Linguistic diversity and universals

Bickel, Balthasar

DOI: https://doi.org/10.1017/CBO9781139342872.006

5

Linguistic diversity and universals

Balthasar Bickel

5.1 Introduction

The phenomenon of human language appears in two opposite manifestations: on the one hand, the phenomenon manifests itself in thousands of individual languages, dialects, and sociolects, and these come with differences that are often so obvious and easy to notice (e.g., different sounds, words, ways of saying things) that people can debate about them and deploy them for marking social or national boundaries. On the other hand, language manifests itself as a universal phenomenon that is shared by our entire species, processed by a brain that is in many fundamental aspects identical for all members of the species, learned efficiently by every infant in the first few years, and used for universally comparable purposes in communication.

Understanding how these two opposite manifestations can be reconciled has been at the core of linguistic research for a long time, and the pendulum has swung back and forth between emphasizing one or the other side. In the first half of the twentieth century, it was taken for granted that diversity is enormous and constantly evolving, while in the second half, linguistics has sent a strong message that despite all diversity, languages are built on a single universal grammar. From outside linguistics, it is difficult to find one's way in this opposition, especially so because the opposition is often fraught with ideological assertions, social factions in the scientific community, and a bewildering proliferation of theories that try to sell the right way of studying language (such as "the Minimalist Program," "Lexical-Functional Grammar," "Cognitive Grammar," and dozens more; see Heine and Narrog 2012 for a recent collection).

But for anthropology, just like for any other discipline in the neighborhood of linguistics, the dual nature of language as both a diverse and universal phenomenon is of key importance: for example, if we want to

I thank Sabine Stoll, Larry Hyman, Martin Haspelmath and Nick Enfield for helpful comments on an earlier draft.

understand the role that language plays in shaping society and ideas, we need to know where variation is played out in language and where it is constrained, and what forces determine universal trends and patterns. The present chapter aims to chart out the relevant issues, trying to stay away from the thickets of ideologies and competing theories in order to highlight what I see as the more fundamental questions. I begin by illustrating the various ways in which languages differ from each other (section 5.2) and then raise the question in what sense one can talk about universals despite this apparent diversity and variation (section 5.3). The concluding section (5.4) summarizes the current state of the art in universals research.

5.2 Diversity

One of the few undisputed universals of language is, ironically, that no language is completely stable across generations. This leads to substantial diversification in virtually all dimensions. In the following I review the extent of the diversity that we know from today's languages in the world. I do this by walking through what are traditionally assumed to be the core domains of language – phonology, morphology, syntax, and semantics – but I exclude issues in pragmatics as these are amply discussed in other chapters of this handbook. What is most striking about linguistic diversity is that it is not limited to the details of what individual languages have within each of these domains. Languages even differ from each other in whether or not they have the defining units of some of the domains: there are languages without phonemes (the key ingredient of phonology), without words (the key ingredient of morphology), without recursive phrase structure, parts-of-speech categories, or grammatical relations (the key ingredients of syntax).

Before we proceed, a word needs to be said about the units of variation. What varies and diversifies is basically individual structures and their organization. In (non-applied) linguistics the term "language" is mostly meant as no more than a rough identifier, a label, for some variants of structures that are sampled (e.g., a set of specific relative clause structures, identified as English structures) and there is no theoretical interest associated with this labeling. The term "language" sometimes does enter theoretical analysis (e.g., in historical linguistics), but then it is understood as a gradient notion: a set of variants of structures that are relatively similar to each other and therefore understandable or relatively easy to learn to understand across users of each variant (then called "dialects," "sociolects"). For the relevant gradient, there are no standard metrics, and therefore, it is not possible to count the number of languages beyond ballpark estimates (which cluster about 6,000–7,000). There are of course yet other uses of the term "language," most notably ideological ones, as when two groups of speakers start to declare that their sets of variants are

distinct languages (as has happened in the case of Serbian and Croatian) or when large and widely diverse sets of structural variants are declared to belong to a single language (as is the case of Chinese).

In this chapter, I use the term "language" in the sense of a structure identifier, and so I apply it to any kind of variant set of structures, including dialects, sociolects, idiolects.

5.2.1 Phonology

The most obvious way in which languages differ is in the nature of their sounds. There are languages with only very few sound distinctions, the known minimum being eleven in Rotokas, a language of Papua New Guinea (Firchow and Firchow 1969). The largest inventory is known from languages in Southern Africa. !Xõõ (from the !Ui-Taa family), for example, has (under one analysis) thirty-seven consonants plus a complex set of eighty-three click sounds and a rich inventory of forty-four vowel qualities (including, e.g., pharyngealized or breathy vowels) and four tones (Traill 1985). Differences like these are abundant,[1] but diversity goes far beyond this, and affects even the very elements that constitute phonology.

The single most important element here is the phoneme: the smallest meaning-differentiating units, which do not themselves bear meaning but recombine to create meaningful expressions. Such units are constitutive in all known spoken languages, and also in most sign languages.[2] For example, in Israeli Sign Language, the signs meaning "send" and "scold" differ only in the location of the sign ("send" being signed near the signer's torso, "scold" near the face), the signs meaning "profit" and "restraint" only in the shape of the hand during signing, etc. This is equivalent to minimal pairs like *send* vs. *lend* in English which differ only in the shape of the tongue in the initial sound. But not all sign languages use phonemes. Sandler *et al.* (2011) demonstrate that although it is operative as a fully fledged language in the third generation, the Al-Sayyid Bedouin Sign Language lacks minimal pairs and stable phonemes. Each sign is basically a holistic unit, although young third-generation signers are now beginning to develop phonemic structure, perhaps self-organized in response to increased conventionalization and a growing lexicon size (also see Sandler *et al.*, Chapter 10, this volume). While no spoken language has been demonstrated to lack phonemic structure, some languages come close (Blevins 2012): given sufficiently rich sound inventories, a spoken language could well create its lexicon directly from meaning-bearing segments, dispensing with phonemes.

For other key elements of phonology the case is clearer: for example, there are languages without syllables, and languages without words. First, consider syllables. In most languages, syllables are the minimal unit that regulate (among other things) the possible distribution of consonants and vowels when these segments are joined together. Mandarin Chinese, for example,

allows twenty-one consonants at the beginning of such units but only *n* and *ŋ* at their end. A comparable differentiation is found in Gokana (Niger-Congo, Nigeria; Hyman 1983, 2011), but here it holds not for syllables but for larger, word-like units (stems plus suffixes): in initial position, any consonant of the language is allowed, but later in the unit, only a subset: e.g., *gɔɔmáá* 'cowrie' and *zaari* 'scatter' are well-formed, but not, say, * *gɔɔzaa*. These constraints cannot be stated in terms of syllables because if one were to split, say, *zaari* into *zaa.ri*, one couldn't specify where the constraint holds (as both *z* and *r* would be syllable-initial); if one were to split into *zaar.i*, this would create a CVVC (consonant-vowel-vowel-consonant) syllable that would not occur on its own. Similar issues arise for other phonological rules in Gokana: unlike in other languages, they do not "work" in terms of syllables (but see section 5.3.1 for further discussion).

