On the relationship of typology to theoretical syntax

MARK C. BAKER and JIM MCCLOSKEY

In these remarks on the relationship between typology and theoretical (morpho)syntax, we touch briefly on three issues: what is their relationship in practice now, what relationship should one in principle expect given the founding goals of each enterprise, and what kind of research could help connect the two fields in a more productive way in the future.

1. Assessing current practice

We see signs that work in theoretical syntax is being influenced more and more by the methodology and results of typological work. The Greenberg universals of word order (Greenberg 1963) have been incorporated into the generative theory of phrase structure for a long time, at least since Stowell’s (1981) reduction of explicit phrase structure rules to more general principles. Perhaps the most discussed formal research of the late 1990s was Cinque’s (1999) study of clause structure, which made use of the typological results of Bybee (1985) and related work; Cinque also performed his own survey of more than 500 languages. Julien (2002) is another important example of this emerging genre. In a similar spirit, Baker (1996) attempted to incorporate and explain some of Nichols’ (1986, 1992) typological observations about head marking languages into his theory of polysynthesis, and Baker’s (2003) theory of lexical categories is informed by typological studies like Stassen (1997) and Wetzer (1996).

More generally, we sampled four recent issues of Linguistic Inquiry and four of Natural Language & Linguistic Theory (two leading venues for formal syntactic work). Of the 27 articles on syntax in these issues, the mean number of languages discussed per article was 3.37 (ranging from 1 to 8 languages). We suspect that this is much higher than it would have been 25 years ago.

But there are also signs that the integration of typology and formal theory is progressing more slowly and more sporadically than one might hope. Cinque’s research has been widely admired, and widely cited, but not widely imitated.
And one gets a different impression of the 27 recent research articles if one counts how many language families are discussed in each article. Sixteen of the 27 articles focus exclusively on languages from a single family (in every case but one, Indo-European). Another nine articles discuss languages from two families—comparing a single East Asian language (Chinese, Japanese, or Korean) with English, for instance, or including a Semitic language in a sample that otherwise contains languages from various branches of Indo-European. One paper considered languages from three families, and one single paper considered languages from five families.

A journal article with strict page limits is of course not the ideal genre for presenting a large scale typology. Still, we get the strong impression that, while crosslinguistic comparison is clearly on the rise in formal work, true typology is not. Conspicuously absent in most of this work is the typologist’s vision of controlling for genetic and areal idiosyncrasies by sampling unrelated languages and language families. Rather, the methodology of comparing closely related languages, pioneered by Richard Kayne, has been much more influential in formal generative circles—partly because of its intellectual merits, but also partly (we suspect) because it is easier to do.

2. On the logical relationship between the disciplines

With this in mind, it is instructive to reconsider the founding goals of typology and formal syntactic theory, to evaluate whether the current situation is healthy and to be expected or not.

A founding goal of typological research since Greenberg has been to discover (and explain) “universals” of human language by sampling widely from among all the attested natural languages. The universals that were originally proposed were of three main types:

(i) truly universal claims, of the form “No language/every language has feature X”
(ii) statistical regularities, of the form “Languages with feature X are very (un)common”
(iii) implicational universals, of the form “If a language has X then it will also have Y”

Two classes of implicational universal can also be distinguished: absolute implicational universals (“If X, then always Y”) and statistical implicational universals (“If X, then usually Y”). This category then reduces, in a certain sense, to the first two.

These different kinds of potential empirical discovery differ in their implications for the program of formal theoretical syntax. This program typically aims to provide an understanding of what it is to be a “possible human language”, an “attainable state of the language faculty”. For this project, universals of the
On the relationship of typology to theoretical syntax

first type (to the extent that they are not attributable to external contingencies) are extremely relevant, by definition. They must either be built into the design of the theory, or the theory must be developed in such a way that it guarantees their truth.

In contrast, the relevance of universals of the second type is much less clear. Obviously languages with even the rarest features – say Hixkaryana, with its OVS word order – are possible human languages (represent attainable states of the language faculty); otherwise they could not be actual human languages. Nor is there is any evidence that we are aware of that such languages are more difficult to acquire or to use than languages that have the more common features. It follows then that nothing in a formal theory of what counts as an attainable human language should rule out the existence of these rare languages. The relevance of statistical universals to the theoretical tasks is thus obscure at best.

