the impression that Chomskian linguistics largely abandoned the quest defined by (1)-(3) in the switch from Principles and Parameters to Minimalism, initiated by Chomsky (1993, 1995) (see also Boeckx, this volume). But I think that there is no deep or meaningful conflict between them. Chomskian linguistics (like other approaches) has always sought to maximize explanation. A simple way to think about degrees of explanation is that explanation is maximized when the ratio of observable phenomena analyzed to theoretical assumptions made is as high as possible. How then might one strive to increase the degree of explanation achieved by (say) the Principles and Parameters theory of circa 1990? There are two obvious strategies to pursue. The first is to reduce the number of assumptions made by the theory, while keeping the phenomena analyzed (ideally, more or less) constant. This is the spirit of the Minimalist Program proper. The second strategy is to increase the number of observable phenomena being analyzed, while keeping the number of assumptions (ideally, more or less) constant. This is the driving force of FGT. Chomsky himself has concentrated primarily on the first approach over the last 15 years, as have quite a few others; hence the Minimalist Program. In contrast, I (for one) have concentrated on the second approach. In practice, the two kinds of projects look rather different when it comes to their day-to-day activities. But there is no principled conflict between the two research thrusts; on the contrary, they should be complementary, even synergistic. And if one looks not so much at Chomsky's own output but at that of Ph.D. students coming out of MIT and allied departments, much of it is in practice making a contribution of the second kind: people are seeking to apply a body of shared theoretical assumptions to Tagalog, or Itelmen, or Ewe, or Passamaquoddy, or whatever.

3. How the questions can be pursued: borrowing the best of both

In recent linguistic history, there have been two somewhat different senses of the term “universal”, and hence two rather different approaches to finding universals. It is easy for the two to misunderstand each other. A hope of FGT is that these two approaches can be found to be complementary, rather than in conflict with each other, or irrelevant to each other.

3.1 Universality and the needs of language acquisition

Chomsky's own answer to question (5b) has famously been built on the assumption that there is a fairly rich initial state of the language faculty, which is what he calls “Universal Grammar” (UG). Moreover, this “Universal Grammar” is thought, in point of fact, to constitute a rather high percentage of the answer to (5a) as well; this is Chomsky's famous nativism. The word “Universal” in “Universal Grammar” is thus used in a somewhat counterintuitive sense. When one first hears this term, it is natural to think that it refers to rules or principles of grammar that are part of all human languages. But that is not Chomsky's primary meaning. More properly, Chomsky's Universal Grammar is the preprogrammed biases that a human being brings to bear on the task of learning any natural language. Once this terminological point is realized, Chomsky's scandalous, counterintuitive, ethnocentric sounding claim that “A great deal can be learned about UG

from the study of a single language” (Chomsky 1981:6) loses most of its scandalous character. Rather, it is virtually a truism.

To make the discussion more concrete, consider the following simplified-but-not-entirely-misleading example. It is well known that English is a rather typical example of a Subject-Verb-Object language, a word order shared by some 35% of languages of the world (Dryer 2005). As such, not only does the verb come before the object, but the auxiliary verb also comes before the main verb, and the directional adposition comes before its object:

(6) John will (quickly) give a book to Mary.

Moreover, standard generative (and structuralist) techniques show that the relationship of the verb to the object in English is also different from the relationship of the verb to the subject in a subtler way, that goes beyond the obvious difference in linear order. The verb and the object join together to form a constituent—the verb phrase—which does not include the subject. One reflection of this is the fact that the verb and the object are adjacent in sentences like (6), whereas the verb and the subject need not be; rather the subject can be separated from the verb by the tense auxiliary, or by an adverb. Other evidence confirms this. For example, the verb and the object can be omitted from a sentence leaving the subject and tense marker behind, but the verb and the subject cannot be omitted, leaving the object and the tense marker behind:

- (7) a. John will read the book tomorrow, and Mary will -- -- too.
b. *John will read the book tomorrow and – will -- the newspaper too.

Similarly, the verb and the object can shift together to the beginning of a sentence together in special contexts, whereas the verb and its subject cannot do this:

- (8) a. Mary promised that she would read the book, and [read the book] Mary will.
b. *Mary promised that she would read the book, and [Mary read] will the book.

For current purposes, we state this familiar fact about English in the following way:¹

- (9) *The Verb-Object Constraint (VOC)*: The verb combines with its object to form a linguistic unit that cannot include the subject.

The Chomskian might then go on to observe that some of the empirical evidence in favor of the VOC is rather subtle. Students of a beginning linguistics class often have no firm intuitions about which is more closely connected with the verb in English, the subject or the object. In readily observable linear terms, the subject is often as close to the verb as the object is. Moreover, it is not hard to imagine that some of the children who learn this property of English do so without ever hearing sentences like those in (8), although these were important to establishing the existence of the verb phrase in the

¹ For current purposes, we can say that the object of the verb is the nominal that denotes the entity whose state or position changes the most in the event described by the verb, whereas the subject of the verb is the nominal that denotes the entity that caused the event to take place—although this is a simplification.

minds of linguists. How then do they acquire the knowledge expressed in the Verb-Object Constraint? The Chomskian might conjecture that this is an expectation that they bring to the task of language learning, that they are predisposed to this hypothesis and do not need a lot of conclusive evidence to arrive at it. In other words, (9) might be part of Universal Grammar.

Given this perspective, and the practices that have grown out of it, it is easy to think that the Chomskian approach to linguistics has very little to do with linguistic typology as it is usually practiced. Nevertheless, one can draw a link between the two, given a few plausible assumptions. First, suppose that Chomsky and others are correct that the Universal Grammar that forms the basis of the answer to question (5b) also constitutes an important part of the answer to question (5a). Second, suppose that language learning is a more or less monotonic process. In other words, Jones the language learner *adds* knowledge to her innate store to arrive at full knowledge of English, but she does not discard or radically change what she already has. Third, suppose that the initial state of the language faculty is essentially the same for all human children.² Then, by parity of reasoning, we expect that the same UG that is a substantial part of Jones's knowledge of English is also a substantial part of Hiro's knowledge of Japanese, and of Mohammed's knowledge of Arabic, and of Sak's knowledge of Mohawk. And from this it follows that all human languages must be similar in these respects.

We can see how one is led to this by giving the example above a crosslinguistic dimension. The Nigerian language Edo, although historically unrelated to English, has the same word order as English in all relevant respects, as shown in (10).³

- (10) Ozo ghá rhié èbé ne Emeri.
Ozo FUT give book to Mary
'Ozo will give a book to Mary.'

Chomskian reflection on the logic of language acquisition leads us to think that children must arrive at the VOC in (9) from their innate predispositions, plus at most observations of data like (6), since there is no guarantee that they will observe examples like (7) and

² This assumption is not logically necessary, but we have good observational reasons to believe it. It is logically possible that the innate assumptions about language that different children start with are significantly different (they have different "UGs"), perhaps as a result of genetic variation. But we observe that—barring serious genetic defects—any child can learn any natural human language perfectly if it is raised in the right environment, regardless of (for example) its genetic background. For example, a child with Chinese genes who is brought up in an integrated suburban American environment acquires an English that is indistinguishable (up to normal idiolectal variation) from the English acquired by a child with Western European genes growing up in the same neighborhood. Furthermore, as far as we know, they learn the language in essentially the same way, passing through the same kinds of intermediate stages.