The unit "word" can also be absent in a language. Words normally constrain the application of certain rules or constraints in phonology. For example, in English the assignment of main stress is constrained by phonological words. Each word has exactly one main stress, and everything that belongs to the same stress group is within the phonological word ("PW"), no matter how this is usually written, thus: [$_{PW}$ *an'argument*], [$_{PW}$ *'take it*] etc. The same unit also constrains for example where the voiced fricative *zh* can appear: only internally, never at the beginning of words (although *zh* can appear at the beginning of syllables, as in *luxurious*). The larger-than-syllable unit that we encountered in Gokana above is also an instance of phonological word:[3] it constrains the distribution of consonants and other features (e.g., there can be at most two different tones within this unit), and it also has a maximum length (four vowels and two consonants and no more; Hyman 1983, 2011). While most languages deploy such domains (indeed often in multiple versions: Bickel *et al.* 2009), there are languages for which there is no phonological evidence for words at all. A case in point is Vietnamese (Schiering *et al.* 2010). In this language, phonological rules and constraints all operate on either the syllable or on sequences of two or three syllables, but it does not matter whether these sequences are lexical units or syntactic phrases: e.g., an expression like *hoa hồng* does not phonologically differentiate between a single-lexeme reading 'rose' and a phrasal reading 'pink flower'. This is different from most languages, where such contrasts play out for example in terms of phonological rules that assign one stress to each word (compare double-stressed *black board* with single-stressed *blackboard*).

5.2.2 Morphology

Languages also differ widely in the type and range of morphological processes that they use for expressing various grammatical categories. Most languages realize categories by segmental affixes, but many languages in Africa rely on tonal distinctions, e.g., Kinyarwanda (Niger-Congo, Rwanda)

differentiates between moods by tone alone (indicative *mukora* 'we work', subjunctive *múkora* 'that we work', relative *mukorá* 'which we work (at)'; Overdulve 1987). The range of categories that are expressed also varies widely. Some languages express only tense morphologically (English *work* vs. *worked*) while others add various distinctions depending on whether the speaker has witnessed an event or not (e.g., in Turkish), is familiar with it or not (e.g., in Tibetan), whether the event should be thought of as a single whole or in the way it unfolded over time (e.g., in Russian), etc. Some languages allow intricate combinations here: Chechen, for example, requires speakers to decide about the internal temporal constituency of events (e.g., whether the event repeated itself or took a long time) even when at the same time declaring that one hasn't witnessed the event (Molochieva 2010). The combinatorial potential in some languages is impressive. In Chintang (Sino-Tibetan, Nepal), for example, we counted in a corpus over 1,800 distinct verb forms, expressing a large variety of categories in complex and variable ways (Stoll *et al.* 2012).

Like in phonology, languages also vary in whether or not, or to what extent, they have the very building blocks of morphology: words, now in the sense not of phonological rule domains, but in the sense of the primary building blocks of syntax. One usually posits this type of word ("grammatical words") in order to capture the fact that some sequences of meaningful elements ("morphemes") form tight units that cannot be altered by syntax. For example, we cannot insert another word between the stem *work* and the suffix *ed* in the word form *worked*. Some languages tolerate limited insertions. In Chintang, for example, some clitics (but not full words) can be intercalated in the middle of stems (e.g., *lak=ta-luse* 's/he danced', based on the stem *laklus-* 'dance', the past tense suffix *-e* and the focus clitic *=ta*, Bickel *et al.* 2007). In other languages, however, there is no limit to such insertions, and then there is no evidence for grammatical words (and hence, no morphology). This is so in Vietnamese, for which we already noted that there are no phonological words either (Nhàn 1984, Schiering *et al.* 2010). For example, strings of syllables can be interrupted by syllables that express another morpheme. It does not matter for this whether the strings themselves constitute an unanalyzable morpheme or a sequence of morphemes. For example, one can insert words like *với* 'and' or *không* 'not' inside polysyllabic strings consisting of a single morpheme like *cà phê* 'coffee' (a French loan) just like in bi-morphemic creations like *đo đỏ* 'reddish' (from *đỏ* 'red'), cf. *cà với phê* 'coffee and the like', *đo không đỏ* 'not reddish'. Even verbs can occur in the middle of lexical units, e.g., the verb *xây* 'build' in the middle of the bisyllabic noun *nhà cửa* 'house' (*Tôi xây nhà xây cửa* 'I build a house').

5.2.3 Syntax

The most obvious ways in which languages differ with regard to syntax is in their range of rules and constructions: some languages have elaborate

rules of case government (Russian), some have reduced versions (English), and some have no trace of it (Vietnamese). Another variable concerns constraints on interpretation: for example, some languages impose constraints on reference in clause-combining: in an English sentence like *she smashed the bottle to the ground and broke*, the syntax demands that we identify the second (unexpressed) referent with the agent of the first clause, against all situational odds (and so we'd rather not drop the second referent and say *... and it broke*). Some languages don't have such a constraint and referential expressions can be freely dropped (e.g., in Chinese or in Chintang). Another well-established variable is the ordering of words and phrases: e.g., some languages place possessors before, some after the head; others allow both, depending on the construction (as in English *my brother's house* vs. *the house of my brother*) or on what one wants to put emphasis on.

But as in the other domains surveyed so far, variation goes beyond these relatively obvious variables and affects the very core of syntax. I briefly survey three such issues. First, for many languages, if one wants to describe how words are put together, one needs to assume recursive stacking of phrases, e.g., the noun phrases (NPs) in [$_{NP}$[$_{NP}$[$_{NP}$*John's*] *mother's*] *car*]. But some languages impose severe limits here, so that there is no descriptive need, and hence no empirical evidence, for recursion. With regard to possession constructions, most ancient and some modern Indo-European languages, for example, frequently use structures with adjective derivations (Russian *mamina kniga* 'mother's book', where *mamin-* 'belonging to mother' is an adjective) that block recursive stacking (* *Ivanova mamina kniga* 'Ivan's mother's book', Witzlack-Makarevich, p.c.). Pirahã (Brazil; Everett 2005) allows only one NP inside another (*xipoógi hoáoíi* 'Xipoogi's shotgun'), but not two (**kó'oí hoagí kai* 'Kooi's son's daughter'). Similar issues arise with clause embedding, such as in: *I told you that* [*I was there*]. Many languages avoid this in favor of juxtaposition: *I told you that. I was there* (Mithun 1984). There is ample evidence that recursion is a historical variable, gradually developing from juxtaposition over time (Deutscher 2007, Viti 2012).

A second area of fundamental variation concerns the role of parts of speech in syntax. In most languages, syntax is to a significant extent driven by parts-of-speech (and related) categories in the lexicon, i.e., one can predict much syntactic behavior from lexical properties (in some theories in fact almost all syntactic behavior). For example, in English, the lexical unit *see* is a verb and its use predicts the creation (technically called a "projection") of a verb phrase (VP) used for predication: [$_{VP}$ [$_{V}$ *sees*] $_{NP}$]. The unit *tree* is a noun and this predicts projection of a noun phrase used for reference: [$_{NP}$ *a* [$_{N}$ *tree*]]. But not all languages work this way. Some languages use syntactic phrases that are defined independently of the lexicon. In Kharia (Austroasiatic, India; Peterson 2011 and p.c.), for example, the syntactic functions of predication and reference are formally

distinguished by dedicated series of clitics that mark off the relevant phrases. These phrases are not projected from the lexicon. In fact they are totally independent of lexical choices and can be filled by any word. The full meaning results compositionally from a fairly abstract lexical content and the clitics: in a referential phrase (RefP), for example, the proper name *Aʔghrom* refers to a town, but in a predication phrase (PredP) it denotes an event: [$_{\text{PredP}}$ *Aʔghrom=ki*] 'it became (came to be called) Aghrom', with middle voice past *=ki*. Similarly, a word like *kayom* 'talk' identifies an event as an object when used in a referential phrase: [$_{\text{RefP}}$ *u kayom*] *onḍor=ki* '(someone) overheard this (*u*) talk'. But the same word expresses the event itself when used in a predication phrase: [$_{\text{PredP}}$ *kayom=ki*], 'it became talk', i.e., 'someone talked'.[4]