This conceptual assessment interacts with an empirical result which seems to emerge from current typological research: universals of the first kind are rare and usually not terribly interesting (e.g., all languages have vowels; all languages have verbs). Most of the striking results that typological research has discovered are statistical. Thus, Bickel (2007) writes that “[I]arge datasets almost invariably reveal exceptions to universals, and this […] has practically done away with notions of absolute universals and impossibilities” (see also Nichols 2007). There are many more statistical generalizations, and they concern interesting contingent properties of language, but, as a matter of logic, their relevance to a theoretician whose goals are those of generative grammar is not obvious. This then partially explains why typological concerns and results have had less influence on the practice of generative grammar than one might have hoped. It is easy to get the impression either that there are no truths of the kind that the formalist is interested in (as Bickel and Nichols suggest), or that standard typological methodologies are unlikely to uncover those truths (as generative linguists might suspect).

In the meantime, those committed to standard typological methods understandably become interested in the statistical regularities that they uncover, quite apart from what they might reveal about what is a possible human language. They naturally seek to explain those regularities in terms of forces and principles that are completely external to the theory of what it is to be a possible human language. It is Newmeyer’s (2005) position that this kind of “externalization” strategy (according to which the task of explaining statistical generalizations is external to the business of constructing a theory of possible grammars) is both logically and empirically justified. On the typological side, this also seems to be the emerging position in the “new typology” reported by Bickel (2007), which focuses on areal, historical, and anthropological factors in creating statistical patterns. In this conception, formal linguists would not
be held responsible for accounting for the statistical distribution of language
types and features. Grammatical theory and typology then tend to evolve into
two separate fields, with distinct goals and objects of inquiry, which rightly
have little to do with each other.

There are, of course, alternatives to the strict “externalist” mode of expla-
nation. Many strands of current work converge on the view that distributional
regularities often reflect the action of functional pressures of one kind or an-
other. One response to this observation is to seek to build those functional pres-
sures directly into individual grammars, as has been done in the functionalist
tradition, and more recently in functional Optimality Theory. Frequent or in-
 frequent grammar-types then reflect the probabilities of different “solutions” to
problems of optimization clumping in one or another region of the possibility-
space.

A different choice would be to resist building functional pressures directly
into grammatical theories, holding instead that functional pressures operate
over time to shape the statistical distribution of possible features and systems.
The time-span over which such pressures operate could either be historical
(shaping the evolution of particular languages) or evolutionary (shaping the
evolution of the language faculty). The former would provide an understanding
of which systems are (im)probable; the latter would provide an understanding
of which systems are (im)possible. One promising path in this direction would
be a marriage between John Hawkins’s research program and Principles and
Parameters/the Minimalist Program – a marriage that seems desirable on other
grounds as well (see again Newmeyer 2005; Polinsky & Kluender 2007 also
offer thoughts along this line).

Consider now the third possible class of universals, those that have the form
of implicational statements. To the extent that these are absolute implica-
tional universals, they have the same importance for the grammatical theorist
that the absolute universals have: they too should be built into the theory or
emerge from it as some kind of logical consequence. Moreover, typological
research gives us hope that this kind of universal might be relatively common
and involve interesting and substantive linguistic properties. If there is to be
a true partnership between grammatical theory and typology, this is where we
might expect it to be centered. It is true that to date even implicational uni-
versals seem to be mostly statistical rather than absolute, but this could be an
artifact of linguists needing to uncover additional terms in the if-clause of the
conditional. For all we know so far, statements of the form “If X and Y, then
usually W” can be turned into statements of the form “If X and Y and Z, then
always W” (where Z itself is a relatively common feature of language). On this
conception the apparent irregularities that statistical generalizations offer up re-
fect the complexity of the interactions involved, together with the limitations
(severe) on current understanding.
There are some encouraging cases where this looks to be so. Consider the classical Greenbergian discovery that SOV languages are only slightly more common than SVO languages, but that both are much more common than VSO languages. Current formal treatments suggest that VSO word order can arise when and only when a number of factors happen to come together. Heads must come before their complements (as in SVO languages), subjects must appear relatively low in the clause structure (perhaps inside VP, or at any rate not moving to the specifier position of InfP), and the verb must move into the Inf node. Suppose that each of these grammatical parameters has a 50% chance of being set in this particular way as opposed to the opposite way. (This is known to be roughly true for the head directionality parameter.) Then only about 1 in 8 languages will be a VSO language, because all three factors have to be set in one particular way for a VSO language to emerge. Any other combination of parameter settings will give either an SOV language or a SVO language (see Baker 2001: Chapter 5 for details). The prediction, then, is that only about 12.5% of languages of the world will be VSO, as compared to 50% SOV language and 38.5% SVO languages. This prediction is not far from the observed frequencies. Dryer (2005) cites 497 SVO languages, 435 SVO languages, and 85 VSO languages. In this large sample (1017 languages), the number of SVO languages is roughly four times the number of VSO languages, whereas the sum of the SVO languages and the VSO languages is roughly equal to the number of SOV languages – approximately as predicted by combining the structure of the formal analysis with some rudimentary (overly simple, no doubt) statistical expectations. If a substantial proportion of statistical universals emerge as complex absolute implicational universals in this way, then the prospects for deep and productive interaction between typology and grammatical theory are much rosier.