³ The abbreviations used in this article are the following: 2sO, second person singular object agreement; 2sS, second person singular subject agreement; 3sO, third person singular object agreement; 3sS, third person singular subject agreement; ACC, accusative case; ADV, adverb; DUP, duplicative; FACT, factual mood; FGT, Formal Generative Typology; FUT, future tense; HAB, habitual aspect; IMPF, imperfective aspect; IND, indicative mood; M, masculine; N, neuter; NOM, nominative case; PAST, past tense; PRT, particle; PUNC, punctual aspect; STAT, stative aspect; UG, Universal Grammar; VOC, the Verb Object Constraint; WALS, the *World Atlas of Language Structures*

(8). But then the child exposed to Edo is in the same position as the child exposed to English—even though there is no process of VP fronting or VP ellipsis in Edo. Therefore, we expect that the object and the verb form a constituent that excludes the subject in Edo as well. And further investigation shows that this is correct—even though the evidence that is available to show that this is correct in Edo is rather different from the evidence available in English (Stewart 2001).

3.2 Universality and observing diverse languages

Now, what about languages in which the basic word order of subject, object, and verb are not the same as in English? The form of argument just sketched is more general, and does not necessarily depend on there being similarities of word order. Its upshot is that if there is a Universal Grammar in Chomsky’s language-acquisition-oriented sense, one expects that there will be observable universals of language in something like Joseph Greenberg’s sense as well.

Greenberg (1963) famously initiated the search for facts about grammatical patterning that are observably true in representative samples drawn from the set of all natural languages. These are patterns that recur in languages that are not (recently) historically related—universals of language in a more obvious sense. (11) is a classic example of this sort of universal:

- (11) **Universal 4:** With overwhelmingly greater than chance frequency, languages with normal Subject-Object-Verb order are postpositional.

Consider, for example, Japanese, a canonical positive example of Greenberg’s Universal 4. The Japanese equivalent of English (6) or Edo (10) is (12).

- (12) John-ga Mary-ni hon-o yat-ta.
 John-NOM Mary-to book-ACC give-PAST

At first glance, the differences between Japanese and English/Edo are more striking than the similarities. In particular, the position of the verb relative to the object is different, the position of the adposition relative to its associated NP is different, and the position of the auxiliary verb with respect to the main verb is different. This is the other very common word order, found in some 40% of languages of the world (Dryer 2005).

Now, given this salient difference in word order, what do we make of the conjecture that the VOC is part of Universal Grammar, an innate bias for language acquisition? If it is, then the VOC will influence the way that a Japanese child learns Japanese too, and hence it will influence the structure of mature Japanese. And indeed there is reason to believe that this is true. First, despite the different word order, Japanese is like English and Edo in that the direct object is adjacent to the verb in the simplest, least marked word order, whereas the subject is not.⁴ Moreover, this phrasal grouping can be confirmed by other, less obvious tests. For example, Japanese has a process

⁴ Of course, both Japanese and English also allow marked word orders for special pragmatic purposes, and we must abstract away from this. There has been much discussion of how to identify basic word orders in both the functionalist-typological and generative literatures.

similar to VP anaphora in English, in which the verb and the object may be omitted, but the subject remains (example from Tsujimura 2006):

- (13) Taroo-ga tori-o uti-korosita-node Hanako-mo soo-sita.
Taro-NOM bird-ACC shoot-to-death-because Hanako-also do.so.PAST
'Since Taro shot a bird to death, Hanako did so too.'

So the generative conjecture that the VOC is universal holds up rather well. Indeed, the Greenbergian universals tacitly presuppose that head-final languages like Japanese and head-initial languages like English have essentially the same phrasal groupings across the board (see Baker (2001: ch.3) for discussion).

Looking more broadly, it is consistent with the VOC that the two most common word orders in the world, together accounting for more than 75% of all languages, both have the object adjacent to the verb. The VOC thus captures something important that those two word orders share that might help account for their popularity. The VOC also has the virtue of playing a role in accounting for both obvious facts of word order and less obvious properties (like possible ellipses) in the same terms.

More generally, then, a core interest of FGT is evaluating whether this sort of convergence between the two sorts of universal is found in general or not. We want to see if it is true in practice that what one is led to attribute to Universal Grammar from the detailed study of (say) English, driven by the need to account for how its details discovered could have been learned from simple and unsystematic data, is really the same as what one is led to attribute to Universal Grammar from the detailed study of Japanese, or Mohawk, or any other language. In this, the conception of what language is that informs FGT is solidly Chomskian, but the data considered and the method used has something important in common with Greenbergian typology, with its emphasis on collecting data from a wide sample of the world's languages and looking for patterns that recur in unrelated languages. Only when we find such patterns can we say with confidence that grammar is truly universal.

3.3 Universals and abstractness

But despite the important point of similarity, there is still a methodological difference between FGT and the functionalist-oriented practice of typology. A crucial issue is the level of theoretical abstraction at which the search for universals and patterned variation takes place. Perhaps the most constant feature of functionalist typology in the tradition of Greenberg, Comrie, and others over the last fifty years is that it focuses on superficial features of languages—features that can easily be described in the relatively atheoretical terms of traditional descriptive linguistic practice. For example, Bickel (2007:242) writes, in describing the current typological scene:

But not everything has changed: most prominently, as in the past, typologists find it useful to develop variables as close to observable data [operationalized criteria] as possible and close to fieldwork. This is first of all a practical decision, because very abstractly defined variables are difficult to survey in sufficiently large samples, and samples can often only be completed by doing additional fieldwork. But the decision

is also theoretically motivated because the definition of abstract variables is also tied to some UG model that itself seeks to abstract away from linguistic diversity, and less so to the kinds of anthropological hypotheses of interest.

This commitment to studying nonabstract, surface-observable properties of language is one of the most constant features of typology as it is usually practiced, a near corollary of its defining interest in studying representative samples of unrelated languages. As Bickel acknowledges, it is a commitment with practical motivations. The notions it uses are usually fairly easy to define, they can often be found in standard descriptive grammars, and the results can be replicated by others (in principle, anyway). But Bickel also correctly points out that this is a theoretical choice as well as a practical one. It amounts to a denial, implicit or explicit, of the value and reality of more abstract generative concepts—concepts such as phrase structure, c-command, “movement”, and the like. Functionalist typology is often motivated by a kind of positivistic empiricism, which wants to see the theory emerge from the data, rather than having a theory imposed onto the data (see, for example, Croft 2002). It is an attempt to avoid the acraneness, the question-begging, and the immunization from counterexamples that generative abstractness at its worst makes possible. And I can easily understand why one would want to avoid these things.

But it should also be acknowledged that traditional typology has in a sense failed, and thus has been led to change its goals.⁵ It has by and large been unable to discover many interesting universal properties of natural human languages. At most, standard typologies find statistical tendencies of various degrees of strength. Thus, Bickel (2007:245) also writes:

Large datasets almost invariably reveal exceptions to universals, and this, together with a substantial increase of newly described languages and assisted by prominent conceptual argumentation (e.g., Dryer 1998, Croft 2002: Chapter 8), has practically done away with notions of absolute universals and impossibilities. Modern studies of typological distributions involve statistical methods, from association tests to multivariate scaling methods.... The general assumption is that if there are large-scale connections between linguistic structures, or between linguistic structures and geography, they consist in probabilistic (and therefore exception-ridden) correlations between independently measured variables.

