DRAFT, December 2001

What is typology? — a short note

Balthasar Bickel
University of Leipzig
bickel@uni-leipzig.de

NOTE: This is a rough and informal draft putting together my personal take on what typology is and should be. I have not taken the time to engage in individual discussions with other authors, and I therefore mostly refrain from attributing views to specific theoreticians or typologists. I originally planned on working this out for publication, but my other projects kept me away from this. Recent personal discussions with students and colleagues, however, suggested to me that I should make this note at least available on my web site. For papers applying in practice the kind of reasoning presented here, see, e.g., *The syntax of experiencers in the Himalayas*, or *Predicate-level vs. clause-level linking*, or *Typology as a historical disciple*, available from my web site.

It is often assumed that the goal of typology is to define the notion 'possible human language'. This view, which I call the Universalist Typology view is shared, for example, by virtually all contributors to Bynon & Shibatani's 1995 volume *Approaches to Language Typology*, and by Moravscik in her review of this volume in *Linguistic Typology* 1 (p.105). In the following I claim that this assumption is fundamentally mistaken. To clarify the theoretical status of what is meant by 'possible human language', I argue here for a distinction between typological theory (theoretical typology) and grammatical theory (theoretical syntax and theoretical morphology) as distinct sub-disciplines of linguistics.

The goals of Universalist Typology are in irresolvable conflict with the methods of typology. The basic method of typological research is crosslinguistic surveying -- whether this involves only a dozen languages, or whether this involves a statistically sophisticated sample of several hundred languages. In either case, all findings are necessarily probabilistic, i.e. more (if the sample is big and representative) or less (if the sample is small and distorted) reliable estimates on what is typical for, i.e. preferred by human languages, and where and when, and why. Such findings reveal how likely structural patterns are in the world's languages, how they are distributed geographically, with what their occurrence is likely to corrrelate, what explains such correlations and patterns of distribution, etc. This all provides essential insight into the mechanisms of language change, into the social and cognitive forces underlying these mechanisms, and into patterns of human migration and prehistory. To establish such findings, to test them statistically, and to propose explanations of them, is the single most important contribution of typology to the rest of linguistics (and indeed, to many of its sister disciplines, from genetics to psychology to social anthropology).

By the same token, however, crosslinguistic surveying cannot in principle contribute to the definition of what is possible and what not in human languages.[1] A probabilistic statement is not, and cannot be converted into, a possibility statement. Even if our sample consisted of all languages that are currently spoken, it would still be a sample of all languages that have ever been spoken (that is, during the time that human language has had its modern characteristics). As such, it only allows probabilistic statements. If a structural pattern is found to be extremely unlikely, that does not mean it has never existed. Indeed, typological research has again and again shown that what is very unlikely still can occur. Recent examples include the discovery of syntactic ergativity in the absence of morphological ergativity (Donohue & Brown 1999), of syntactic ergativity in complementation (Bickel & Nichols 2001), or of pronoun borrowing (Thomason & Everett 2001). An earlier examples was object-before-subject constituent order (Pullum 1981). Since they involve small quantities of languages, these discoveries by no means undermine probabilistic statements about the unlikelihood of the relevant phenomena, and by the same token, they leave mostly intact the explanations of why the phenomena are unlikely and rare.

But exploring, testing, and explaining probabilities is not the same as defining what is possible for a human language and what not. What is possible for human languages is what is describable in a given descriptive framework. The definition of such a framework, and the exploration of what the framework allows to describe, i.e. predicts to be possible, is the goal of grammmatical theory, not of typological theory. Specific proposals of grammatical theory can be falsified by individual languages.[2] But they cannot be falsified by probabilistic findings from typology. And by the same token, grammatical theory cannot be *based* on probabilistic typologies. Instead, theories of grammar are best based on external evidence: on principles of philosophy and mathematics, as Chomsky would have, or on findings from neurology and psychology, as more empirically-minded linguists would have it.

Proponents of Universalist Typology sometimes claim that typology is a theory of grammar along such other theories as LFG, P&P, RRG, HPSG, etc.. Typologists usually define descriptive notions for their survey, and sometimes develop mini-theories that allow deriving such notions from general principles (cf. e.g., Dixon's A/O/S system for deriving alignment notions). Dixon has labeled this enterprise Basic Linguistic Theory (BLT). While BLT is the very foundation of typological surveying (since it defines the very objects that are surveyed, i.e., the collection of crosslinguistically applicable notions), it is not a theory of grammar. A theory of grammar might, and in an ideal world should, be able to derive BLT notions from its principles and interpret these notions in terms of its principles. Now, typologists may go further and aim at developing precisely a theory that derives BLT notions. But that enterprise is no longer typology, but, indeed, theory of grammar.

---

[1] We can certainly list what is possible, but the list will not be finite, i.e. not *defined* in a strict sense.

[2] Theories of grammar that are not falsifiable are not scientific theories, and I intend the term theory here in a strictly scientific sense.

There is one other way in which the theory of grammar could be typological in a more genuine sense. This is so if a theory of grammar directly predicts probabilities of occurrence. The one theory that at first sight comes close to this is OT, especially Stochastic OT. However, OT is a theory of distribution and predicts probabilities of occurrence. Indeed, in a sense, it is the first fully formalized typological theory. Yet by the same token, it is not a theory of grammar; rather OT *presupposes* a theory of grammar, i.e. representational formats and other descriptive tools. OT, like any typological theory, is compatible, and indeed, has been used with, many different grammatical theories – be it LFG, P&P or LDG. Whether OT (by itself, abstracting away from the specific theory it is used with) provides successful accounts of typological distribution will have to be seen. A big challenge is that typological distributions, even if they evidence universal preferences, very often also include areal skewings, and that has to my knowlegde not been addressed yet in OT. Moreover, typological distributions are historically grown distributions, and there are good reasons to assume that they are better captured by diachronic type transition probabilities rather than by synchronic or achronic type dependencies (cf. e.g, Maslova in *Linguistic Typology* 4, 2000; also Greenberg's chapter in the Bynon & Shibatani volume mentioned above, and my presentation on *Typology as a historical discipline*, available from this web site). But I am not aware of a diachronic reinterpretation of OT (although I do not see a principled reason why this should not be possible).

In general, however, it seems that none of these efforts of merging typology and the theory of grammar are conceptually very convincing. The problem is that deriving descriptive notions and structural types from first principles is fundamentally different in method and nature from surveying and explaining the distribution of these types in the world's languages. There is nothing one could gain from fusing these two tasks and indeed both suffer from attempts at doing so.

Practical implications. What does this all mean for the practice of doing typology? First of all, it refocuses the field onto its proper task: defining types (qualitative typology), surveying them statistically (quantitative typology), and explaining their distribution (theoretical typology). Second, once typology is understood as being about distributions and probabilities, it becomes self-evident that the field requires powerful statistical methods. Yet at this point our statistical methods are poorly developed; they will remain so as long as typology is not seen as *fundamentally* quantificational (and many typologists, especially those of a Universalist persuasion, indeed seem to shy away from statistics). Third, the second half of the 20th century has suffered from an over-emphasis on theory development at the expense of knowledge generation. As long as typologists see their ultimate goal in the construction of theories that define what is a possible human languages, the 21st century risks the same bias. But if typology emancipates itself as a sub-discipline with its own method and agenda, the 21st century promises a vast and fast expansion of our knowledge of the ways languages are structured and why they are so structured.