How likely is it that this program can be successfully carried out? The question, it seems to us, remains open. Functionally-oriented typologists tend to extrapolate from their experience to the conclusion that all is statistical, and that nothing is truly impossible to human language. Even the formally-minded pessimist will note that the parameters proposed in the earliest work on the topic (e.g., the pro-drop parameter or the configurationality parameter) have tended to shatter, on close investigation, into smaller-scale and independent “microparameters”. Relatively few new proposals of the same scale have arisen to take their place. Rather, current “parameters” tend to allow or disallow, say, Object Shift or Verb Second patterns but predict few correlations.

The optimist, however, will note that there does seem to be something true and significant about the headedness parameter, which underlies the Greenbergian word order universals, and there is no obvious reason why this should be the only parameter of its kind. Moreover, the impression that everything is possible (though perhaps unlikely) might arise in part from an insufficient
appreciation for how large the space of logical possibilities actually is.

3. Mixed methodologies for future progress

It is in the end an empirical matter whether statistical universals can be properly understood in implicational terms, and hence it is also an empirical matter how the relationship between grammatical theory and typology should be properly conceived. However, it seems clear to us that the kind of research that could best decide these questions is done too rarely under either the generative or the typological banner.

Generative linguists, when they compare languages at all, typically compare languages from a single family, or only two languages at a time. The positive side of this is that it enables them to dig deeply into the languages in question. But the obvious drawback is that they find many spurious generalizations resulting from their very small sample size, and miss many true generalizations by not looking at enough languages. Why haven’t these linguists found more interesting large or medium scale parameters? It could be that these are common enough, but too few formal linguists have spent enough time looking at a range of non-Indo-European (non-Semitic, non-East-Asian) languages to see them.

Typological linguists are unable to find complex implicational universals for the opposite reason, because they look at so many languages at once. This means that their samples inevitably contain errors, either inherited from the descriptions they draw on, or introduced by their own misreadings of those descriptions. This introduces an additional source of noise into their results, so it is unsurprising that most of their generalizations look to be statistical in nature. Probably even more significant is the fact that complex implicational universals are harder to investigate at the level of analysis that these studies tend to demand. There is an issue of combinatorial explosion here: there are many complex conditionals to consider once one imagines nesting implications or allowing factors to be conjoined. These cannot all be tested against a typological sample in a bottom-up, data-driven manner. Rather, one must have some kind of deeper conceptual analysis that can tell one what combinations might be worth testing – a kind of analysis that is more typical of the formal-generative tradition. Furthermore, the more factors that go into a complex implicational universal, the less likely it is that the descriptions of individual languages will discuss all of the relevant factors, making it harder to use a large sample. These pressures make it difficult for typological methods on their own to find these sorts of universals if they exist.

What kind of methodology might tell us whether or not language is characterized by complex implicational universals embedded in a formal theory of grammar? We would like to identify two possibilities. The first is a gold-
On the relationship of typology to theoretical syntax

standard, best-practice ideal that the field as a whole can aim toward over time. The second is a more practical approach that individual linguists might adopt to help move the field as a whole toward the ideal.

First the gold standard. There is no reason in principle why one could not combine the use of a motivated degree of abstractness in the analysis of particular languages, as is typical in formal linguistics, with an interest in sampling widely from a range of languages, as is typical in typological studies. We understand in part why many functionalist-typologists are either intrinsically suspicious of abstract analyses or find them impractical to work with. Nevertheless, we find such techniques to be well-warranted when looking in depth at the grammars of particular languages, as necessary to capture meaningful generalizations and achieve genuine descriptive depth. It then seems obvious that these techniques must have a role to play in larger-scale studies as well. We also understand in part why many formal linguists never get around to comparing a significant number of unrelated languages: the initial investment in sketching out a basic formal analysis of each new language seems so high. Nevertheless, if there is a rich and substantive universal grammar (as we still tend to believe), then the variation in unrelated languages will not in practice be so large as to make this unfeasible. Indeed, we have not found it so hard in practice to compare (say) the grammar of Mohawk with that of Mapudungun, or the grammar of Irish with that of Chamorro or Lebanese Arabic.