This retreat from a universalist vision is also borne out in the massive, wonderful, and frustrating *World Atlas of Language Structures (WALS)* (Haspelmath et al. 2005), which maps out some 140 linguistic features over hundreds of languages each. Haspelmath (personal communication) has observed that computational techniques have been used to systematically mine through the work, looking in a purely mechanical way for statistically significant correlations among the properties cataloged (Cysouw et al. 2007). The result was that very few correlations emerged, and most of the ones that did looked crazy, correlating (say) some phonological feature with some syntactic feature,

⁵ Of course, many researchers within the paradigm (such as Bickel) would describe this not as a failure, but as a healthy discovery that moves the field forward toward greater truth and enlightenment.

which no plausible linguistic theory could relate directly. It thus seems to many that the search for substantive linguistic universals has been tried and has failed.

It is, however, entirely possible that many absolute and implicational universals are out there, but they can only be seen at a higher level of abstraction. Consider again the Verb-Object Constraint. When this is translated into a prediction about word order in the simplest possible way, we get a statement like “the verb and its object will be (nearly) adjacent to one another, whereas the verb and its subject need not be.” This has the status of a strong statistical tendency (true for some 92% of languages with fixed word order) but not an absolute universal (false for some 8%). In other words, it is too promising to abandon, but there are certainly exceptions to it, just as Bickel says. Almost all of the exceptions are languages that have Verb-Subject-Object word order. Although much noticeably less common than Subject-Verb-Object order and Subject-Object-Verb order, it is not uncommon, and is found in different parts of the world. Closest at hand are the Celtic languages, which have been fairly well-studied from a generative perspective:

- (14) Gwelodd Siôn ddraig. (Welsh (Sproat 1985))
See.PAST.3sS John dragon
'John saw a dragon.'

Now there are several things to note in assessing this “counterexample” to the VOC. First, by all accounts, Verb-Subject-Object order is a relatively minor variant of Subject-Verb-Object order; only the position of the verb and the subject relative to each other is different. Adpositions still come before NPs, auxiliaries before main verbs, and so on. It is also notable that, in Welsh, SVO order surfaces when the tense marker and the verb do not fuse into a single word:

- (15) Gwnaeth Siôn weld draig. (Sproat 1985)
do.PAST.3sS John see dragon
'John saw a dragon.'

We also know independently that the verb moves to join with the tense marker in some languages and not others. For example, French and English are taken to differ from each other in just such a way. As a result, finite verbs come before adverbs in French ((16c)) but after adverbs in English ((16d)), even though nonfinite verbs come after adverbs in both languages ((16a-b)) (Pollock 1989).

- (16) a. Jean a souvent embrassé Marie. (French)
b. John has often kissed Marie. (English)
c. Jean *embrasse souvent* Marie. (French)
d. John *often kisses* Marie. (English)

Second, there is some reasonably subtle evidence that the subject moves from the beginning of VP to the beginning of the clause in English. The evidence is that a quantifier associated with the subject can be “left behind” in the immediate pre-verbal position (Sportiche 1988):

- (17) a. All the children will – find a candy.
b. The children will all – find a candy.

If we accept these two kinds of movement as legitimate grammatical processes, motivated to explain certain details about English and how it differs from French, then an easy account of Welsh emerges: Welsh is simply a language in which verbs move the way they do in French, but subjects do not move the way they do in English (Koopman and Sportiche 1991). So the possibility of Verb-Subject-Object languages emerges naturally out of possibilities that generative linguistics is committed to anyway. They are not exceptions to the VOC, when it is understood as an abstract claim about linguistic structure, not as a surface property of word order. And indeed, there is other evidence for the VOC in Celtic languages (Sproat 1985, McCloskey 1996).

Note also that a Verb-Subject-Object language is created only if a series of properties falls into place in a certain way: heads must come first in their phrases, verbs must raise to the tense position, and subjects must not move to the beginning of the clause. If each of these parameters of variation is set this way roughly 50% of the time, then we would expect to observe Verb-Subject-Object order in roughly 12.5% of the languages of the world—and that is not far from the observed figure. FGT can thus explain why one language type is less common than another when a constellation of factors is needed to produce that type (see Baker 2001: ch.5 for more discussion).⁶

There are some obvious dangers here. It is certainly true that using abstract linguistic theories makes it harder to apply those theories to new languages. It also opens up opportunities for various kinds of cheating when it comes to evaluating hypotheses. For example, generative linguists might hide themselves from all counterexamples by making hasty appeals to movement, even when there is no evidence for it, the way that there is in Welsh.

But do we have any reason to believe that we can do without it, even if it seems both safer and more convenient to do so? I think not, for at least two reasons.

First, when one studies one particular language in detail, one finds that a degree of abstractness is needed to give the best possible account of that language within its own terms. One familiar case in point is agreement paradigms: it is often necessary to posit morphologically null agreement affixes to complete a paradigm that has uniform grammatical behavior, or to distinguish two homophonous affixes that are actually agreeing with different underlying categories. Similarly, in the study of case inflections on noun phrases, one might have to posit a unified notion of (say) accusative case at an abstract level, even though it is realized by different observable morphemes in different circumstances (see, for example, Legate 2008, who shows how this is crucial to understanding the phenomenon of ergative case marking in some languages). Linguists may of course debate just how much abstraction is warranted when it comes to truly and fully describing the grammar of a particular language; I personally think that the answer is “quite a bit”, especially as one tries to approach the generative ideal of a full description, one that does not presuppose “linguistic common sense” on the part of the

⁶ See also Chung 1998 for a different proposal for deriving Verb-Subject-Object order in Chamorro from a structure that obeys the VOC by way of a nontrivial movement process.

grammar-user but that tries to explicate what that “common sense” consists of.⁷ But putting this debatable matter partly aside, suppose that we accept in principle that the grammars of natural languages studied in their own terms are found to be abstract to *some* nontrivial degree. Then it inevitably follows that comparing grammars to see how they are the same and different will have to be done at this same level of abstractness. There is no reason to think that one could avoid this abstractness except perhaps in a few lucky, nonrepresentative cases in which the abstract categories happen to map straightforwardly onto surface categories.

A second way of making essentially the same point comes from thinking again about the logic of Chomsky’s acquisition-based sense of Universal Grammar. What kinds of knowledge of language does the generative linguist most want to attribute to the initial state of the human child? In point of fact, we typically want to attribute those facts that are *more* abstract and remote from obvious every day experience to that initial endowment. There is (maybe) no serious language acquisition problem for the more obvious facts about word order, case marking, agreement, and use of lexical categories in a language. These matters are saliently and abundantly attested in the sample of language that is presented to any child, so they could be learned from the data by some kind of inductive process. The real acquisition puzzles come from those subtle and surprising but robust and replicable discoveries at the corners of a grammatical system that are often discovered by a formal-generative attention to detail, explicitness, and the special issues that can arise when simple structures are composed to form more complex structures. These include distinctively generative discoveries such as the Binding Conditions on pronouns and anaphors, the so-called Island Conditions that constrain movement operations, conditions on the interpretation of quantified expressions, and so on. Often distinctive facts about these matters show up clearly and unambiguously only in sentences of a certain complexity, sentences with a very particular combination of properties. For such properties, it is hard to imagine that every language user that demonstrates the knowledge was exposed to a sufficient number of the relevant structures—and that they noticed them, *and* that they realized their significance for the grammatical point in question. These then are the grammatical properties that we have the most reason to attribute to Universal Grammar qua the innate endowment for language. It then follows that these are the grammatical properties that we have the most reason to expect to be universal also in the Greenbergian sense of being observable in all human languages. So we expect that the most abstract properties of language—the very hardest properties of that language for a linguist to discover, and thus the issues most rarely discussed in descriptive grammars—also to be the most universal properties of language.