A third area of variation concerns grammatical relations, i.e., constrained sets of arguments. In most languages, if one wants to state what is possible in syntax and what not, one needs to refer to specific sets of arguments. For example, the constraint on reference in English clause combination that we noted above is best captured by the set comprising the sole argument of intransitive verbs ("S") and the most agent-like argument of transitive verbs ("A"). The same set defines the range of arguments that occur directly before the verb in English and that can trigger third-person agreement (S in *she works*, A in *she sees me*, but not, e.g., patients of transitive verbs: **I sees her*). Cross-linguistic variation here is enormous (Bickel 2011a) because some languages arrange roles in different ways (e.g., treating S arguments like patients, not agents, a phenomenon called "ergativity"), or they define their grammatical relations not, or not only, in terms of argument roles, but in terms of reference (e.g., the argument that is most topical) or rely on different relations depending on the construction and verb used. A more fundamental question is whether some languages may lack any constraints on arguments whatsoever, i.e., where arguments and adjuncts are all treated the same way, regardless of their role, reference, or whatever. Above, we noted that languages may lack specific constraints, such as the one on reference in coordination. In addition, languages may lack fixed word-order rules, case assignment, verb agreement, etc. If all of these are lacking, there is no evidence for argument sets and grammatical relations to exist. This has been claimed for the Riau (and a few other) varieties of Indonesian (Gil 1999, 2005). Clause positions are not constrained to specific argument roles or types, and so a sentence like *ayam makhan*, literally 'chicken eat', can mean 'the chicken is eating' (with first word being an agent) just as well as 'someone is eating chicken' (with the first word being a patient). So far, no constructions have been observed in Riau that would delimit one set of arguments from all other arguments or adjunct. Relative clauses, for example, use the same syntax for all purposes: compare, e.g., *yang di-pukul Ali* [REL PAT-hit Ali] 'the one who was hit by Ali' with *yang di-cari wang* [REL PAT-look.for money] 'the one who is looking for money' (the latter would be out in

Standard Indonesian because this language has grammatical relations and these constrain the *yang*-construction to patient arguments if the verb contains the patient marker *di-*, as it does here).

5.2.4 Semantics

Substantial variation in semantics is obvious to anyone trying to translate between languages: every language has words with meanings that do not correspond to similar meanings in other languages and that need lengthy circumlocutions if we want to recast them. Such differences can even be observed between closely related languages or dialects: e.g., Zurich German has a word *schärme* denoting a place where it does not rain when it rains all around. There is no equivalent in Standard German. Every language has lexical peculiarities of this kind.

More interesting are differences that are systematic across entire lexical fields: in kinship terminology, for example, languages are well known to differ substantially in how they group people. Chintang, for example, differentiates between elder and younger siblings but for younger siblings it does not distinguish between the sexes (*-nisa* 'younger brother or sister'), and this term also includes cousins. Some languages have color words exhausting the range of visible colors, others don't (e.g., Yélî Dnye, Rossel Island; Levinson 2001) and when they fully partition the range of colors, languages differ in how they do this more substantially than has often been thought (Roberson *et al.* 2000, Roberson 2005).

Variation in semantics reaches furthest, however, when it is linked to differences in morphology and syntax. As noted above, languages differ in the range and type of categories they express by morphology (or through periphrastic constructions) and therefore in what they force speakers to attend to ("thinking for speaking" in Slobin's 1996 terms). An example of the impact of syntax comes from noun-phrase syntax: in some languages (especially in Southeast Asia and Mesoamerica), most nouns have the semantics of what are called mass nouns in English grammar: they denote "stuff," not "things." In Yukatek Maya, for example, a word like *k'éek'en* denotes any "material" related to pigs and is neutral as to what kind of reference it is used for. The concrete referential meaning is only established in the syntax through the use of what are called classifiers in combination with numerals (e.g., *'un-túul k'éek'en* 'a pig (alive)', *'um-p'éel k'éek'en* 'a whole pig (dead or alive)', *'un-shóot' k'éek'en* 'a slice of pork', etc.; Lucy 1992). A similar but potentially even more radical abstraction arises in languages that lack a noun vs. verb distinction in the lexicon (as in Kharia). In these languages, the lexical semantics of individual words abstracts away not only from the specific type of reference or predication that a word is used for but even from the difference between things and events; e.g., a word like *kayom* 'talk' is general across these ontological types. All further specification arises only from use of the word in a specific syntactic environment.

Interactions with syntax are particularly important because they suggest that languages differ not only in the kind of semantic units they employ, but also in the principles of how they combine these units to form full concepts: lexically (English *pig* entails not only relevant properties associated with pigs but also reference to an identifiable entity with a definite shape) or syntactically (e.g., Yukatek *k'éek'en* only denotes the relevant properties, leaving reference and shape concerns to syntax).

5.3 Universals

From a biological and psychological perspective the diversity we find is less surprising than one might think. As Levinson (2003, 2012) and Evans and Levinson (2009) have emphasized, human cognition is fundamentally tuned to variation: our brain is extremely flexible and open to environmental input (e.g., Jäncke 2009); its development and expression is fundamentally affected by genetic variation (e.g., Thompson *et al.* 2010); the physiology of language processing is highly variable (e.g., Bornkessel-Schlesewsky and Schlesewsky 2009); children's learning strategies are geared towards extracting key grammar information from distributional signals in whatever speech environment they happen to grow up in (e.g., Tomasello 2003).

Given this, it comes a bit as a surprise that over the past fifty years linguistics has managed to send a strong message that languages are far less diverse than meets the eye and that behind the apparent diversity there is a richly articulated universal grammar. In order to understand this message we need to distinguish two senses of "universal": the absolute (exceptionless) and the statistical (tendency-based) universal. As we will see, these concepts are more complex than they would seem (and than how they are sometimes treated in the literature). Absolute universality does not simply mean "found in all languages" and the concept is deeply intertwined with the theoretical foundations of modern linguistics. Statistical universality too cannot simply be translated "found in the vast majority of languages" and instead refers to complex models of probabilistic trends anchored in the social and physiological eco-system that languages are embedded in.

5.3.1 Absolute universals

Absolute universals are exceptionless. However, since we cannot survey all possible human languages, it is not clear how one would categorically establish that a universal has no exception.[5] More importantly, how can we actually *decide* whether a given language really has syllables or not, or whether it has recursive phrase structure or not? Unless this is clear, we cannot even remotely hope to know whether syllables or recursive phrase

structures are universal. The issue reflects a fundamental problem identified by Chomsky (1964, 1965, 1975): how can we justify any specific analysis (or even a whole grammar) among a range of possible analyses (grammars)? I illustrate the problem with one of the examples of diversity surveyed above.

We noted that Gokana lacks evidence for syllables as constitutive elements of phonology. This was based on Hyman's insight that all relevant phonological rules and constraints in Gokana can be best captured in terms of a larger unit than the syllable, the phonological word. As in quite a few other languages, this unit has a maximum size. The size can be described as a sequence of consonants and vowels that match the template CV(V)(C)(V)(V) and an additional constraint against CVVC words (which are allowed by the template but not possible in Gokana). However, as Hyman (2011) points out, one could in principle analyze Gokana also by positing syllables and thereby assume the syllable to be an absolute universal: the maximum word could then be analyzed as allowing maximally two syllables, each at most weighing in what is known as two moras (CVV or CVC). In addition one would need a constraint limiting C-final syllables to monosyllabic words (because strings like CVV.CVC or CVC.CVV are impossible). The asymmetry in the range of consonants in the first vs. second syllable can then be accounted for by positing a trochaic (strong before weak) ordering of syllables.