Functionalist-typologists might not think that there is much to gain from making use of generative abstractness, but they should be open to changing their view if this brings real (near) absolute implicational universals to light. Generative linguists might think that there is little to learn about a topic of interest (say, anaphor binding) from a superficial study of Mapudungun that cannot be learned from a deep study of (say) Greek, but they should be open to changing their view if such comparisons do bring relevant new patterns to light. Furthermore, any proponent of Universal Grammar should welcome in principle typological evidence that the properties in question truly are universal in the most obvious sense.

One can imagine, then, a kind of mixed method that could draw on the best of what both traditional approaches have to offer. The task would admittedly be huge, in that it involves doing deep generative analysis of a large sample of languages. But the task is in principle no larger than the project of mapping the human genome, which the biologists have faced. It simply needs to be approached gradually, over time, making use of collaborative efforts by communities of researchers. In other words, it is something that we could collectively aim for, even if it is not an appropriate standard for evaluating individual research contributions. Further, it is our sense that such a research program is a more practical prospect now than it has been before, in that it can now draw on a very substantial body of high-quality analytical work on a range of language
families and language types that has been built up over the past thirty years of work within the broadly generative tradition.

One of us (Baker) attempted an experiment in this sort of large scale generative typology over the last year. My detailed formal-generative investigation into agreement phenomena suggested to me that the Bantu languages that I had studied differed systematically in their agreement behavior from familiar Indo-European languages in the following two ways:

(1) The agreed-with noun phrase must c-command the agreement-bearing category in Bantu languages, whereas there is no such requirement in Indo-European languages.

(2) The agreed-with noun phrase must determine the case features of the agreement-bearing category or vice versa in Indo-European languages, whereas there is no such requirement in Bantu languages.

This led to the conjecture that these are two discrete formal parameters that largely characterize the syntax properties of agreement in particular languages. This conjecture was then tested against data from the core sample of 100 languages described in the *World Atlas of Language Structures* (Haspelmath et al. (eds.) 2005). Simplifying somewhat, the goal was to confirm or disconfirm implicational universals like the following:

(3) The tense node must be asymmetrically c-commanded by a NP (the subject) in order to agree with it in language X if and only if the verb must be asymmetrically c-commanded by a NP (the object) in order to agree with it (and similarly for each pair of agreement bearing heads).

(4) If a language does not permit the verb to agree with oblique-case subjects and objects, then it will not permit both the auxiliary verb and the main verb to show full person-number-gender agreement with the same NP (the subject) in complex tenses. (The results are reported fully in Baker in press.)

These proposed universals are abstract theoretical claims that require some structural analysis to be done on each language before they can be evaluated. The truth of (4) can only be judged, for example, if one is able to distinguish reliably between an object NP in dative case (which can be agreed-with in languages of the non-Indo-European type) and PPs with a directional adposition like English to (which cannot be agreed with in either type of language). Settling such questions takes considerable work, work of exactly the type that the methodology of generative syntax requires and encourages.

Similarly, one can get a good first pass idea about whether a language is like a Bantu language or like an Indo-European language with respect to parameter
(1) and the universal in (3) derived from it by seeing if agreement is sensitive to the relative word order of the noun phrase and the agreeing category or not. This works because c-command and surface word order are heavily correlated. But word order is only a proxy for the real notion, which is c-command, and somewhat noisy results become less noisy when one moves from the proxy to the real thing. For example, heads agree only with NPs that precede them in Kinande, whereas they agree only with NPs that follow them in Apurinã (Facundes 2000). At first glance, these look like opposite conditions, but in fact they are not. Kinande is a head-first SVO language, whereas Apurinã is a head-last language that alternates between the rare orders OSV and OVS. There is independent reason to think, then, that the basic phrase structures of Apurinã are the exact mirror image of what they are in Kinande. If that is right, then the two language-particular statements fall together as special cases of the general universal in (3). Implicational universals stated in generative terms can thus legitimately turn out as being less noisy and more absolute than versions stated in more surface-oriented terms.