Informal experience suggests that this may well be true. Of the ten or so non-Indo-European languages from various families that I have done serious fieldwork on, every one has phenomena that are recognizably like Chomsky’s (1981) Binding Conditions and Ross’s (1967) Island conditions. (18) shows some familiar-looking contrasts from Mohawk, a language that is otherwise very different from English (Baker

⁷ Note that this scientific ideal is much more ambitious than what normally counts as “describing a language” in the current linguistic scene. Therefore, the fact that one might be able to do without significant abstractness in completing what now counts as a decent descriptive grammar of a language does not at all imply that abstractness is unnecessary for the ultimate goal of language description.

1996: ch.2); note that the Mohawk examples have the same grammatical status as their English translations.

- (18) a. U'hka í-hs-ehr-e' Uwári ruwa-núhwe'-s?
 who Ø-2sS-think-IMPF Mary 3sS/3sO-like-HAB
 'Who do you think Mary likes?'
- b. *Ka nikáyΛ áthere' she-yΛtéri ne yakó-hs-u.
 which basket 2sS/FsO-know.STAT PRT 3sO-finish-STAT
 'Which basket do you know the woman who made (it)?'
- c. *U'hka wa'-te-sa-hΛ'reht-e' ne tsi Uwári
 who FACT-DUP-2sO-shout-PUNC because Mary
 wa-huwa-rasΛ'tho-'
 FACT-3sS/3sO-kick-PUNC
 'Who did you shout because Mary kicked (him)?'

It is easy to multiply such examples, and no one is surprised these days when such things are discovered in a new language. But these matters are not considered at all in the *World Atlas of Language Structures*, nor in the sorts of typological databases that Bickel refers to. Of the 140 maps in *WALS*, not a single one concerns anaphor binding, quantifier scope, extraction from a complex noun phrase—or even the Verb Object Constraint. They all have to do with word order, agreement, case marking, category membership, and other superficial morphological categories. Thus, standard typologists have looked hardest for universals in exactly those domains where generativists least expect to find them, and have hardly looked at all in those domains where generativists predict that they exist. It does not come as a surprise, then, that functionalist typology by itself has found little in the way of linguistic universals.

Why hasn't anyone done a *WALS*-style map that documents the distribution of (say) island effects in languages of the world? Even if one wanted to do so, it would be a complicated endeavor. While it is true that all the languages I have worked on have recognizable island effects, they are not all the same in this regard. These matters necessarily have to do with the interactions among phenomena. Therefore, any differences in the phenomena themselves will inevitably cause differences in the interactions that one would expect. It matters to island effects whether the phrase being extracted from sits in the object position or has been extraposed to the edge of a domain. It matters whether the extracted phrase is moved overtly or covertly. It matters whether the original position is occupied by a gap or by a resumptive pronoun. And so on. So even if a grammatical condition is universal in the strongest and simplest possible way, its observable effects will not be universal and invariant. For example, one can extract a question word out of the understood direct object in English but not in Mohawk:

- (19) *?U'hka se-núhwe'-s ne ako-kára'?
 who 2sS-like-hab PRT her-story
 'Who do you like stories about?' (OK in English)

I believe that this difference does not undermine at all the claim that the same island conditions apply in both languages; on the contrary, the difference is expected once one realizes that the “object” is in a different position in Mohawk than it is in English (see next section). The point is that the observed effects inevitably depend on how the condition under investigation interacts with other aspects of the language, which may themselves vary. So the observables will be complex and multidimensional; it is not clear that they could be expressed in a useful and meaningful way in a simple map format, even once we pull together enough relevant material.

It is possible that the generativist’s impression that there is something universal about the core island conditions (for example) will turn out to be an illusion. Perhaps the various language-specific island-like phenomena we observe cannot truly be unified into a single condition, and it is only because of vagueness in our understanding of the phenomena that we can see the conditions at work in different languages as being the same. It has happened in the history of generative grammar that some attempts to unify similar-looking conditions into a single, more general condition have either failed or proved to be Pyrrhic victories—claimed successes that were so complex or inflexible that they turned out to be the equivalent of a failure in the end. Much depends on whether the differences between the island effects observed in language A and those observed in language B can be correctly and insightfully attributed to independently observable differences in the phrase structure, or in the word order, or in what exactly is moving, or in where it is moving to, or in what is left behind by the movement (or something else). There are many encouraging results of this sort, and also many outstanding problems. But this discussion is enough to show why the putative result of functionalist typology that there are no solid linguistic universals does not really speak to the issue as it arises within a generative perspective.

3.4 Interim summary

Part of the motivation, then, for a distinct formal generative typology comes from the belief that standard functionalist typology is inadequate as a way of fully answering questions (1)-(3) because of its deep-set aversion to abstractness in linguistic analysis. Its techniques are not even adequate to tell us if something relatively straightforward like the Verb-Object Constraint is true or not. Many typologists have been guilty of a degree of laziness in not striving to understand the grammatical structures of the languages they are drawing on in their surveys, with the result that they may not ask the most important questions, cannot always filter out interfering factors, and are not in a position to recognize indirect and unforeseen consequences of the factor they are interested in. For the most part, they have not found the level of abstraction at which questions (1)-(3) can truly, insightfully, and productively be answered.

The other impetus for a distinct formal generative typology is that formal-generative linguists have also been guilty of a degree of laziness. We have usually not bothered to do the work we need to do to prove the genuine universality of our claims about Universal Grammar. It is all very well to predict that the surprising details discovered in the corners of one language will be universal, but we need to have methods for testing and refining these predictions. Despite the healthy sense that crosslinguistic comparison plays a larger role in generative work than before, it still falls far short of

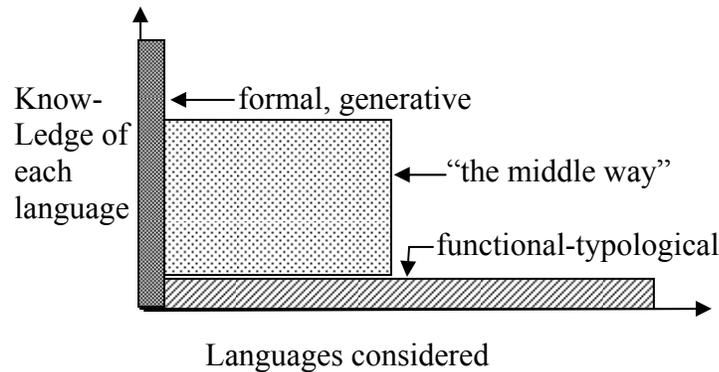
what is attractive and right about the typological ideal. It is now commonplace to find examples from four or five languages cited in a generative linguistics article, but typically those languages are all from a single language family (usually Indo-European), or at most from two language families (for example, comparing an East Asian language to English). So while crosslinguistic comparison is on the rise, true typology is not. Conspicuously absent in most of this work is the typologist's vision of controlling for genetic and areal factors by sampling from unrelated languages and language families, and looking for naturally occurring replications of particular linguistic systems in other parts of the world. As a result, generative linguistics also has not made that much progress in establishing whether something like Verb-Object Constraint is universal or not. Generative linguistics has something important to learn from typological approaches in this regard.