The issue is typical for many linguistic analyses: there are competing ways of analyzing structures, and the different analyses go together with different sets of universals. For example, it is perfectly possible to posit phonological words in Vietnamese and then to declare phonological words an absolute universal (as is done for example by Vogel 2009). If one does, the variation will no longer consist in whether or not languages have words, but in what kinds of audible effects these domains have on phonology (none in Vietnamese, many in English). Also, one can posit grammatical words in Vietnamese, or indeed any other, language: the variation would then consist in the fact that in some languages, grammatical words cannot be interrupted by free morphemes while in Vietnamese they can (recall the *cà với phê* 'coffee and the like' example from above). The same issues arise in syntax and semantics: one can declare recursion a universal (e.g., Nevins *et al*. 2009) and posit it for Pirahã as well, even though it has no direct syntactic effects; the variation will then concern the depth of recursion allowed by individual languages (limited to 1 in Pirahã, and some other limit elsewhere: cf. Karlsson 2010 for a survey). Also, one can posit universal parts of speech and universal grammatical relations and declare them valid for Kharia and Riau Indonesian; in return one would have to add specific rules and constraints that explain why phrase structure is not exhaustively projected from parts of speech in Kharia, and why grammatical relations range over all argument and adjunct roles in Riau but not in other languages. For semantics too, one can posit English-like noun types

for Yukatek and declare them to be absolute universals if one specifies rules that explain why Yukatek nouns behave differently from English nouns in phrases with numerals. And so on and so forth.

As a result, the range and nature of absolute universals depends entirely on the nature of the analysis and the descriptive metalanguage that one uses for a particular phenomenon. The choice between alternative sets of absolute universals is not about "more" vs. "fewer" universals; it is only a choice as to where the diversity is located. This brings us back to Chomsky's problem: how can we justify one analysis over the other? In order to develop an answer, it is useful to distinguish what Chomsky (1964, 1965) has called "levels of adequacy" (though in slightly different ways than Chomsky originally proposed). The first level of adequacy is relatively trivial: an analysis obviously needs to describe the phenomena correctly, i.e., capture all and only structures that native speakers accept as part of their language. For example, saying that tonal melodies in Gokana operate on C(V)(C)(V)(C)(V) (rather than on CV(V)(C)(V)(V)) strings would simply be incorrect, as it would predict tonal melodies that Gokana speakers wouldn't produce and would reject as not being part of Gokana.[6]

Another level that is fairly uncontroversial is that the analysis of a language should aim at using only concepts (such as "C" and "V") that are also found in other languages, i.e., concepts that are cross-linguistically applicable. This criterion of cross-linguistic adequacy means that analytical concepts must come from a universal inventory and must be licensed by a universal theory of how this inventory is structured.[7] Obeying such a constraint automatically limits the absolute range of possible languages: only those with structures that can be covered by the universal theory; everything else is ruled out as violating the universal theory. Of course it is difficult to know exactly what such a universal theory should look like. There are many proposals, and this is where the many theories of grammar differ (e.g., Lexical-Functional Grammar vs. Construction Grammar vs. Role and Reference Grammar, etc.; see Heine and Narrog 2012 for a survey). In addition, there is a fairly consolidated stock of traditional, non-formalized analytical concepts that most linguists would agree as being part of the inventory (see, e.g., Shopen 2007 for one collection) and that is sometimes called Basic Linguistic Theory (Dryer 2006, Dixon 2010–12). The criterion of cross-linguistic adequacy would rule out an analysis of Gokana in terms of lists of Gokana phonemes (e.g., as set strings like {p, t, k^j, k, kp, . . .}{a, e, i, . . .}). But both an analysis with syllables and one without would fare equally well: most theories would allow both syllables and CV-templates.[8] This is true in fact of most of the alternative analyses mentioned above. To the extent that one theory is just as good as the other (e.g., a theory with and one without requiring syllables to be universal), one set of universals is just as good as another one.

Still, is there a way to decide? Chomsky's (1964, 1965, 2004) key insight is that there are relevant criteria for this, and that these criteria are

cognitive and biological in nature, i.e., that they are about whether or not a particular theory fits with what one needs in order to explain how language is embedded in the nature of our brain/mind.[9] Specifically, analytical concepts and the set of absolute universals they imply are considered adequate only if they

(i) allow formulating grammars that are learnable by children;
(ii) are psychologically realistic;
(iii) are biologically and evolutionarily realistic;
(iv) account for the creative use of language.

The criteria themselves are not really controversial since they are relatively abstract. But heavy controversies arise whenever the criteria are made more concrete. With regard to (i), for example, it has for a long time been taken as a given that grammars are learnable only if we assume a rich set of absolute universals that are innate (e.g., Chomsky 1965, 1980). This has been challenged by empirical research on how children actually acquire language (Tomasello 2003) and by computational learnability theory (Clark and Lappin 2011, Perfors et al. 2011). Both suggest that innate universals of grammar are not needed for learnability (because children apply powerful probabilistic learning strategies).

With regard to (ii), it is often assumed in the Chomskyan tradition (e.g., Chomsky 1980) that psychological reality is already ensured if a grammar captures all generalizations that a speaker intuitively "knows" (in the sense of "has the competence of"). Under a stronger reading, criterion (ii) requires that the theory is in line with our understanding of language processing. This has been a foundational point for Lexical-Functional Grammar (Bresnan and Kaplan 1982) and similarly-minded theories (e.g., Jackendoff 2002). Processing research is advancing quickly, and it is difficult to predict which kinds of analytical concepts and thereby which representations will eventually fare best on this criterion. For example, in Chomskyan models, the properties of arguments are typically defined in phrase structure, but Bornkessel and Schlesewsky (2006) suggest on the basis of neurophysiological evidence that we need independent, non-phrase-structural dependencies for defining these properties.

With regard to (iii), it has been argued that the evolution of language requires a very specific universal, namely that grammars are modeled as the most efficient link between sign production and thinking (Chomsky 1995). But there are alternative accounts that are equally compatible with plausible evolutionary models and that do not require specific universals of grammar – for some recent perspectives see, for example, Tomasello (2008), Evans and Levinson (2009), Chater and Christiansen (2010), or Hurford (2011).

Finally, with regard to (iv), it has long been taken for granted that in order to allow for creativity one needs to assume that the set of sentences in a language is infinite and that infinitude can only be ensured by taking

recursive phrase structure as an absolute universal (allowing stacks like [*she said* [*that he knew* [*that she said* [*that. . .*]]]]). But Pullum and Scholz (2010) note that creativity does not require the possibility of infinitude (as witnessed by the creativity in, say, Haiku poetry) and that, even if infinitude were needed for creativity, it could be achieved without recursive phrase structure (e.g., by iteration with a Kleene star, or by constrained combination of constructions and template matching) and, conversely, recursive phrase structure alone does not guarantee infinitude. And of course, hierarchical or symmetrical structure does not by itself entail recursion. Criterion (iv) is therefore not very useful in distinguishing the adequacy of theories.

The overall conclusion from this is that if one wants to know the best candidates of well-motivated absolute universals, one needs to look at the latest developments in psychology and biology. Note that this implies that once we have the relevant nonlinguistic evidence, a universal can be considered demonstrated. No individual language or set of data could falsify such a universal, because, as we have seen, any counterexample can be removed by reanalyzing the data so as to fit the universal (e.g., by reanalyzing Gokana with syllables, if syllables were demonstrated to be universal on the basis of psychological or biological evidence). The problem is of course that psychology and biology do not easily give clear evidence for or against specific analytical concepts in linguistics (words, syllables, nouns vs. verbs, etc.) because our understanding of how grammar is embedded in the brain is still limited and what evidence there is typically allows multiple interpretations. This limits the practical use of Chomsky's approach to criteria of adequacy. The standard way out is to simply *assume* a universal and then to see how far one gets (e.g., Chomsky 2004, 2010).[10] The risk is of course that assumptions (e.g., about recursive phrase structure as a universal) quickly become dogmas (e.g., when they are taught to students), and many linguists therefore reject this research strategy.