Was this methodological experiment a success? Others will have to be the judge of that, but it was certainly an interesting project to undertake, and the results are at the very least promising. No counterexamples to implicational universal (4) emerged from the 108-language sample. As for the implicational universal (3), 16 languages were identified in which c-command was consistently needed for agreement, 43 languages in which it was consistently not needed, and only 2 (Berber and Nez Perce) which seemed to be counterexamples. So (3) and (4) are surprisingly good candidates for being nearly-absolute implicational universals. Furthermore, there is some relevant generative literature on both Berber and Nez Perce, and defensible alternative analyses exist under which even they may not be true counterexamples to (3). This suggests that there may in fact be more large-scale parameters – and hence more bona fide implicational universals – out there to be found once the right methodology is adopted.

On the other hand, the project at times seemed to push the limits of what was feasible. In many cases, it proved necessary to construct toy-generative analyses for interesting languages on the basis of limited information and no real expertise. It is certainly possible that this gave undue scope to wishful thinking that subsequent research will (hopefully) correct. The research would have been easier to do and more reliable if we had fuller syntactic analyses of the many languages involved, which had been worked out by actual experts of those languages. Thus, while I (Baker) am quite confident that Kinande is correctly placed in the typology, I am less confident about Apurinã, and even less confident about my interpretation of the possible counterexamples in Nez Perce. This sort of inquiry may, then, be slightly premature, even for a topic like agreement, where one can expect to find a fairly full description in any
Mark C. Baker and Jim McCloskey

decent reference grammar. How premature it is will hopefully be revealed by whether or not the rough-and-ready analyses constructed on the basis of easily available materials turn out to have been reasonably reliable in the light of future research. But the larger point is surely that the monograph stands as a kind of existence proof that research of this type is now feasible, at least in a preliminary way. What is needed to make the enterprise more reliable and less premature is that the range of languages for which reasonably deep and accurate analyses exist must be greatly expanded. In the absence of such a foundation of reliable language-particular work the typological generalizations can only emerge in obscured and incomplete form.

This brings us, then, to our suggestion for a practical research strategy that typologically-minded people can profitably pursue while we wait for our understanding of a wide range of languages to increase. We suggest that there is a “Middle Way” which will shed light on the crucial underlying issues. This Middle Way style of research would look at fewer languages than a typical typological study does, but at more languages than a typical generative study does. It would dig into the internal workings of each language to an intermediate degree, so as to cull out superficial counterexamples and identify additional factors that could impinge on the topic under study, while still leaving time to look at more than one or two languages. More concretely, we might expect followers of the Middle Way to base their research on five to ten languages that share a relevant feature but that are genetically and areally unrelated. That would greatly reduce (although not eliminate) the danger of spurious generalizations that besets formal work, while at the same time reducing (although not eliminating) the danger of errors introduced by superficiality of analysis that besets some typological work. In fact, Baker’s 108-language survey mentioned above did not change much the proposed parameters that he arrived at by thinking about a much smaller set of languages in some depth (several Bantu languages, several Indo-European languages, and a couple of others). This suggests that smaller databases may in practice be almost as reliable as larger ones for sorting out the formal possibilities, at least in areas where the space of possibilities is constrained fairly tightly by universal grammar – although there is ultimately great value in confirming any results by larger scale studies as this becomes possible.

Practice of the Middle Way could also provide a useful channel of communication between prototypical typological work and prototypical generative work, so that the two communities see their potential relevance for each other more clearly. The results of Middle Way research would probably make more sense to typologists than most generative results do; the typologists could see how to generalize that work by investigating parts of it in the larger samples that they often work with. The results of Middle Way research could also make more sense to generative linguists than most typological results do, because
On the relationship of typology to theoretical syntax

some sources of statistical noise would be reduced, leading to sharper (though less certain) results of a kind that invite theoretical reasoning. The Middle Way could be pursued either by one researcher in a multi-year research project, or by a team of researchers, each doing careful work on one or two languages with a commitment to combining their results. A few models of both kinds exist in the field.

If these potentials could be realized, the Middle Way could thus form a kind of conduit for the best results of typology to flow into formal syntactic theory and vice versa, as already seems to be true in the phonological domain (see Hyman, this volume). This could provide an escape from those sterile debates between “functionalist” (typological) and “formal” approaches to language that have too long hindered dialogue and deepening of understanding.

Received: 17 December 2005 Rutgers University
Revised: 12 January 2007 University of California at Santa Cruz

Correspondence addresses: (Baker) Department of Linguistics, Rutgers University, 18 Seminary Place, New Brunswick, NJ 08901, U.S.A.; e-mail: mabaker@ruccs.rutgers.edu; (McCloskey) Humanities Academic Services, Humanities One, University of California at Santa Cruz, Santa Cruz, CA 95064, U.S.A.; e-mail: mcclosk@ucsc.edu

References

Mark C. Baker and Jim McCloskey