4. A distinctive method

Of course, we are all lazy. More charitably put, we all have limited time, energy, and resources. Therefore, we need some realism as well as some idealism in our research enterprises. Although it would be great to know everything about everything, it is not feasible in this life, and we need to make choices. We also need to identify intermediate stopping points along the way where we can evaluate and take stock of our progress. How can FGT address this need?

My primary suggestion is to aim for what Baker and McCloskey (2007) dub “The Middle Way”. This is simply to do an intermediate amount of linguistic research on an intermediate number of languages. There was a joke at the engineering school I attended that our school's strategy was to teach us “more and more about less and less until we knew everything about nothing.” This was contrasted with the strategy of the liberal arts school up the street, which was said to be to teach “less and less about more and more until the students knew nothing about everything”. The point of the joke is, of course, that despite the differing educational approaches, both student bodies end up in the limit knowing exactly nothing. Stereotypical formal-generative research risks achieving irrelevance via the tech school route, whereas stereotypical typological research risks achieving irrelevance via the liberal arts school route. The obvious alternative is simply to try to know something about something—the Middle Way. It is a simple mathematical fact that the way to maximize the area of a rectangle given a fixed perimeter is to make its height and its breadth equal. In the same way, linguistic understanding given finite resources is likely to be maximized by striking a careful balance between the range of languages considered and the depth of knowledge about each language, as sketched in (20).

(20)



In practice, this might involve working with samples of 10-20 unrelated languages, rather than with sample sizes one or two orders of magnitude greater, as is now common in typological studies. Each language, however, would be considered in generative-inspired detail, looked at in terms of its underlying grammar, not just its easily observed surface properties. Surprisingly few linguists are attempting this as a consciously-chosen strategy to identify universals and recurring patterns in diversity.

To illustrate the advantages of knowing an intermediate number of languages in an intermediate level of detail, consider evaluating the relevance of the Verb-Object-Constraint for the Mohawk language. On first glance, this language is perhaps an even stiffer challenge to the VOC than Verb-Subject-Object languages like Welsh are. It is a language with no identifiable basic order at all (Mithun 1987), in which all of the following word orders are attested:

- (21)
- | | | |
|----|--------------------------------|------------------------|
| a. | Sak ranuhwe's ne atya'tawi. | (Sak likes the dress) |
| b. | Ranuhwe's ne atya'tawi ne Sak. | (Likes the dress Sak.) |
| c. | Ranuhwe's ne Sak ne atya'tawi. | (Likes Sak the dress.) |
| d. | Sak atya'tawi ranuhwe's. | (Sak the dress likes.) |
| e. | Atya'tawi Sak ranuhwe's | (The dress Sak likes.) |
| f. | Atya'tawi ranuhwe's ne Sak. | (The dress likes Sak.) |

Hence no one well-defined and independently motivated kind of movement will save the day for the VOC, the way it did in Welsh. So the VOC might seem like a nonstarter for Mohawk.

But someone with a broader knowledge of Mohawk is in a position to see that this conclusion would be hasty. The VOC may not be visible in *syntax* of Mohawk, but it is visible in the morphology of Mohawk—in particular, in the phenomenon of noun incorporation. The object of the verb can optionally be realized inside the inflected verb as a kind of noun+verb compound ((22b)), but the subject can not be ((22c)).

- (22)
- | | |
|----|---------------------------------------|
| a. | Owira'a wahrake' ne o'w <u>ahru</u> . |
| | baby ate the meat |
| | 'The baby ate some meat.' |

- b. Owira'a waha'wahrake'.
baby meat-ate
'The baby ate some meat.'
- c. *O'wahru wa'kawirake.
Meat baby-ate
'The meat was baby-eaten.'

Some other illustrative examples of object (but not subject) incorporation in Mohawk are given in (23).

- (23) a. Wa'eksohare' 'She dish-washed.'
- b. Wa'kenaktahninu' 'I bed-bought'
- c. Wa'khwistatshvri' 'I money-found'
- d. Wahana'tarakwetare' 'He bread-cut'

Why does this difference exist? Although the details of different accounts vary, it is very plausibly related to the Verb-Object Constraint in (9). Indeed, it follows immediately from this constraint if we now interpret the "linguistic unit" that can contain the object and the verb but not the subject as being the inflected verb, rather than verb phrase, a morphological unit rather than a syntactic one. We then observe the same constraint applying to different but analogous linguistic units to explain a significant parallelism in the data.

Our confidence that this is a related phenomenon is increased by considering English compounding. An understood object can also appear compounded with a deverbal noun in English, whereas an understood subject cannot. The examples in (24) and (25) are thus parallel to the Mohawk examples in (22) and (23) in this respect.

- (24) a. meat-eating is disgusting. (= the eating of meat)
- b. #baby-eating is disgusting.
 (not eating *by* babies is disgusting; only the eating *of* babies)
- (25) a. meat-eating, meat-eater
- b. dishwashing, dishwasher
- c. cheese-cutting, cheese-cutter
- d. car-buying, car-buyer

So English compounding serves as a conceptual bridge. It is easy to see noun incorporation in Mohawk and compounding in English as two manifestations of a common underlying truth. It is also not hard to see compounding in English and phrase structure in English as being related phenomena—a generalization first captured (for English) as Selkirk's (1982:37-38) "First Order Projection Condition". Therefore, by transitivity, all three are related, and the Verb-Object Constraint applies to Mohawk as well as to English. This shows that, when doing typological investigation, we need to know a good percentage of the languages in our typologies well enough to recognize

when there are nonobvious ramifications of potentially universal conditions or properties, or we may conclude that there are fewer genuine universals than there are.

We can also illustrate the advantages of the Middle Way from the negative side. Would more knowledge about Mohawk allow us to say anything more useful why Mohawk *sentences* do not show evidence of the VOC, even though Mohawk *verbs* do? Here I can give only the briefest sketch of a form of explanation that I worked out in detail in Baker (1996). The following sentences in English show that, although true objects must be next to the verb in English, “dislocated objects” need not be:

- (26) a. That dress, Sak really likes it
b. Sak really likes it, that dress.

We would not say that a sentence like (26a) violates the VOC: the object for grammatical purposes is the pronoun, and this pronoun does form a phrase with the verb. *That dress* is not the direct object, but a topicalized phrase that the direct object refers to. Now Mohawk could also have object dislocation, and that would not tell against the VOC. But what would dislocation look like in Mohawk? It so happens that in Mohawk, weak subject and object pronouns are phonologically null. Thus, a sentence like ‘He likes it’ can be expressed in Mohawk just by the inflected verb.

- (27) Ra-nuhwe’-s
3sS.M/3sO.N-like-HAB
‘He likes it.’

Nor is it surprising that Mohawk allows null pronouns. Mohawk verbs bear prefixes that agree with both the subject and the object; for example, the prefix *ra-* in (27) indicates that the subject is masculine singular third person and the object is neuter third person. Because this information is expressed in the verbal morphology, overt pronouns are not needed, just as subject pronouns are needed in English but not in Spanish or Italian.