But there is a completely different perspective on language that bears more promise for deciding between analyses and how one allocates the balance between universals and diversity. To see this we need to move beyond static representations and investigate another fundamental dimension of human language: its dynamics, i.e., the way representations change and diversify. This dynamics defines another criterion of adequacy: a good analysis should also be *historically realistic*, i.e., match with what we know about how languages change over time. More than the others, this criterion is utilitarian and systematically favors theories that minimize universals and maximize the potential for diversity. We prefer analyses that maximize the detection of pathways of change and general patterns in development over time. From this point of view, some analyses clearly seem to fare better than others. For example, if the existence of words is taken to be variable rather than an absolute universal, we can see strong

patterns in historical development (e.g., noting that the existence and importance of phonological words is a surprisingly stable family-specific trait: Bickel *et al.* 2009; for many similar examples, see Blevins 2004). Likewise, if we take the existence of phonemes as a variable rather than a universal, we can gain insight into the emergence of phonological patterning when the lexicon grows and becomes ever more conventionalized in a newly established language (Sandler *et al.* 2011, and Chapter 10, this volume). And for syntax, if we take recursive phrase structure to be a variable rather than an absolute universal, we can study fundamental changes in the syntax of many languages (e.g., Deutscher 2007, Sampson *et al.* 2009, Viti 2012).

5.3.2 Statistical universals

The criterion of historical realism is foundational for research on statistical as opposed to absolute universals because statistical universals are fundamentally historical in nature. The basic idea of statistical universals is this: when languages change over time, these historical processes are influenced by a large array of factors and because these factors are in highly dynamic competition, universal effects manifest themselves only statistically, never categorically. Let us explore this idea a bit more closely.

The relevant factors underlying statistical universals draw from the same psychological and biological issues that also matter to the theory of absolute universals, and it is no surprise therefore that research on statistical universals is also labeled "cognitive" or "biological" (e.g., Givón 2001, 2010, Croft 2003).[11] But the sense is a different one. For statistical universals, psychological and biological facts define causal theories that predict how languages tend to develop over time. A classic example comes from word-order research: Hawkins (1994, 2004) proposes that the way languages are processed by the brain causes languages to favor certain structures over others, for example "harmonic" over "disharmonic" patterns. Harmonic patterns are ones where nested phrase structures are oriented in the same way, e.g., in English, the head (constitutive element) of noun phrases is initial (*the development* in *the development of science*) and so is the head of verb phrases (*study* in *study the development*); in Turkish or Hindi, both heads are final.[12] This is indeed what we find in most languages, but there are exceptions (e.g., Chinese, which puts verbs before objects like English, but has head-final noun phrases like *kēxué de fāzhǎn* 'the development of science', where *fāzhǎn* means 'development'). Rather than trying to reanalyze the exceptions so as to fit an absolute universal, in research on statistical universals, the exceptions are taken at face value and the relevant universal is taken to be probabilistic. Why?

The reason is that language change is not only affected by psychological (or biological) demands such as ease of processing. The dynamics of languages is also deeply embedded in the social eco-system within which

speakers communicate: we tend to adjust our ways of speaking to how others speak. Words and grammars can become fashion and fall out of fashion again for reasons that have nothing to do with how language is processed or represented in the brain. For example, English loan words have become fashionable in many languages around the world (sometimes triggering cultural discomfort or even outright political protest, sometimes not). Similarly for grammar: some time during the great migrations after the fall of the Roman Empire it became increasingly fashionable in Europe to coin new expressions of the perfect tense, built on an auxiliary verb meaning 'have', e.g., *I have eaten* or Italian *ho mangiato* (Heine and Kuteva 2006). Effects of this kind pervade the world: whenever speakers learn another language, they tend to adapt their grammar to this other language, a process known as "(areal) diffusion." And there are many reasons for learning other languages: for example, speakers migrate to another place, they want to establish trade or other relations, or they give up their ancestral language. There are thousands of opportunities for diffusion.

A prominent example is again word order. This aspect of syntax is particularly prone to diffusion (e.g., Thomason and Kaufman 1988, Johanson 1992, Dryer 1992, Nettle 1999). As Heine (2008) shows, in most cases the relevant process consists in turning an alternative order pattern that was previously special (less frequent, carrying overtones) into the normal way of speaking so as to minimize the difference between languages in contact (as if a special construction like *Beans I like* were to become the norm in English). As a result of many such processes over hundreds and often thousands of years, specific word-order patterns spread over large geographical regions. For example, statistical analysis suggests that the odds for a linguistic family of showing a diachronically significant preference towards object-after-verb (VO) order is on average 5.4 times higher in Europe, Africa, and the coastal area around the Pacific than in the interior parts of Eurasia, Australia, and the Americas (and the difference is statistically highly significant).[13]

Given such strong effects of historically and geographically contingent events, universals of historical change cannot be expected to be categorical and must be statistical instead. However, the strong effects of historical contingencies also mean the statistics cannot simply consist in counting what is most frequent in the world. For example, looking at the raw figures given in the *World Atlas of Language Structures* (Dryer 2005), one is tempted to hypothesize a statistical universal against putting the verb first in a clause, i.e., languages like Welsh, Maa (the language of the Maasai), Tagalog, or Yukatek would seem exceptional and dispreferred (Baker 2001), perhaps because they would be more difficult to process or require exceedingly complex representations. However, such a conclusion is not warranted. The current distribution could just as well result from the accidents of history. For example, if Insular Celtic had spread throughout Europe and beyond, or if it hadn't been Bantu but Nilotic speakers (say, the ancestors of

the Maasai) who expanded into the better part of the Sub-Saharan regions, verb-initial languages might be far more common now.

For establishing statistical universals it has therefore become standard (since at least Dryer 1989 and Nichols 1992) to control for historical contingencies like areal diffusion. For the example of harmonic word-order patterns mentioned above, prospects look good: all analyses so far that control for historical contingencies reveal strong worldwide preferences for harmonic head placement (e.g., Dryer 1992, Haspelmath and Siegmund 2006, Bickel 2011b, each using different methods; but also see Dunn *et al.* 2011 for dissent and Bickel in press for further discussion).

Unlike in the case of absolute universals, establishing the validity of statistical universals is not an issue of theory choice. It is an empirical question. As such it requires appropriate statistical methodology just like empirical questions do in other sciences. Settling on the right methods here is of course not trivial. Various options are currently debated, mostly concerned with the best ways in which we can estimate diachronic trends and probabilities of language change on the basis of synchronic data, and with how we can in fact mine synchronic data for trends without missing important signals from variation.[14]

For a better understanding of statistical universals, one issue in this debate is particularly important, however: universal effects are often quite weak (perhaps not as weak as the statistical signals of certain particles in physics, but nevertheless weak enough to require large datasets). The reason is this: given the many contingencies that influence language change, it is easily possible that languages end up in a situation that appears to contradict a universal principle. Chinese is a case in point, since it places heads initially in the verb phrase, but finally in the noun phrase. And this has been so for generations, not impeding the acquisition or use of the language. Unlike absolute universals, statistical universals do not declare dispreferred structures unlearnable or unusable. Instead, they state that, given the choice between two structures (say between a harmonic and a disharmonic one), there is a very slight preference in language change for speakers to (unconsciously) choose the one structure that is better in line with the universal (the harmonic one) and make it the dominant structure in language use. This can happen either by keeping a structure if it already is in line with the universal or by changing it so as to become in line with the universal. The result and the statistical signal is the same. However, what is critical in this is that the relevant choices are given at all. Choices can arise from many different sources – for example, through developing a new structure from scratch, by borrowing it in language contact, or through fashion-driven expansion of the use of specialized constructions (e.g., one that puts objects before verbs even though the other order would be more common), etc. But, importantly, there is no need for any such choices to arise in the first place, and then the universal cannot exert any effect. Moreover, an individual may select a universally preferred variant in usage, but unless such a usage

spreads and becomes a social norm in the community,[15] the selection may remain a one-off phenomenon and the universal may have no effect either. Averaged over thousands of cases of language change, therefore, a universal may leave only small statistical signals.