Now, given this simple fact about Mohawk, what would dislocation sentences analogous to (26) look like in Mohawk? They would look like (28).

- (28) a. Atya’tawi Sak ranuhwe’s
dress Sak likes (it)
b. Sak ranuhwe’s ne atya’tawi.
Sak likes (it) the dress

In fact, they would look like the grammatical sentences in (21a) and (21e). We thus have the appearance of free word order of the object, and with it violations of the VOC. But in English we say that these are not real violations of the VOC, because the true object is the pronoun, in the verb phrase, adjacent to the verb, where it belongs. We can say exactly the same thing about Mohawk. (28a)/(21c) is not a counterexamples to the VOC, either: the object is in the verb phrase, next to the verb, in Mohawk too, for all we know—we just don’t see it because object pronouns are null in Mohawk, for reasons that are both predictable and independently observable.

Is this just a slick trick, or is there other evidence that sentences like (27) are more like dislocations in English than they are like simple (6) in English? In Baker (1996: ch 2) I argued at length that there is lots of independent evidence for this. One class of evidence comes from nonreferential NPs, such as anaphors, nonreferential quantifiers, and idiom chunks. We know that in well-studied languages like English and Italian these sorts of nonreferential NPs cannot be dislocated (**Nobody/*himself, John saw him in town*; Cinque 1990). Furthermore, Mohawk does not have independent NPs of these sorts; there is no NP that is an exact equivalent of nobody or herself in Mohawk, for example. This is just what we expect if overt “objects” are actually dislocated adjuncts in Mohawk.

In this light, recall from section 3.3 that one cannot extract a question word from a putative direct object in Mohawk the way one can in English (see (19)). Now it so happens that one cannot extract a question word out of a dislocated object in English either, as shown by the contrast in (29).

- (29) a. Who did you hear a story about?
b. *Who did you hear it, a story about?

The contrast is expected: dislocated object is really a kind of adjunct, so extracting from it is a kind of adjunct island violation, bad for the same reason that the English version of (18c) is. If apparent objects in Mohawk really have the status of dislocated NPs that are indirectly related to (null) pronouns which are the true objects, then we would expect that one could never extract from these apparent objects in Mohawk—especially since we already have some evidence that the adjunct island condition holds in Mohawk as well as in English ((18c)). Thus, we actually predict the badness of (19). At the time, this seemed like an anomalous fact, calling into question the universality of the island conditions. But now it does not seem anomalous at all; rather it is expected given what else we know about the language.

So the story holds together. Once we understand the grammar of Mohawk to some degree, we realize that we do *not* expect to observe the VOC in Mohawk syntax, because its agreement properties mean that object dislocation is always a possibility and will not look obviously different from nondislocation on the surface. So some real knowledge of the grammar of the language is necessary to evaluating the universality of a condition like the VOC (or the adjunct island condition), not only so that we can recognize reflections of the VOC that we might not have thought to look for, but also so that we do not look for simplistic evidence of the VOC in places that we really should not expect it. Practicing the Middle Way makes this feasible.

Will the results of research done in these ways scale up to larger, more impressive sample sizes? There is no guarantee, of course; this is research, and once we know for sure exactly how to do it, that phase of research is probably almost over. But my recent investigation of the principles and parameters of agreement systems (Baker 2007) suggests that the answer is positive. In this study, I began with a close comparison of agreement in some Bantu languages with agreement in Indo-European languages, then moved to a pilot study of 10-15 languages (the Middle Way stage), and then tested my two parameters against a larger sample of 108 languages, based on the core languages survey of WALS. I found that the move from 15 languages to 108 did not give

appreciably different results, although it did of course serve to illustrate those results in a much wider range of language types. (For example, one could see what different agreement systems look like within languages with different word orders, including unusual ones like Object-Verb-Subject and Object-Subject-Verb.) My guess is that one does not have to look at a very large number of unrelated languages to get most of the benefit of guarding against parochialism in one's linguistic analysis. Nor does it take that long to learn what one needs to know about a language in order to guard against superficial errors and misinterpretations. Along both dimensions, the famous 80-20 rule that 80% of the results are gained from 20% of the effort probably applies. If so, "Middle Way" methodology should be very effective.

Saying that this is what needs to be done by the field as a whole does not imply that this is what each linguist needs to do. There are obvious ways to contribute to an FGT-style program by attempting only pieces of it. For example, doing a generative analysis of a single understudied non-European language that discovers the right level of abstraction for capturing facts about that language, using data and terminology that are accessible to others, is a huge contribution. And linguists can work in teams, use each other's results, and corroborate each other's finding to accomplish a larger portion of this than any of them could do by themselves. In a variety of such ways, one can realistically hope to learn something about something.

5. On the relationship between the subfields of linguistics

What does FGT have to say about the relationships between the different subfields of linguistics? Since it is more of a research methodology than an independent theory, it has no fixed distinctive position on those matters. What answers it has are either taken over from its Chomskian/generative inheritance or are discoveries it claims to have made in the course of its pursuits.

FGT certainly assumes a lexicon, which is at least a list of the atoms that can be used in a syntactic representation, together with the properties that distinguish one from another. It is quite possible that the lexicon is also no more than this (Marantz 1997). Syntax, then, is the system by which such elements can be combined into larger, well-formed and coherent linguistic representations. Phonology and semantics are both taken to be interpretive, in the usual Chomskian sense. In other words, we assume for theoretical purposes that a formal syntactic representation is constructed, and then phonological principles calculate how to pronounce that syntactic representation and semantic principles decide what it could be used to mean in a particular context (see also below).⁸

5.1 Morphology and Syntax

⁸ Exactly what sequence of calculations an actual language-user might go through in accomplishing a particular task that uses language is taken to be a partially different matter, and is left largely open. This is a matter of *using* a language rather than a matter of *knowing* a language, and understanding what it is to know a language is assumed to be a necessary although not sufficient condition to understanding what it is to use a language.

The place of morphology is a particularly interesting question within this approach. The layout just sketched leaves it as an open empirical question whether the atoms listed in the lexicon correspond more closely to the traditional idea of a word, or to the traditional idea of a morpheme. In fact, there turn out to be many similarities between word structure and sentence structure, such that we can capture significant generalizations by having abstract principles that are relevant to the formation of both. The Verb-Object Constraint is an excellent case in point. The theme/object NP but not the agent/subject NP combines with the verb to make a verb phrase in English. Similarly, a noun expressing the theme/object but not a noun expressing the agent/subject can compound with the verb root to make a complex verbal stem in Mohawk. We want to say that this is not an accidental similarity; rather, the two are both reflections of the same underlying fact about human language. This is most naturally done if we do not consider morphology and syntax to be two separate domains, but rather include both in the same comprehensive representational system, so that both are subject to the same representational laws.

In fact, there are many ways in which what shows up as a constraint on the syntax in a language like English shows up as a constraint on word structure in Mohawk and similar languages. A second example is the so-called Extended Projection Principle which stipulates that clauses must have subjects, but makes no such requirement about objects (Chomsky 1981). This can be seen in English syntax in the fact that meteorological predicates like *rain* must have a dummy placeholder pronoun in the subject position, but not in the object position:

- (30) a. *Rained yesterday.
 b. It rained yesterday.
 c. *Rained it yesterday.
 d. *It rained it yesterday.