Not all universal effects are weak, however. For some patterns, the pressure from psychological and biological factors seems to be so strong as to yield universal effects that are virtually never canceled out by other factors. For example, Dryer (2011) finds only one out of nearly 1,000 languages that have no overt signal that would mark the difference between a statement and a question. There seems to be heavy pressure from efficient processing in communication that favors overt signaling of the distinction. Or take the case of syllables discussed above: there seems to be a very strong trend towards organizing the distribution of consonants (and vowels) in terms of syllables and Gokana seems quite exceptional. A likely cause of this trend is grounded in aspects of speech rhythm, and constraints here appear to be very strong. Note that any insights into such factors, indeed any research on them, is impossible if syllables (or question vs. statement distinctions, for that matter) are declared absolute universals and the whole discussion is locked in debates of adequacy. As a result, one would lose an important avenue for studying how language is embedded into its speakers.

Current research has established many hypotheses on statistical universals.[16] However, only a small fraction has received detailed testing. Also, many hypotheses may turn out to be spurious correlations (like the infamous correlation between the declining number of pirates in the world and global warming), because obviously any statistical pattern is only as good as the explanatory theory behind it. Where they are worked out in any detail, the relevant theories are mostly grounded in language processing (e.g., Hawkins 1994, 2004) or in principles of communication (e.g., economy, iconicity, and markedness; Croft 2003 for extensive review), conceptualization (e.g., optimizing the partition of perceptual spaces; Regier *et al.* 2007) or social structuration (adapting languages to social structures; Evans 2003). What is shared by all theories is the idea that languages tend to develop in such a way as to fit into the natural and social eco-system of speakers: that they are easy to process, that they map easily to patterns in nonlinguistic cognition, and that they match the social and communicative needs of speakers.

This fundamental and intrinsic connection to other disciplines makes research on statistical universals particularly rewarding: instead of exploring the natural and social basis of human language by means of theoretical working assumptions (as is typical for absolute universals), we can now explore this basis empirically and in close partnership with other disciplines. For example, it becomes possible to take neurophysiological findings on principles of language processing and test the statistical footprints that such principles might leave in the historical development of languages. One study along these lines (Bornkessel-Schlesewsky *et al.* 2008)

currently explores the effects that certain measurable preferences in the processing of clause-initial noun phrases across languages (namely the preference for interpreting these agents) have on the way grammatical relations develop in languages worldwide (namely in such a way as to group agents with intransitive subjects, avoiding what is called ergativity).

What facilitates such undertakings is that work on statistical universals relies on the same mathematical toolkit that is standardly used in other sciences, from data-mining techniques to regression modeling. While research on absolute universals often proceeds along strictly theory-bound argumentation in arcane formalisms, statistical universals are explored in the same theory–hypothesis–statistics triangle that characterizes most sciences.

5.4 Summary

On the face of it, languages differ from each other in every aspect of their phonology, grammar, and lexicon. When linguists nevertheless speak of universals of language they mean one of two things: absolute or statistical universals.

Understood as "absolute," universals are whatever is posited as a necessary ingredient of the metalanguage that we need for describing and analyzing languages. This ranges from trivial notions of linearization (one sentence precedes another one) to sophisticated and highly controversial ideas like recursive phrase structure. Thus, the right answer to a question about absolute universals is: "It depends on the metalanguage you use!" The evidence for or against a specific metalanguage is hard to get within linguistics alone since metalanguages are hopelessly underdetermined by the data. In response, evidence is often sought in the extent to which a given metalanguage fits criteria or psychological and biological adequacy. But assessing this fit is difficult and controversial. As a result, proponents of absolute universals often proceed by assuming universals as working hypotheses and then explore how well these hypotheses allow them to analyze languages. A central problem of this approach is that working assumptions easily become dogmas.

Understood as "statistical," universals are probabilistic inequalities stating that it is more likely for languages to develop over time into some preferred state than to develop away from this state. Hypotheses on preferred states are in turn grounded in causal theories of how the nature of the human brain, the nature of societies and of our communicative needs determine language change. A well-established (though not uncontroversial) example is the preference for harmonic word-order patterns, which is grounded in processing ease. However, any such universal preference is constantly confounded by historical contingencies, above all by diffusion of structures across languages. Diffusion is mostly motivated not by

choosing what might best fit our nature, but simply by what happens to be around and fashionable at a given time and place. In order to detect clear signals among all these confounding factors, statistical universals can thus normally only be tested against large-scale global databases. Developing these has become a prime issue for many linguists, especially also because we are currently witnessing an unprecedented loss of linguistic diversity in the wake of globalization.

Notes

1. For a survey, see Maddieson (2011).
2. It has become a convention to extend phonological terminology to sign languages despite the etymology in the Greek word for "sound." This move is motivated by the rich parallelism between the formal grammar of signed and spoken languages (e.g., Sandler and Lillo-Martin 2006 for review).
3. More accurately, a specific subtype of the phonological word, technically called a prosodic stem. The difference is irrelevant for current purposes.
4. English allows some flexibility too (cf. *to talk* and *a talk*), but unlike in Kharia, this involves lexical derivation between noun and verb. Evidence: the meaning effects are not predictable from the syntax, e.g., the semantic relation between *to talk* and *a talk* is very different from that between *to shovel* and *a shovel*. In Kharia, these relations are determined by the rules of syntax (Peterson 2011).
5. Piantadosi and Gibson (2013) have recently shown that given the number of languages we know, we normally cannot even *estimate the probability* that a universal can be reasonably assumed to be exceptionless. The number of languages we know is indeed very small: assuming that a language stays the same for maximally about 1,000 years, that on average there have been at least 5,000 languages in the world, and that human language emerged at least 100,000 years ago, there would have been at least half a million languages so far – perhaps many more! The total sample we know now, i.e., that we normally think of when talking about "all languages," makes up less than 1% of this.
6. What I call here the first level of adequacy corresponds to "descriptive adequacy" in Chomsky (1965). Chomsky (1964) also speaks about "observational" adequacy, which only requires that there is an analysis for every string in a corpus and, unlike descriptive adequacy, does not require that analyses capture all underlying regularities and native speaker intuitions about them. The distinction is not relevant for our purposes.
7. See Chomsky (1975: 81) for an early explication of this, and Haspelmath (2010) vs. Newmeyer (2010) for recent debate. Controversies mostly arise when terms are very coarse, as for example with a term like "noun," because such terms hide much (in fact, too much) variation across languages (Bickel 2007, in press).

8. In fact, if formulated as a regular expression like `\bCV([BDG]|V{1,3}|(V?([BDG]V{1,2})?))\b`, all formal grammars (in the sense of formal language theory) would have to accept the template analysis because regular expressions are necessarily a subpart of any formal grammar (Chomsky 1975).
9. Chomsky (1965) analyzes these criteria as requests for "explanatory adequacy," emphasizing the idea that they explain why we find only those structures that are licensed by a given universal theory.
10. From this perspective, absolute universals might as well be called "theoretical universals." This would highlight the fact that absolute universals are justified solely by assumptions about the best universal theory of the analytical concepts we need. But the term is also misleading because, as we will see, statistical universals are also based on theories.
11. Unlike in research on absolute universals, there is no single father figure for this research tradition, although perhaps the work of Greenberg – especially his 1963 article and the monumental four-volume oeuvre from 1978 – can be said to have played one of the most decisive roles in developing the current state of the field. For state-of-the-art surveys, see a 2007 special issue of the journal *Linguistic Typology* or the *Oxford Handbook of Linguistic Typology* (Song 2011).
12. Correlations between the properties of two structures (here, the linear organization of verb phrases and noun phrases) are traditionally called "implicational universals" and stated in the form "If the head is initial in the verb phrases, then it is also initial in the noun phrase." The terminology is misleading because the universal is a statistical correlation, not a logical implication (Dryer 1997, Cysouw 2003, Bickel in press).
13. For the analysis I applied the methods proposed in Bickel (2011b, 2013) to the data collected by Dryer (2005). The geographical regions are defined as in Nichols and Bickel (2009).
14. A glimpse of the current state of the art can be obtained from the 2007 and 2011 special issues of the journal *Linguistic Typology*.
15. For detailed discussion of the complex mechanisms behind this, see Enfield, this volume, Chapter 3.
16. The *Universals Archive* at http://typo.uni-konstanz.de/archive (Plank and Filimonova 2000) catalogues over 2,000 hypotheses that have been proposed so far (with or without good empirical support).

References

Baker, Mark C. 2001. *The Atoms of Language: The Mind's Hidden Rules of Grammar*. Oxford: Oxford University Press.

Bickel, Balthasar. 2007. Typology in the 21st Century: Major current developments. *Linguistic Typology* 11: 239–51.

2011a. Grammatical Relations Typology. In *The Oxford Handbook of Language Typology*, ed. Jae Jung Song, 399–444. Oxford: Oxford University Press.

2011b. Statistical Modeling of Language Universals. *Linguistic Typology* 15: 401–14.

2013. Distributional Biases in Language Families. In *Language Typology and Historical Contingency: Studies in Honor of Johanna Nichols*, ed. Balthasar Bickel, Lenore A. Grenoble, David A. Peterson, and Alan Timberlake, 415–44. Amsterdam: Benjamins.

In press. Distributional Typology: Statistical inquiries into the dynamics of linguistic diversity. In *The Oxford Handbook of Linguistic Analysis*, 2nd edition, ed. Bernd Heine and Heiko Narrog. Oxford: Oxford University Press [pre-print available at www.comparativelinguistics.uzh.ch/bickel/publications].

Bickel, Balthasar, Goma Banjade, Martin Gaenszle, Elena Lieven, Netra Paudyal, Ichchha P. Rai, Manoj Rai, Novel K. Rai, and Sabine Stoll. 2007. Free Prefix Ordering in Chintang. *Language* 83: 43–73.

Bickel, Balthasar, Kristine Hildebrandt and René Schiering. 2009. The Distribution of Phonological Word Domains: A probabilistic typology. In *Phonological Domains: Universals and Deviations*, ed. Janet Grijzenhout and Bariş Kabak, 47–75. Berlin: Mouton de Gruyter.

Blevins, Juliette. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. New York: Cambridge University Press.

2012. Duality of Patterning: Absolute universal or statistical tendency? *Language and Cognition* 4: 275–96.

Bornkessel, Ina, and Matthias Schlesewsky. 2006. Generalized Semantic Roles and Syntactic Templates: A new framework for language comprehension. In *Semantic Role Universals and Argument Linking: Theoretical, Typological and Psycholinguistic Perspectives*, ed. Ina Bornkessel, Matthias Schlesewsky, Bernard Comrie and Angela D. Friederici, 327–53. Berlin: Walter de Gruyter.

Bornkessel-Schlesewsky, Ina, Kamal Kumar Choudhary, Alena Witzlack-Makarevich, and Balthasar Bickel. 2008. Bridging the Gap Between Processing Preferences and Typological Distributions: Initial evidence from the online comprehension of control constructions in Hindi. In *Scales* (Linguistische ArbeitsBerichte 86), ed. Andrej Malchukov and Marc Richards, 397–436. Leipzig: Institut für Linguistik [www.uni-leipzig.de/~asw/lab/lab86/LAB86_Bornkessel_et_al.pdf].

Bornkessel-Schlesewsky, Ina, and Matthias Schlesewsky. 2009. *Processing Syntax and Morphology: A Neurocognitive Perspective*. Oxford: Oxford University Press.

Bresnan, Joan, and Ronald M. Kaplan. 1982. Grammars as Mental Representations of Language. In *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, xvii–liii. Cambridge, MA: MIT Press.

Chater, Nick, and Morten H. Christiansen. 2010. Language Acquisition Meets Language Evolution. *Cognitive Science* 34: 1131–57.

Chomsky, Noam. 1964. The Logical Basis of Linguistic Theory. In *Proceedings of the 9th International Congress of Linguists, Cambridge, Mass., August 27–31, 1962*, ed. Horace G. Lunt, 914–1008. The Hague: Mouton.

1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

1975. *The Logical Structure of Linguistic Theory*. New York: Plenum Press [first circulated in 1955].

1980. *Rules and Representations*. Oxford: Blackwell.

1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

2004. Beyond Explanatory Adequacy. In *Structures and Beyond: The Cartography of Syntactic Structure*, ed. Adriana Belletti, 104–31. Oxford: Oxford University Press.

2010. Some Simple Evo Devo Theses: How true might they be for language? In *The Evolution of Human Language: Biolinguistic Perspectives*, ed. Richard K. Larson, Viviane Déprez, and Hiroko Yamakido, 45–62. Cambridge: Cambridge University Press.

Clark, Alexander S., and Shalom Lappin. 2011. *Linguistic Nativism and the Poverty of the Stimulus*. Malden, MA: Wiley-Blackwell.

Croft, William. 2003. *Typology and Universals*. 2nd ed. Cambridge: Cambridge University Press.

Cysouw, Michael. 2003. Against Implicational Universals. *Linguistic Typology* 7: 89–110.

Deutscher, Guy. 2007. *Syntactic Change in Akkadian: The Evolution of Sentential Complementation*. Oxford: Oxford University Press.

Dixon, R. M. W. 2010–12. *Basic Linguistic Theory*. 3 vols. Oxford: Oxford University Press.

Dryer, Matthew S. 1989. Large Linguistic Areas and Language Sampling. *Studies in Language* 13: 257–92.

1992. The Greenbergian Word Order Correlations. *Language* 68: 81–138.

1997. Why Statistical Universals Are Better Than Absolute Universals. In *Papers from the 33rd Annual Meeting of the Chicago Linguistic Society*, 123–45.

2005. Order of Subject, Object, and Verb. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie, 330–41. Oxford: Oxford University Press.

2006. Descriptive Theories, Explanatory Theories, and Basic Linguistic Theory. In *Catching Language: Issues in Grammar Writing*, ed. Felix Ameka, Alan Dench, and Nicholas Evans, Berlin: Mouton de Gruyter.

2011. Polar Questions. In *The World Atlas of Language Structures Online*, ed. Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library [http://wals.info/].

Dunn, Michael J., Simon J. Greenhill, Stephen C. Levinson, and Russell D. Gray. 2011. Evolved Structure of Language Shows Lineage-Specific Trends in Word-Order Universals. *Nature* 473: 79–82.

Evans, Nicholas. 2003. Context, Culture, and Structuration in the Languages of Australia. *Annual Review of Anthropology* 32: 13–40.

Evans, Nicholas and Stephen C. Levinson. 2009. The Myth of Language Universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429–48.

Everett, Daniel L. 2005. Cultural Constraints on Grammar and Cognition in Pirahã: Another look at the design features of human language. *Current Anthropology* 46(4): 621–46.

Firchow, Irwin, and Jacqueline Firchow. 1969. An Abbreviated Phoneme Inventory. *Anthropological Linguistics* 11: 271–6.

Gil, David. 1999. Riau Indonesian as a Pivotless Language. In *Tipologija i teorija jazyka*, ed. Ekatarina V. Rakhilina and Jakov G. Testelec, 187–211. Moscow: Jazyk Russkoj Kultury.