This constraint is not readily observable in the syntax of Mohawk or Mapudungun, since these languages omit all unstressed pronouns. But it can be seen in the morphology of Mohawk and Mapudungun. These languages require the verb to bear a subject agreement affix, but they do not require the verb to bear an object agreement affix:

- (31) a. v-yo-kvnor-e' (*v-kvnor-e') (Mohawk)
 FUT-3sS-rain-PUNC FUT-rain-PUNC
 'It will rain.'
- b. Petu mawün-üy. (*mawün-fi-y) (Mapudungun)
 ADV rain-IND.3sS rain-3sO-IND.3sS
 'It is raining.'

Yet another example of this involves concerns anaphora. English contains a difference between the overt pronouns *him* and *himself*, such that the marked pronoun *himself* must be used if and only a reflexive interpretation is intended (Reinhart and Reuland 1993). Mohawk does not have a difference between two pronominal forms; there is only one masculine singular pronoun, *rauha*. But Mohawk does have a parallel

morphological difference: a verb marked by the special prefix *-atat-* must be used if and only if a reflexive interpretation is intended.

In all three cases, the same (abstract!) constraint seems to be at work, even though its effects show up most clearly in the syntax of some languages and in the morphology of others. Hence, typological research that is open to abstract generalizations discovers that morphology and syntax are subject to the same potentially universal principles. This in turn suggests that they are not fundamentally different areas of study, but aspects of the same grand compositional system. (See also Cinque 1999 for ordering conditions that govern both the positioning of adverbs in some languages and the order of tense-mood-aspect morphemes in other languages.)

6.2 Syntax and pragmatics

The results of FGT also have implications for controversial questions about the relationship between syntax and pragmatics. Everyone agrees that the two are interrelated in some way. The controversial question is which one leads and which one follows. On one view, pragmatics is the more basic study, and syntax is the crystallization (grammaticization) of pragmatic functions into more or less iconic grammatical forms. On the other view, syntactic principles determine what sentences can be formed, and then pragmatics takes the range of syntactic structures that are possible and assigns to each of them some natural pragmatic use(s) that take advantage of the grammatical forms that are available. The first view is characteristic of functionalist approaches to linguistics; the second is the traditional Chomskian position.

What we have seen about the structure of Mohawk seems relevant to this matter. English permits both structures in which the direct object is in place and structures in which the direct object is dislocated. These two structures have distinct pragmatic uses, in a consistent, natural, iconic way. In contrast, Mohawk permits only the dislocation structure, (in part) because agreement with the object is grammatically necessary in Mohawk. Therefore, only one of the two structures is available. The dislocation structure is thus forced to cover a wider pragmatic range in Mohawk than in English. For example, only definite NPs with a topic reading are dislocated in English, but there is no such requirement on dislocation in Mohawk: indefinite, nontopical NPs are possible in the very same positions that definite NPs are in Mohawk.

There is another domain in which Mohawk allows a larger set of possibilities than English does: Mohawk allows an object to be incorporated into the verb, whereas English does not. As a result, Mohawk speakers can draw a pragmatic distinction between incorporated nouns and unincorporated ones: incorporated nominals are backgrounded, and contrastive focus is incompatible with incorporation. In contrast, ordinary independent objects are used for backgrounded as well as for focused interpretations in English, that being the only relevant structure available.

Therefore, it seems then that pragmatic considerations—which are taken to be more or less universal—cannot in themselves explain the different ranges of structures that are available in different languages. What pragmatic considerations can do is, given a range of well-formed structures, say something about what each one might naturally be used for. That is just what one would expect if semantics/pragmatics has an interpretative relationship to syntax, rather than a formative one. I am not in a position to

say that this state of affairs cannot be made sense of from a functionalist, pragmatically-driven perspective; they are certainly well aware of such facts. But at the very least, there is no reason here to abandon the syntax-driven approach inherited from FGT's generative roots. On the contrary, FGT-style research seems to confirm that form is partially distinct from function.

6. Types and Levels of Explanation

This leads naturally to the question of what kind of explanation does FGT seek to offer, and what relationship does it see between language and other areas of cognition. The short answer is that it offers language-internal explanations, rather than language-external ones. One feature of a language is explained in terms of its similarity to another, at first different-seeming feature of that language and another language, by saying that both are consequences of the same general principle. As such, the explanation of a feature is *not* primarily in terms of its pragmatic value, or its iconic relationship to other cognitive systems, or in terms of its diachronic origins.

To take a familiar example, FGT says that languages have NP-P order also have object-verb order (see (11)) simply because NP and P constitute a phrase, the object and the verb also constitute a phrase (see (9)), and there is one general rule for how phrases are built in a language, with the "head" of a phrase either coming first (English) or last (Japanese), all things being equal. These two sorts of phrases have common properties because they both products of the same phrase-building machine, just as two bullets might have similar markings because they were both shot from the same gun.

This is not to deny that there may ultimately be external, functional explanations for some of these matters as well. But even if the language internal explanations of FGT are not ultimate explanations, they are still real, significant, plentiful, and valuable. Indeed, part of their value is that they clarify the full nature of natural language, with all its interconnections, and thus reveal what its deepest properties are that may call for another level of explanation. For example, the discussion above suggests that the Verb-Object Constraint might well be universal property of human languages. If so, why should this be? I do not know, and I have never heard someone attempt a really serious explanation, even though the basic claim has been known to generativists for some time. It may very well show us something deep and important about how human beings conceive of events and the entities that take part in them, something that might have been difficult or impossible for us to discover otherwise.

Some might take it as a weakness of FGT that it offers no specific tools for answering "type (4)" questions like this. But what I find exciting about FGT is that it does succeed in *raising* such questions. The VOC could be a true universal of language, and it is certainly true of a very wide range of languages, even though it shows up in different ways in different languages. No other approach could really make and confirm this discovery: it is too abstract for functionalist typology to find, whereas it is too sweeping a claim for a nontypological generative approach to confirm. Much FGT style research has gone into this discovery, and much more may still be needed to see if it is really valid for all the languages of (say) Amazonia and New Guinea.⁹ If in the long run

⁹ A full discussion of the VOC should talk about ergative languages, including the gradual rejection of Marantz's (1984) "deep ergative" hypothesis in favor of analyses that are compatible with VOC (for

FGT does nothing more than set the proper stage for a true Minimalism or a true eliminative functionalism, I think that it has done a lot of valuable work that it can be proud of. (I also think it is a virtue of the approach that it does not push one to offer hasty and premature pseudo-explanations, which may offer some rhetorical satisfaction but offer little in terms of discovery or prediction. But I do press this point here.)

8. Types of evidence

The sources of data for FGT research are simply the traditional ones of the field. For example, it sees the value of both targeted elicitation techniques and the study of naturally occurring data in recorded narratives or conversations. Each kind of data can be used to complete and correct the other. One can even go back and forth between the two. A naturally occurring example might be a great starting point, as a model for targeted elicitation. Conversely, it might help to ask a speaker to build a small narrative around a sentence that arises in elicitation. Anything behavior that is a product of true linguistic knowledge and use is in principle welcome.