2005. Word Order Without Syntactic Categories: How Riau Indonesian does it. In *Verb First: On the Syntax of Verb-Initial Languages*, ed. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 243–64. Amsterdam: Benjamins.

Givón, Talmy. 2001. *Syntax*. Amsterdam: Benjamins.

2010. The Adaptive Approach to Grammar. In *The Oxford Handbook of Linguistic Analysis*, ed. Bernd Heine and Heiko Narrog, 27–49. Oxford: Oxford University Press.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In *Universals of Language*, ed. Joseph H. Greenberg, 73–113. Cambridge, MA: MIT Press.

ed. 1978. *Universals of Human Language*. Stanford, CA: Stanford University Press.

Haspelmath, Martin. 2010. Comparative Concepts and Descriptive Categories: Consequences for language description and typology. *Language* 86: 663–87.

Haspelmath, Martin, and Sven Siegmund. 2006. Simulating the Replication of Some of Greenberg's Word Order Generalizations. *Linguistic Typology* 10: 74–82.

Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Heine, Bernd. 2008. Contact-Induced Word Order Change Without Word Order Change. In *Language Contact and Contact Languages*, ed. Peter Siemund and Noemi Kintana, 33–60. Amsterdam: Benjamins.

Heine, Bernd, and Tania Kuteva. 2006. *The Changing Languages of Europe*. Oxford: Oxford University Press.

Heine, Bernd, and Heiko Narrog, eds. 2012. *The Oxford Handbook of Linguistic Analysis*. 2nd ed. Oxford: Oxford University Press.

Hurford, James. 2011. Linguistics From an Evolutionary Point of View. In *Handbook of the Philosophy of Science: Linguistics*, ed. Ruth M. Kempson, Tim Fernando, and Nicholas Asher, 473–98. Amsterdam: Elsevier.

Hyman, Larry M. 1983. Are There Syllables in Gokana? In *Current Approaches to African Linguistics*, ed. Jonathan Kaye, Hilda Koopman, and Dominique Sportiche, Vol. 2, 171–9. Dordrecht: Foris.

2011. Does Gokana Really Have No Syllables? Or: what's so great about being universal? *Phonology* 28: 55–85.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jäncke, Lutz. 2009. The Plastic Brain. *Restorative Neurology and Neuroscience* 27: 521–38.

Johanson, Lars. 1992. *Strukturelle Faktoren in türkischen Sprachkontakten*. Stuttgart: Steiner.

Karlsson, Fred. 2010. Syntactic Recursion and Iteration. In Harry van der Hulst (ed.), *Recursion and Human Language*, 43–67. Berlin: De Gruyter Mouton.

Levinson, Stephen C. 2001. Yélî Dnye and the Theory of Basic Color Terms. *Journal of Linguistic Anthropology* 10. 3–55.

2003. Language and Mind: Let's get the issues straight. In *Language in Mind: Advances in the Study of Language and Thought*, ed. Dedre Gentner and Susan Goldin-Meadow, 25–46. Cambridge, MA: MIT Press.

2012. The Original Sin of Cognitive Science. *Topics in Cognitive Science* 4: 1–8.

Lucy, John A. 1992. *Grammatical Categories and Cognition*. Cambridge: Cambridge University Press.

Maddieson, Ian. 2011. Consonant and Vowel Inventories. In *The World Atlas of Language Structures Online*, ed. Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library [http://wals.info/].

Mithun, Marianne. 1984. How to Avoid Subordination. In *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, 493–509.

Molochieva, Zarina. 2010. *Tense, Aspect, and Mood in Chechen*. Leipzig: University of Leipzig dissertation.

Nettle, Daniel. 1999. *Linguistic Diversity*. Oxford: Oxford University Press.

Nevins, Andrew, David Pesetsky, and Cilene Rodrigues. 2009. Pirahã Exceptionality: A reassessment. *Language* 85: 355–404.

Newmeyer, Frederick J. 2010. On Comparative Concepts and Descriptive Categories: A reply to Haspelmath. *Language* 86: 688–95.

Nhàn, Ngô Thanh. 1984. *The Syllabeme and Patterns of Word Formation in Vietnamese*. New York: NYU dissertation.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: The University of Chicago Press.

Nichols, Johanna, and Balthasar Bickel. 2009. The AUTOTYP Genealogy and Geography Database: 2009 release. Electronic database.

Overdulve, C. M. 1987. *Kwiga ikinyarwanda: manuel d'apprentissage de la langue rwanda*. Kabgayi.

Perfors, Amy, Joshua B. Tenenbaum, and Terry Regier. 2011. The Learnability of Abstract Syntactic Principles. *Cognition* 118: 306–38.

Peterson, John. 2011. *A Grammar of Kharia*. Leiden: Brill.

Piantadosi, Steven T., and Edward Gibson. 2013. Quantitative Standards for Absolute Linguistic Universals. *Cognitive Science* 1–21 [DOI: 10.1111/cogs.12088].

Plank, Frans, and Elena Filimonova. 2000. The Universals Archive: A brief introduction to prospective users. *Language Typology and Universals* 53: 109–23.

Pullum, Geoffrey K., and Barbara C. Scholz. 2010. Recursion and the Infinitude Claim. In *Recursion and Human Language*, ed. Harry van der Hulst, 113–37. Berlin: De Gruyter Mouton.

Regier, Terry, Paul Kay, and Naveen Khetarpal. 2007. Color Naming Reflects Optimal Partitions of Color Space. *Proceedings of the National Academy of Sciences* 104: 1436–41.

Roberson, Debi. 2005. Color Categories are Culturally Diverse in Cognition As Well As in Language. *Cross-cultural Research* 39: 56–71.

Roberson, Debi, Jules Davidoff, and Ian Davies. 2000. Color Categories are Not Universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General* 129: 369–398.

Sampson, Geoffrey, David Gil, and Peter Trudgill, eds. 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.

Sandler, Wendy, Mark Aronoff, Irit Meir, and Carol Padden. 2011. The Gradual Emergence of Phonological Form in a New Language. *Natural Language and Linguistic Theory* 29: 503–43.

Sandler, Wendy, and Diane Lillo-Martin. 2006. *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Schiering, René, Balthasar Bickel, and Kristine Hildebrandt. 2010. The Prosodic Word Is Not Universal, But Emergent. *Journal of Linguistics* 46: 657–709.

Shopen, Timothy, ed. 2007. *Language Typology and Syntactic Description*. 2nd ed. Cambridge: Cambridge University Press.

Slobin, Dan I. 1996. From "Thought and Language" to "Thinking for Speaking". In *Rethinking Linguistic Relativity*, ed. John J. Gumperz and Stephen C. Levinson, 70–96. Cambridge: Cambridge University Press.

Song, Jae Jung, ed. 2011. *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press.

Stoll, Sabine, Balthasar Bickel, Elena Lieven, Goma Banjade, Toya Nath Bhatta, Martin Gaenszle, Netra P. Paudyal, Judith Pettigrew, Ichchha P. Rai, Manoj Rai, and Novel Kishore Rai. 2012. Nouns and Verbs in Chintang: Children's usage and surrounding adult speech. *Journal of Child Language* 39: 284–321.

Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Thompson, Paul M., Nicholas G. Martin, and Margaret J. Wright. 2010. Imaging Genomics. *Current Opinion in Neurology* 23: 368–73.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.

Traill, Anthony. 1985. *Phonetic and Phonological Studies of !Xóõ Bushman*. Hamburg: Buske.

Viti, Carlotta. 2012. *Variation und Wandel in der Syntax der alten indogermanischen Sprachen*. Zurich: UZH Habilitationsschrift.

Vogel, Irene. 2009. Universals of Prosodic Structure. In *Universals of Language Today*, ed. Sergio Scalise, Elisabetta Magni, and Antonietta Bisetto, 59–82. Berlin: Springer.