Large scale corpus techniques are probably not especially useful in practice, simply because large corpora are not available in the range of languages that need to be considered. But you never know. The same point holds for psycholinguistic and neurolinguistic experimentation. My impression is that, in order to answer the explicitly comparative questions in (1)-(3), what we most need is first pass generative-style analyses of a wider range of languages, rather than new techniques that can mainly be applied to more familiar and accessible languages. There are hard practical problems about how to do psycholinguistic and neurolinguistic tests on languages spoken in remote areas by elderly people of a very different culture, and these may just not be worth solving, for now. But I hasten to add that questions (1)-(3) are not all there is to linguistics. Other techniques may be needed to answer other important questions, and convergence across different modes of inquiry is always to be looked for and hoped for.

Finally, FGT makes no special use of sociolinguistic and diachronic data. That may partly be a deficiency in our training. Surely these areas of inquiry are related, at least in the sense that diachronic changes have produced the range of languages we now have to study, and sociolinguistic variation probably helps to create diachronic change. In answering the question of whether crosslinguistic differences are patterned or not (question 3), looking at which languages can develop from a given language type could be particularly valuable. For example, the fact that uniformly head-initial languages like English and French have evolved from the same source as uniformly head-final languages like Hindi helps to suggest that there is a unified word order parameter. Similarly, Baker (1996) takes the fact that uniformly polysynthetic languages have evolved out of nonpolysynthetic languages in Northern Australia as evidence that there is a polysynthesis parameter, and Baker (2007) draws similar morals about agreement from comparing Makhwa with its Bantu cousins and Ecuadorian Quechua with Peruvian Quechua. We thus use some quasihistorical examples on an opportunistic basis, and there may be opportunities in the future to do this more deeply and systematically.

example, Bittner and Hale 1996). On the negative side, Warlpiri is an interesting case of a nonconfigurational language that has been fairly extensively studied, but has yielded little evidence for the VOC, at least in syntax. I for one remain uncertain just how this language fits in.

But there are limits to the potential of such inquiry too, which keep it off the top of our (my) agenda. First, it is not always necessary to know where the diversity we see came from answer questions (1)-(3) and the implications of those answers for human cognition more generally. Second, FGT accepts Chomsky's point that the most real and primary notion of language is the mental representations in the minds of individual speakers. Those typically do not represent the historical changes that led to the current language. In principle, they would be exactly the same even if very different paths of historical change led to the same primary linguistic data. In short, the history of a language is not known to the average native speaker of a language, so it can have at most an indirect role.¹⁰ Added to this is the practical problem that we have very little recorded history for any of the languages spoken outside of Eurasia, and reconstruction of syntax by internal reconstruction and the comparative method seems like a rather speculative enterprise. We do know that there has been enough historical change to get linguistic diversity in different areas of the world; therefore the universals that we observe do not seem to be simply due to inheritance from a common ancestor. That might be most of what we need really need to know to face problems (1)-(3).

8. Conclusion

Formal Generative Typology is a pragmatic and somewhat eclectic approach that is built on the idea of combining a generative-style toleration for abstractness in analysis with a typology-inspired interest in testing claims over a sample of unrelated languages. This seems to be a powerful combination, with the potential to greatly advance the quest for answers to questions about what is truly universal to natural human languages, what can vary, and whether the variation is patterned or not.

For Further Reading:

An accessible, book-length overview of this approach is Baker 2001. More specialized methodological discussions are included in Baker and McCloskey 2007 and Baker In press. Some canonical examples of the approach in action include Baker 1996, Baker 2003, Baker 2008, and Cinque 1999. For more information about the generative substrate that the approach is built on, see the references in Boeckx (this volume).

References

- Baker, Mark. 1996. *The polysynthesis parameter*. New York: Oxford University Press.
Baker, Mark. 2001. *The atoms of language*. New York: Basic Books.
Baker, Mark. 2003. *Lexical categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
Baker, Mark, and McCloskey, James. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11:285-296.

¹⁰ In contrast, some sociolinguistic variables probably *are* represented in the minds of speakers to some extent, since speakers can use and understand different styles in different social contexts. Perhaps, then, these should have a larger role in FGT (and generative linguistics more generally) than they have to date.

- Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Baker, Mark. In Press. The macroparameter in a microparametric world.
- Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11:239-251.
- Bittner, Maria, and Hale, Kenneth. 1996. Ergativity: toward a theory of a heterogeneous class. *Linguistic Inquiry* 27:531-604.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, eds. Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam, and Lasnik, Howard. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann, 506-569. Berlin: Walter de Gruyter.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chung, Sandra. 1998. *The Design of Agreement*. Chicago: University of Chicago Press.
- Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, Mass.: MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford.
- Comrie, Bernard. 1981. *Language universals and linguistic typology*. Chicago: University of Chicago Press.
- Croft, William. 2002. *Typology and universals*. Cambridge: Cambridge University Press.
- Cysouw, Michael, Albu, Mihai, and Dress, Andreas. 2007. Analyzing feature consistency using dissimilarity matrices. *Strachtypologie & Universalienforschung*.
- Dryer, Matthew. 1998. Why statistical universals are better than absolute universals. *Chicago Linguistics Society* 33:123-145.
- Dryer, Matthew. 2005. Order of subject, object, and verb. In *The world atlas of language structures*, eds. Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie, 386-389. New York: Oxford University Press.
- Greenberg, Joseph. 1963. *Universals of language*. Cambridge, Mass.: MIT Press.
- Haspelmath, Martin, Dryer, Matthew, Gil, David, and Comrie, Bernard eds. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Koopman, Hilda, and Sportiche, Dominique. 1991. The position of subjects. *Lingua* 85:211-258.
- Legate, Julie. 2008. Morphological and abstract case. *Linguistic Inquiry* 39:55-102.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Mass.: MIT Press.
- Marantz, Alec. 1997. No escape from syntax. In *University of Pennsylvania Working Papers in Linguistics*, eds. A. Dimitriadis and I. Siegel, 201-225. Philadelphia: U. Penn Department of Linguistics.
- McCloskey, James. 1996. On the scope of verb movement in Irish. *Natural Language and Linguistic Theory* 14:47-104.

- Mithun, Marianne. 1987. Is basic word order universal? In *Coherence and grounding in discourse*, ed. R Tomlin, 281-328. Amsterdam: Benjamins.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.
- Reinhart, Tanya, and Reuland, Eric. 1993. Reflexivity. *Linguistic Inquiry* 24:657-720.
- Ross, John. 1967. Constraints on variables in syntax, MIT.
- Selkirk, Elisabeth. 1982. *The syntax of words*. Cambridge, Mass.: MIT Press.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425-450.
- Sproat, Richard. 1985. Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3:173-216.
- Stewart, Osamuyimen. 2001. *The serial verb construction parameter*. New York: Garland.
- Tsujimura, Natsuko. 2006. *An introduction to Japanese linguistics*. Cambridge, Mass.: Blackwell.

Biographical Sketch: Mark Baker is a professor in the Department of Linguistics and the Center for Cognitive Science at Rutgers University. He received his Ph.D. in Linguistics from MIT in 1985, and has also taught at McGill University. He is the author of numerous technical articles, and four research monographs: *Incorporation* (1988), *The Polysynthesis Parameter* (1996), *The Lexical Categories* (2003), and *The Syntax of Agreement and Concord* (2008). He has also written one book for a more general audience, *The Atoms of Language* (2001). His primary research interest is the syntax and morphology of less-studied languages, especially Native American languages and African languages.